

Der Udānavarga als Datenbankproblem¹

Katharina KUPFER

1. Formale Klassifikation

Ein elektronisch erfaßter Text, der in einer Textdatenbank gespeichert ist, bietet Möglichkeiten, die in einem herkömmlichen Katalog nicht zur Verfügung stehen, so vor allem die Möglichkeit von Verknüpfungen. Ein Fragment, das in der aktuellen Zählung der tocharischen Texte die Nummer THT 1052 trägt, soll als Beispiel dienen, um Probleme zu thematisieren, die sich bei der Erfassung eines nicht konsequent normierten Textes in einer Datenbank ergeben.²

Nach den Angaben des von der Staatsbibliothek zu Berlin erstellten Katalogs³ handelt es sich um ein Fragment der Tocharischen Handschrift aus den Turfanfunden Nummer 1052, in der Textedition von SIEG/SIEGLING (1921) unter der Nummer 418 erfaßt und als Tocharisch A klassifiziert, ein Stück aus der zweiten Expedition zur Turfanose an der Seidenstraße, gefunden in Sängim und dort vor Ort mit der Nummer 47 versehen. Insofern liegt eine eindeutige Bestimmung des Textfragments nach den Fundumständen vor, zusätzlich durch die THT-Nummer eine aktuelle Klassifikation auf Grund der Sammlung in der Berliner Staatsbibliothek. Aber auch wenn das Fragment durch diese Angaben eindeutig bestimmbar ist, so kann man es dadurch doch noch nicht inhaltlich zu anderen Texten in Beziehung setzen.



¹ Der Beitrag ist eine überarbeitete Fassung meines Vortrags. Ich danke allen Diskussionsteilnehmern für ihre Hinweise, die ich in den Text eingearbeitet habe. Für alle etwaigen Fehler bin ich allein verantwortlich.

² Die Graphik ist der vorläufigen Internet-Edition entnommen: <http://titus.fkidg1.uni-frankfurt.de/texte/tocharic/tht/10/1052rt.jpg>.

³ THT = Tocharische Handschriften aus den Turfanfunden. Für weitere Informationen s.: <http://www.sbb.spk-berlin.de/kohd>.

2. Inhaltliche Klassifikation

2.1. Textinhalt

Die Schwierigkeit der inhaltlichen Klassifikation beginnt damit, daß es innerhalb des Buddhismus zwar Schriften gibt, die überall als zentral gelten; jedoch besitzt jede Schule ihren eigenen Kanon. Wie dem Kommentar der tocharischen Textedition zu entnehmen ist, handelt es sich bei der zitierten Nummer um einen Teil des *Udānavarga*. Dies ist der Titel eines buddhistischen Werkes, das eine Darlegung der Friedfertigkeit, speziell des Mitgefühls mit anderen Lebewesen zum Inhalt hat. Die Quelle für diesen Text sind die Lehrreden des Buddha.

2.2. Überlieferte Fassungen (Auswahl)

Ausgangssprache (des Textkanons)	Zielsprache der Übersetzung	abhängige Überlieferungszeige
Pāli		
(Buddhistisches) Sanskrit	Tocharisch A	
	Tocharisch B	
	Uigurisch	
	Sogdisch	
	Tibetisch	Mongolisch
	Chinesisch	Japanisch

Wie aus der Tabelle⁴ zu ersehen ist, ist der Udānavarga in verschiedenen Sprachen überliefert, wobei im Hinblick auf das Tocharische eine Auswahl mit Schwerpunkt Zentralasien getroffen worden ist. Im Pāli-Kanon, dem Dhammapāda, lassen sich Entsprechungen zu den tocharischen Textstellen finden, aber die Zählung der Abschnitte ist eine andere. Der

⁴ Die Tabelle ist nach den Angaben von BERNHARD (1968) zusammengestellt.

Text des Udānavarga ist in buddhistischem Sanskrit nur in Bruchstücken erhalten, oft in Form von Bilinguen; so auch in der anderen tocharischen Sprache, Tocharisch B. Weiter ist eine Bilingue in Sanskrit und Uigurisch, einer Turksprache, auf uns gekommen. Darüber hinaus ist auch eine unvollständige Version in den iranischen Sprachen Sogdisch und Sakisch bezeugt. Im Tibetischen, das einer anderen buddhistischen Tradition als die bisher genannten Sprachen folgt, findet sich gleichfalls eine Übersetzung des Udānavarga. Das Tibetische ist mit seiner Version wiederum die Grundlage einer mongolischen Übersetzung geworden und eine chinesische Vorlage ist letztlich auch die Quelle für die japanische Übersetzung.

2.3. Der Sanskrit-Kanon

Es ist wissenschaftlich umstritten, ob es jemals einen Kanon der buddhistischen Schriften auf Sanskrit gegeben hat, da ein solcher Kanon nicht als ganzer überliefert ist. In Frage kommt nur das sog. buddhistische, nicht das klassische Sanskrit. Wenn man von der Existenz eines Sanskrit-Kanons ausgeht, dann gehört der Udānavarga in den Āgama, d.h. in den Abschnitt, der einem Nikāya des Suttapiṭaka im Pāli-Kanon entspricht, und zwar den fünften Āgama des Suttapiṭaka.

Tipiṭaka		
Suttapiṭaka	Vinayapiṭaka	Abhidamma
1. Nikāya
2. Nikāya ...		
3. Nikāya ...		
4. Nikāya ...		
5. Nikāya: Udānavarga		

Der Udānavarga ist nun seinerseits ein längerer Text, der in sich wiederum in Kapitel untergliedert ist. BERNHARD (1965) hat eine Sanskritfassung des Udāna rekonstruiert, die sich aus 33 Kapiteln zusammen-

setzt. Parallel zu dieser Abschnitteinteilung in Kapiteln läuft jedoch eine Zählung von Strophen, die von 1 bis 1050 geht:

5. Abschnitt: <i>Udānavarga</i> (33 Kapitel)	1050 Strophen
1. Anityavarga	Str. 1-45
2. Kāmavarga	Str. 46-65
3. Tṣṅāvarga	Str. 66-83
...	...

Dies ist eine Möglichkeit, wie die Sanskritfassung des Udānavarga ausgesehen haben könnte, es sind jedoch auch andere Textfassungen vorgeschlagen worden, die eine andere Zuordnung von Abschnitten (*vargas*) und Strophen vorsehen. Das bedeutet, daß eine Referenz auf den Sanskrittext nur dann sinnvoll ist, wenn die Textedition als bibliographischer Eintrag vorhanden ist. In einem solchen Eintrag wäre es notwendig, die Textedition⁵ in einer Kurzform zu vermerken und auf eine entsprechende Literaturdatenbank zu verweisen. In anderen buddhistischen Schulen hat sich eine abweichende Einteilung durchgesetzt, so z.B. in der tibetischen Textfassung. Hier liegt eine Strophenzählung mit nur 1005 Strophen statt 1050 vor, und die einzelnen Bücher sind statt in 5 Abschnitte in nur 3 Teile eingeteilt. Die chinesische Übersetzung des Udānavarga enthält 949 Strophen, weicht also von ihrer Vorlage um mehr als 50 Strophen ab.

2.4. Strophenzählung

In beiden tocharischen Sprachen, Tocharisch A bzw. Osttocharisch und Tocharisch B bzw. Westtocharisch, sind Textfragmente überliefert, die ein Metrum erkennen lassen und eine Strophenzählung erlauben.

⁵ In meinem Beispiel BERNHARD (1965).

Udānāṅkāra:

THT 0001 B		1. Varga (Anityavarga) des Udānavarga
= Nr. 1 (SIEG/SIEGLING 1949)		
Metrum, Silben: 21x21x18x13	Rhythmus: 8/7/6 9/9 4/5 7/6	recto Zeile 2 = Vers 8 (des Sanskrittextes)
a Zeile 2 = Vers 8		
Str. 75 d - Str. 87 a		

So findet sich in dem in Tocharisch B geschriebenen Textfragment THT 0001 B ein Kommentar zum Udānavarga überliefert, das Udānāṅkāra. In diesem Fragment wird der 1. Varga, der Anityavarga, kommentiert, und es ist ein Metrum von 21x21x18x13 Silben im Rhythmus 8/7/6, 9/9, 4/5, 7/6 zu erkennen. Von den Herausgebern SIEG/SIEGLING (1949) wird als Strophenzählung 75d-87a angegeben. Nachdem die Zählung im Sanskrittext bei BERNHARD (1965) nur die Strophen 1-45 umfaßt, liegt offensichtlich eine andere Einteilung vor. Die Entscheidung der Herausgeber liegt darin begründet, daß der tocharische Text selbst einzelne Abschnitte mittels Zahlzeichen abtrennt. Bei Sieg/Siegling ist nun zusätzlich vermerkt, daß sich in Zeile zwei der Vorderseite die Entsprechung zu Strophe acht des Sanskrittextes findet, die jedoch in der tocharischen Fassung mit der Nummer 75 gezählt ist. Allerdings ist die Bezeichnung von Vorder- und Rückseite in der Tocharologie uneinheitlich. So verwenden Sieg/Siegling *a* und *b* als Bezeichnung hierfür, aber in der Literatur finden sich auch *V* und *R* für Vorder- bzw. Rückseite sowie *r* und *v* als Abkürzungen von *recto* bzw. *verso*. Daher ist in der elektronischen Textausgabe die Zusatzinformation, welche Seite gemeint ist, über die einfache bibliographische Angabe hinaus in einer eindeutigen Kodierung notwendig.

2.5. Querverweise

Neben den bereits besprochenen, abweichenden Fassungen des Udānavarga, die in den verschiedenen Einzelsprachen vorliegen, gibt es weiterhin eine ganze Reihe von Werken, die durch Zitate auf den Udānavarga Bezug nehmen. Angesprochen wurde gerade der Kommentar Udānālāṅkāra, dessen Strophen dem Udānavarga entnommen sind, während der Prosatext von diesem unabhängig ist in dem Sinn, daß er nicht gleichfalls übernommen worden ist, sondern Erklärungen und Kommentare zu seiner Vorlage gibt. So gibt BERNHARD (1968) im zweiten Band seiner Textausgabe über Hundert Konkordanzen zu anderen Werken. Bereits erwähnt wurde der Pāli-Kanon, wo sich Entsprechungen im Dhammapada finden, aber auch innerhalb des Sanskrit selbst finden sich in den großen Epen Zitate, ebenso wie in den buddhistischen Geburtslegenden oder in Inschriften.

2.6. Stellenangaben

Folgende Möglichkeiten notwendiger Referenzen wären daher denkbar:

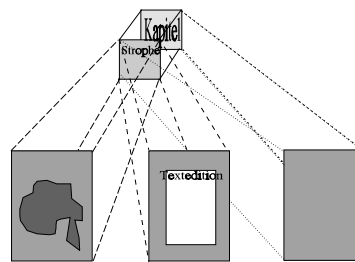
1. Literaturangabe(, Band)(, Seite)(, Zeile)
 2. Literaturangabe(, Band)(, Seite)(, Abbildung)
 3. Literaturangabe(, Band)(, Editionsnummer)
 4. Inventarnummer
 5. Werk(, Kapitel)(, Abschnitt)
 6. Strophenzahl(, Vers)
- + Kombinationen dieser Möglichkeiten

Die verschiedenen Möglichkeiten, wie zitiert werden muß, um eine eindeutige Referenz zu erlauben, hängen letztendlich von der Frage ab, wie man wirklich jede Literaturangabe korrekt benennen kann. Damit zeichnet sich aber auch ab, daß die Frage nach der Normierung des Textes auf einen Urtext hin gar nicht die entscheidende ist.

Ob ein Sanskrit-Kanon mit dem Udānavarga existiert hat und wie er genau ausgesehen hat, ist für die Behandlung in einer Datenbank nicht entscheidend. Grundsätzlich ist es notwendig, daß man einerseits den Text korrekt erfaßt und andererseits die inhaltliche Gliederung korrekt und

vollständig bestimmt hat. Eine Graphik soll die Elemente, die im vorhergehenden genannt worden sind, zeigen und ihren Zusammenhang veranschaulichen (s.u.).

Es ist dabei im Hinblick auf den Text unerheblich, ob man sich auf die graphische Edition der Textfragmente bezieht, wie ich es zu Beginn mit THT 1052 getan habe – in der Graphik durch das linke Rechteck „im Vordergrund“ symbolisiert –, oder auf eine bestimmte Textedition, wie ich es bei dem



Udānāṅkāra angedeutet habe, der von Sieg/Siegling ediert ist, oder auf eine Gliederung im überlieferten Text. Die genannten drei Formen von Text sind in der Graphik durch die verschiedenen Rechtecke symbolisiert. Wenn die elektronische Textfassung in einer eindeutigen Form⁶ vorliegt, kann man im weiteren bei solchen nicht normierten, doch bestimmten inhaltlichen Gesichtspunkten folgenden Texten nach dem Prinzip von Schlagworten verfahren. Schlagworte können grundsätzlich beliebig viele für einen Text vergeben werden, jedoch wird man die Anzahl gering halten, um eine Suche ökonomisch zu gestalten. Die einzelnen Schlagworte stellen wiederum eine Art von Gliederung für sich dar. Dieses Schlagwortprinzip ist in der Graphik durch die Linien repräsentiert, die zwischen den Elementen gezogen sind. Zu beachten ist dabei, daß sie sich von den „vorderen“ drei Elementen auf ein komplexes Gebilde „im Hintergrund“ beziehen – hier bestehend aus Kapitel- und Stropheneinteilung –, das sich auf Grund mangelnder Normativität nicht auf ein einfaches Element reduzieren läßt. Erforderlich ist eine Referenz, die parallel zur ersten Klassifikation (z.B. innerhalb des 3. Kapitels eines bestimmten Buches) einen Bezug auf unterschiedliche Segmente der zweiten Klassifikation erlaubt. Es muß also möglich sein, sich sowohl auf eine bestimmte Stro-

⁶ Auch eine hinsichtlich der Wiedergabe der Schrift eindeutige Kodierung ist notwendig.

phe (in meinem Beispiel) als auch auf mehrere oder keine Strophe zu beziehen. Notwendig ist es damit auch, daß es ein Element bzw. einen Datensatz gibt, der ausdrückt, daß keine Referenzstelle in einer bestimmten Textversion existiert.

Ziel dieses Beitrages war der Versuch zu zeigen, warum die Erfassung von nicht normierten, komplex strukturierten Texten in einer einfachen, d.h. „zweidimensionalen“ Datenbank mit Datensätzen, die auf ein Feld als Lemma festgelegt sind, nicht bzw. nur unbefriedigend lösbar ist: Es handelt sich um mindestens drei Elemente, die jeweils für sich behandelt werden, gleichzeitig jedoch auf einander bezogen werden müssen. In der Praxis wird man dieses Problem am besten dadurch lösen, daß man eine bestimmte Textausgabe als Norm setzt und sich im weiteren auf diese bezieht. Für die tägliche Arbeit ist eine solche Lösung auf jeden Fall vorzuziehen.⁷

Literatur

- BERNHARD, Franz (1965-1968): Udānavarga. Sanskrittexte aus den Turfanfunden X. Abhandlungen der Akademie der Wissenschaften in Göttingen, phil.-hist. Kl., 3. Folge, Nr. 54, Bd. I (1965), Bd. II (1968). Göttingen.
- GIPPERT, Jost (1999): „Linking methods as a basis for cross-linguistic text retrieval: problems and solutions.“ In: *Proceedings of 1999 EBTI, ECAI, SEER & PNC Joint Meeting*, Academia Sinica, Taipei, Taiwan, 18-21 January 1999, 339-346. Taipei.
- SIEG, Emil & SIEGLING, Wilhelm (1921): Tocharische Sprachreste, I. Band. Berlin / Leipzig.
- (1949): Tocharische Sprachreste, Sprache B, Heft 1. Abhandlungen der Akademie der Wissenschaften in Göttingen, 133. Göttingen.

⁷ S. hierzu – auch unter Berücksichtigung des Udānavargaproblems – GIPPERT (1999).