

Achtung!

Dies ist eine Internet-Sonderausgabe des Aufsatzes
„Indoiranistisches Text-Retrieval“
von Jost Gippert (1998).

Sie sollte nicht zitiert werden. Zitate sind der Originalausgabe in
*Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen
Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen,*
hrsg. von Bernhard Forssman und Robert Plath,
Wiesbaden: Reichert Verlag 2000, 133-145
zu entnehmen.

Attention!

This is a special internet edition of the article
“Indoiranistisches Text-Retrieval”
by Jost Gippert (1998).

It should not be quoted as such. For quotations, please refer to the original edition in
*Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen
Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen,*
hrsg. von Bernhard Forssman und Robert Plath,
Wiesbaden: Reichert Verlag 2000, 133-145.

Alle Rechte vorbehalten / All rights reserved:

Jost Gippert, Frankfurt 2011

Indoiranistisches Text-Retrieval Elektronische Bearbeitungen altiranischer und vedischer Texte

Zusammenfassung: Das im Rahmen des TITUS-Projekts angestrebte Ziel, bis zum Jahre 2000 die altiranischen und vedischen Textüberlieferungen vollständig in digitalisierter Form aufzubereiten, scheint mehr und mehr realisierbar. Über die reine Textfassung hinaus sind in jüngster Zeit, v.a. für die RV-Saṃhitā und das Avestacorpus, zusätzliche elektronische Bearbeitungen entstanden, die neuartige Analysemöglichkeiten im Hinblick auf sprachwissenschaftliche, metrische und philologische Fragestellungen eröffnen.

Karl Hoffmann drückte seine Hochachtung für Albrecht Weber einmal mit der Bemerkung aus, dieser habe in seinem Leben mehr altindische Texte herausgegeben, als andere Leute hätten lesen können. Die Leistungen Webers, aber auch anderer Editoren wie z.B. Karl Geldner oder Willem Caland können in der Tat nicht hoch genug eingeschätzt werden, insbesondere wenn man bedenkt, daß ihnen zu ihrer Zeit nicht die elektronischen Hilfsmittel zur Seite standen, die heute bei wissenschaftlichen Publikationen aller Art unentbehrlich sind. Wie sich zeigen läßt, können elektronische Verfahren aber nicht nur bei der Drucklegung wissenschaftlicher Werke hilfreich sein, sondern auch bei deren Vorbereitung, d.h. bei der Erarbeitung der zu publizierenden Erkenntnisse selbst; dies gilt gerade auch für Textzeugnisse aus alter Zeit wie die altiranischen und vedischen Überlieferungen, die Gegenstand indoiranistischer und indogermanistischer Forschung sind.

Die Bemühungen, digitale Medien für die wissenschaftliche Bearbeitung vedischer und altiranischer Texte zu nutzen, laufen bereits seit über 20 Jahren. Sie begannen mit der Eingabe der Ṛgveda-Saṃhitā sowie des Śatapatha-Brāhmaṇa (in der Mādhyamdi-nīya-Rezension), die in den 70er Jahren durch H.S. Ananthanarayana unter der Leitung von W.P. Lehmann an der University of Texas in Austin erfolgte. Mitte der 80er Jahre wurde durch S. Gippert-Fritz eine erste Erfassung des Avesta-Textes in Angriff genommen. Es versteht sich von selbst, daß beide Projekte zu ihrer Zeit mit enormen Schwierigkeiten zu kämpfen hatten, die v.a. die Verarbeitbarkeit der erforderlichen Zeicheninventare betrafen; so kann es nicht verwundern, daß bei der Bearbeitung der Ṛgveda-Saṃhitā mithilfe eines sog. „Betacode“ die Bücher 4 und 5 zunächst ohne Akzentmarkierungen eingegeben wurden und für die Eingabe des Avestatextes (nach der Geldnerschen Ausgabe) eine eigene Hilfstranskription entwickelt werden mußte.

Derartige Anfangsschwierigkeiten können, nachdem Transkriptionssysteme und sogar Originalschriften für praktisch alle Computersystemen verarbeitbar geworden sind, inzwischen als überwunden gelten, auch wenn ihre Nachwirkungen bis in die heutige Zeit reichen mögen, wo immer es um die Konvertierung „älterer“ Datenbestände zwecks einer weiteren Nutzung geht. Mit der leichteren Handhabbarkeit ist inzwischen aber auch der Datenvorrat gewaltig angewachsen. Seitdem mit einem Aufruf in „Die Sprache“ 32/2, 1986 [1987], 429 das Projekt einer indogermanistischen Textdatenbank ins Leben gerufen wurde, sind von unterschiedlichsten Beiträgern in stetiger Folge elektronische Bearbeitungen indoiranistisch relevanter Texte bereitgestellt worden, und das seinerzeit noch utopisch erscheinende Ziel, bis zum Jahre 2000 die gesamte altiranische und vedische Überlieferung in digitalisierter Form verfügbar zu haben, nimmt immer konkretere Konturen an.

Die indogermanistische Textdatenbank, die seit der III. Tagung zum Computereinsatz in der Historisch-Vergleichenden Sprachwissenschaft (Dresden 1994) unter dem Namen TITUS geführt wird („Thesaurus Indogermanischer Text- und Sprachmaterialien“), umfaßt derzeit die folgenden indoiranistisch relevanten Textbearbeitungen¹:

Das Altiranische ist durch eine komplette Bearbeitung des Avestacorporis nach der Geldnerschen Ausgabe sowie der bei Geldner nicht erfaßten Fragmente (zu darüber hinausgehenden Bearbeitungsschritten s. weiter unten) sowie durch ein altpersisches Gesamtkorpus (Eingabe unter der Leitung von G.E. Dunkel durch S. Gindro, S. Scarlata, P. Widmer, alle Zürich; Ergänzungen und Korrekturen durch G. Keydana, Münster; Überarbeitung durch J. Gippert und F.J. Martínez-García, Frankfurt) nahezu vollständig erfaßt. Auch für die mitteliranischen Überlieferungen liegen bereits umfangreiche Corpora vor.

Das Vedische ist durch die folgenden Textbearbeitungen repräsentiert²:

Ṛgveda:

- Ṛgveda-Saṃhitā* (s. dazu weiter unten);
- Ṛgveda-Khilani*: Eingabe durch C. Jordán Cólera (Zaragoza) und F.J. Martínez (Frankfurt; i.B.);
- Aitareya-Brāhmaṇa*: Eingabe durch J. Gippert und F.J. Martínez (Frankfurt);
- Kauṣṭhiki-Brāhmaṇa*: Eingabe durch K. Kupfer (Würzburg; i.V.);
- Aitareya-Upaniṣad*: Eingabe durch P. Magnone (Mailand);
- Śāṅkhāyana-Gr̥hyasūtra*: Eingabe durch Th. Zehnder (Freiburg);
- Vāsiṣṭha-Dharmasūtra*: Eingabe durch H. Falk (Berlin);
- Nighantu*: Eingabe durch M. Kobayashi (Kyōtō; i.B.);
- Nirukta*: Eingabe durch M. Kobayashi (Kyōtō; i.B.);

Sāmaveda:

- Sāmaveda-Saṃhitā*: Eingabe durch A. Sathaye (Lexington; i.B.);
- Pañcaviṃśa-Brāhmaṇa*: Eingabe durch M. Kümmel (Freiburg; i.B.);
- Mantra-Brāhmaṇa*: Eingabe durch Th. Zehnder (Freiburg);
- Jaiminīya-Brāhmaṇa*: Eingabe durch G. Ehlers (Berlin) und M. Kobayashi (Kyōtō; i.B.);
- Jaiminīya-Upaniṣad-Brāhmaṇa*: Eingabe durch M. Fujii und M. Kajihara (Kyōtō);
- Chāndogya-Upaniṣad*: Eingabe durch P. Schreiner (Zürich);
- Kena-Upaniṣad*: Eingabe durch N. Abhyankar (Albuquerque); Überarbeitung durch M. Albino und J. Narten (Erlangen);
- Gobhila-Gr̥hyasūtra*: Eingabe durch C. Jordán Cólera (Zaragoza) und F.J. Martínez (Frankfurt);
- Jaimini-Gr̥hyasūtra*: Eingabe durch Th. Zehnder (Leiden);
- Gautamīya-Dharmasūtra*: Eingabe durch N. Watase und Y. Ikari (Kyōtō);

Schwarzer Yajurveda:

- Kāṭha-Saṃhitā*: Eingabe durch S.H. Rosenfield, M. Witzel (Harvard) und Chl.H. Werba (Wien; i.B.);
- Maitrāyaṇī-Saṃhitā*: Eingabe durch V. Petr (Prag; i.B.);
- Taittirīya-Saṃhitā*: Eingabe d. M. Fushimi (Ōsaka); Korrektur durch Y. Ikari (Kyōtō);
- Kāṭhaka-Brāhmaṇa* (Überreste): Eingabe durch S.H. Rosenfield und M. Witzel (Harvard);
- Taittirīya-Brāhmaṇa*: Eingabe durch M. Fushimi (Ōsaka) und Y. Ikari (Kyōtō);
- Taittirīya-Āraṇyaka*: Eingabe durch A. Griffiths (Leiden; i.B.);
- Kāṭha-Upaniṣad*: Eingabe durch N. Abhyankar (Albuquerque); Überarbeitung durch M. Albino und J. Narten (Erlangen);

¹ Über den aktuellen Stand informiert jederzeit die Internetseite <http://titus.uni-frankfurt.de/texte/texte.htm>. Ausführliche Zwischenberichte sind: TITUS. Das Projekt eines indogermanistischen Thesaurus; in: *LDV-Forum* 12/2, 1995, 35-47 (auch: <http://titus.uni-frankfurt.de/texte/titusldv.htm>); TITUS — Von der Keilschrifttafel zur Textdatenbank; in: *Forschung Frankfurt*, 4/1995, 46-56 (auch: <http://titus.uni-frankfurt.de/texte/titusff.htm>); TITUS — Alte und neue Perspektiven eines indogermanistischen Thesaurus; in: *SIMA* 2, Prag 1997, 54-89.

² Abkürzungen: i.B. = in Bearbeitung, i.V. = in Vorbereitung.

- Taittirīya-Upaniṣad*: Eingabe durch K. Jayaraman (Auburn);
Āpastamba-Śrautasūtra: Eingabe durch M. Fushimi (Ōsaka);
Baudhāyana-Śrautasūtra: Eingabe durch M. Fushimi (Ōsaka);
Āpastamba-Dharmasūtra: Eingabe durch K. Kano und Y. Ikari (Kyōtō);
Baudhāyana-Dharmasūtra: Eingabe durch M. Fujii, M. Kajihara, T. Yagi und Y. Ikari (Kyōtō);
Vaikhānasa-Dharmasūtra: Eingabe durch H. Hikita, Y. Matsuda und Y. Ikari (Kyōtō);
Vaiṣṇava-Dharmasūtra: Eingabe durch T. Hayashi und Y. Ikari (Kyōtō);
Taittirīya-Prātiśākhya: Eingabe durch M. Fushimi (Ōsaka);
Vaikhānasa-Mantrapraśnā: Eingabe durch H. Resnick (Hrdayānanda) und M. Witzel (Harvard);
- Weißer Yajurveda:
Vājasaneyi-Saṃhitā: Eingabe durch M. Kümmel (Freiburg);
Śatapatha-Brāhmaṇa (Mādhyamdiṇīya): Eingabe unter der Leitung von W.P. Lehmann (Austin, Texas) durch H.S. Ananthanarayana (Hyderabad); Überarbeitung durch J.R. Gardner (Iowa; i.B.);
Śatapatha-Brāhmaṇa (Kāṇvīya): Eingabe durch J. Gippert (Frankfurt; i.B.);
Brhad-Araṇyaka-Upaniṣad: Eingabe durch M. Albino und J. Narten (Erlangen);
Īśa-Upaniṣad: Eingabe durch N. Abhyankar (Albuquerque); Überarbeitung durch M. Albino und J. Narten (Erlangen);
Kātyāyana-Smṛti: Eingabe durch A. Akamatsu und Y. Ikari (Kyōtō);
Yājñavalkya-Smṛti: Eingabe durch M. Tokunaga und Y. Ikari (Kyōtō);
- Atharvaveda und Postvedica:
Atharvaveda-Saṃhitā (Śaunaka): Eingabe durch V. Petr; Kollationierung durch P. Vavroušek (Prag);
Atharvaveda-Saṃhitā (Paippalāda): Eingabe durch M. Witzel (Harvard), C. Bastiaansen, A. Griffiths and Th. Zehnder (Leiden; i.B.);
Māṇḍūkya-Upaniṣad mit Gaudapādīya-Kārikā: Eingabe durch P. Magnone und A. Pelissero (Mailand);
Muṇḍaka-Upaniṣad: Eingabe durch P. Magnone (Mailand);
Praśna-Upaniṣad: Eingabe durch P. Magnone (Mailand; i.B.);
Brahmasūtra: Eingabe durch Sh. Rao (Pittsburgh);
Bṛhaspati-Smṛti: Eingabe durch Y. Ikari und A. Akamatsu (Kyōtō);
Nārada-Smṛti: Eingabe durch Y. Ikari (Kyōtō);
Parāsara-Smṛti: Eingabe durch M. Tokunaga (Kyōtō);

Darüber hinaus sind auch das epische, klassische und buddhistische Sanskrit bereits durch umfangreiche Textcorpora vertreten.

Das TITUS-Projekt versteht sich nun nicht etwa nur als ein Sammelbecken zu archivierender elektronischer Datenbestände; es soll vielmehr vorrangig zu Distribution der erarbeiteten Materialien beitragen, so daß diese weiterer wissenschaftlicher Bearbeitung zugänglich werden. Zu diesem Zwecke wird derzeit ein Server eingerichtet, auf den von außen über das internationale Datennetz zugegriffen werden kann und auf dem die Texte für ein „Retrieval“ bereitgehalten werden³. Dabei soll es nicht nur um einen einfachen Transfer von Textfiles gehen, sondern um eine Bereitstellung von Untersuchungsverfahren und -ergebnissen, die die Leistungsfähigkeit heutiger Computersysteme im Hinblick auf die Analyse von Texten ausnutzen.

Für ein elektronisches Textretrieval mit wissenschaftlicher Zielsetzung kommen in der Tat verschiedenartige Erfordernisse in Betracht, die weit über das Suchen einzelner Wortformen in einem Textganzen (als Standardfunktion aller Textverarbeitungssysteme) hinausgehen. Hierzu gehören z.B. Such- und Indizierungsfunktionen, bei denen unterschiedliche Wortformen unter einem Lemma subsumiert werden können (z.B. *ásti* und *sánti* als Präsensformen von *√as* „sein“); Suchfunktionen, die sich auf (phonologisch

³ Zusätzlich können die Texte über CDs bereitgestellt werden (ein Austausch über Disketten ist angesichts der bestehenden Datenmengen nicht mehr sinnvoll möglich). Eine erste Probe-CD liegt seit Ende Februar 1998 vor; sie kann derzeit nur den Mitgliedern der TITUS-Arbeitsgruppe zugänglich gemacht werden, da u.a. noch Urheberrechtsfragen zu klären sind.

oder morphologisch definierte, zusammenhängende oder unzusammenhängende) Wortelemente statt auf ganze Wortformen beziehen (sog. „Substringsearch“, z.B. nach *-úm* als Akk.-Endung ved. *u*-Stämme); Suchfunktionen, die das gemeinsame Auftreten mehrerer Wortformen in bestimmten Kontexten betreffen („verbundene Suche“, z.B. nach *agním* und *devám* innerhalb eines RV-Verses) u.ä. Bei den hier interessierenden Texten ist ein wissenschaftliches Retrieval darüber hinaus mit besonderen Aufgaben verbunden, wozu die Ausweitung von Suchvorgängen auf Texte gleichen Genres oder gleicher Herkunft (z.B. alle vedische Texte, alle Texte des Schwarzen Yajurveda usw.) und die „Synchronisation“ verwandter Texte (z.B. der beiden verschiedenen Rezensionen des Śatapatha-Brāhmaṇa oder des Yasnatexts mit seiner Pahlavī-Übersetzung) oder paralleler Informationen (z.B. Angaben zu handschriftlichen Lesarten, unterschiedlichen Deutungen etc.) gehören. Dies setzt u.a. voraus, daß verschiedensprachliche Elemente (z.B. Pāzend-Zitate im Avesta-Kontext) getrennt verwaltbar bleiben, gegebenenfalls auch mit Wiedergabe in unterschiedlicher Schriften; dasselbe gilt weiter auch für linguistische Zusatzinformationen (morphologisches oder syntaktisches „Tagging“, metrische Angaben etc.). Wünschenswert bleibt letztlich die gemeinsame Verwaltbarkeit textueller und graphischer Daten, so daß z.B. Handschriftenabbildungen in Koordination mit der dazugehörigen Textbearbeitung abrufbar sind. Als eine Softwarelösung, die den aufgezeigten Erfordernissen bereits heute in hohem Maße gerecht wird, hat sich für das TITUS-Projekt das von der Brigham Young University (Utah) entwickelte WordCruncher-System herausgestellt, das die Grundlage für den im Aufbau befindlichen Internet-Server bilden wird.⁴ Auch die im folgenden aufgeführten Analysebeispiele wurden mit ihm erstellt.

Der gegenwärtige Stand und die Auswertungsmöglichkeiten eines indoiranistischen Textretrievals seien nun an den beiden herausragenden Textzeugnissen, der Ṛgveda-Saṃhitā und dem Avestatext, illustriert.

Die Ṛgveda-Saṃhitā ist heute in insgesamt vier elektronischen Textfassungen verfügbar, die für das Retrieval in synoptischer Form wie in Abb. 1 dargestellt zusammengefaßt wurden. Im einzelnen sind dies:

- A) der überlieferte Text in Transkription (Grundlage: die „Texas“-Fassung von H.S. Ananthanarayana; Reformulierungen durch S. Atkins, G. Dunkel, J. Gippert; Korrekturen nach den Editionen von Th. Aufrecht und M. Müller durch F.J. Martínez García, A. Lubotsky, B. van Nooten / G. Holland und J. Gippert)⁵;
- B) dasselbe, in Einzelverse zerlegt, gegebenenfalls mit markierter Sandhi-Bereinigung an der Versgrenze (erarbeitet von J. Gippert in Zusammenarbeit mit F.J. Martínez García und A. Lubotsky);
- C) eine sandhibereinigte, weitgehend dem Padapāṭha entsprechende Fassung (erstellt von A. Lubotsky);
- D) der metrisch-restaurierte Text der Ausgabe B. van Nooten / G. Holland (Harvard Oriental Series).

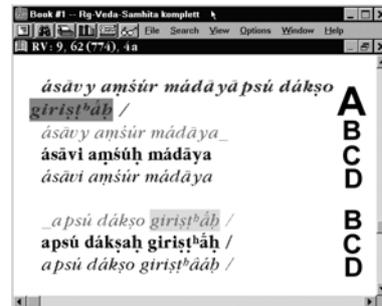


Abb. 1: Vier RV-Fassungen in Synopse

⁴ Ermöglicht wurde dies durch eine großzügige Spende seitens der Vertriebsgesellschaft (WordCruncher Co., s. <http://www.wordcruncher.com>). Leider arbeitet das Retrievalprogramm ausschließlich unter MS-Windows-Betriebssystemen; eine Lösung für andere Betriebssysteme wird noch gesucht.

⁵ Eine bei der Überarbeitung entstandene Liste von Korrekturen gegenüber der Aufrechtschen Ausgabe ist in *SIMA* 2, 1996 [1997], 66-67 abgedruckt.

Alle vier Versionen bleiben nach der Indizierung gesondert abrufbar.

Im Rahmen eines von der DFG geförderten Projektes, das der elektronischen Auswertung von Ṛgveda und Avesta gewidmet ist („AUREA“)⁶, wurden darüber hinaus folgende Zusatzinformationen erarbeitet, die über eine Synchronisierung mit dem eigentlichen Textfile verknüpfbar sind:

- Eine versweise metrische Analyse, die vorwärts und rückwärts sortiert ist (s. Abb. 2). Die dabei berücksichtigten Silbentypen und ihre (vorläufigen) Symbole⁷ sind:

Kürzen:

U: Kurzvokal in offener Silbe vor Konsonant ≠ Aspirata (z.B. *ásāvi*)

V: Kurzvokal in offener Silbe vor Vokal (*ásāvi amśúh*)

Y: Kurzvokal in offener Silbe vor Aspirata (z.B. *ab^hi*)

Längen:

–: Langvokal oder Diphthong (z.B. *mádāya*)

=: Kurzvokal mit Positionslänge (z.B. *amśúh*)

~: Kurzvokal vor *l, ch* (z.B. *ilām*)

Uneindeutig / Anceps:

_: Kurzvokal vor muta cum liquida (z.B. *pavītram*)

W: Kurzvokal plus einfacher Konsonant im Wortauslaut vor vokalischem Anlaut (z.B. *śub^hrām āndho*)

Ü: Erster Vokal innerhalb von metrischen „Zerdehnungen“ (z.B. *giriṣṭ^hāh ≈ giriṣṭ^hāh*)

X: Position am Versende

- Eine metrische Analyse in bezug auf die Position von **Wortformen** im Vers (s. Abb. 3), ausgehend von:

E) der sandhifreien Version, z.B.:

ásāvi 8:3-1'U-V *amśúh* 8:5-4'== *amśúr mádāya* 8:8-6'U-X

F) der metrisch-restaurierten Version, z.B.:

ásāvi 8:3-1'U-V *amśúr* 8:5-4'== *amśúh mádāya* 8:8-6'U-X

Bedeutung: *ásāvi* tritt im gegebenen **achtsilbigen** Vers an der Position der **1. bis 3. Silbe** mit der metrischen Struktur U-V auf etc.

- Eine metrische Analyse in bezug auf die Position von **Silben** im Vers (s. Abb. 4), ausgehend von:

G) der sandhifreien Version; Beispiel:

á-1:8'0U' -*sā*-2:8'U- -*vi* 3:8'-V' *am*-4:8'V=

-*śúh* 5:8'== *śúr má*-6:8'=U' -*dā*-7:8'U- -*ya* 8:8'-X'

H) der metrisch-restaurierten Version; Beispiel:

á-1:8'0U' -*sā*-2:8'U- -*vi* 3:8'-V' *am*-4:8'V=

-*śúr* 5:8'== *śúh má*-6:8'=U' -*dā*-7:8'U- -*ya* 8:8'-X'

Bedeutung: -*sā*- tritt im gegebenen **achtsilbigen** Vers an der Position der **2. Silbe** mit der metrischen Struktur einer Länge (-) nach einer Kürze (U) auf etc.



Abb. 2: Metrische Analyse (Verse)



Abb. 3: Metrische Analyse (Wortformen)

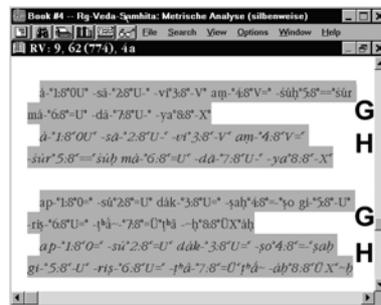


Abb. 4: Metrische Analyse (Silben)

⁶ „Avesta und Ṛgveda: Elektronische Auswertung“; das Projekt wird im Frühjahr 1998 abgeschlossen sein.

⁷ Für eine ausführliche Begründung des Analyseverfahrens sei auf den Aufsatz „Neue Wege zur sprachwissenschaftlichen Analyse der vedischen Metrik“ verwiesen (erscheint voraussichtlich 1998 in der *Ge-denschrift für Jochem Schindler*, Wien / Prag).

- Eine Sandhianalyse der Wortformen, ausgehend von
 I) der zugrundeliegenden Struktur, d.h. der sandhifreien Version;
 Beispiel: *ásāvi_a=ásāvy_a*
 (sandhifreie Version > überlieferte Form; s. Abb. 5)
 J) der überlieferten Form; Beispiel:
ásāvy_a=ásāvi_a-ásāvi_a
 (überlieferte Form > sandhifreie Version > metrisch restaurierte
 Version; s. Abb. 6).

Die Erkenntnismöglichkeiten, die derartige Bearbeitungen bieten, lassen sich z. B. an der Suche nach sprachwirklichem, d.h. metrisch realisiertem Kṣaipra-Sandhi illustrieren. Die Fragestellung lautet dabei: Wo entspricht einer Folge *i/ī* bzw. *u/ū* + Vokal der zugrundeliegenden Struktur (d.h. der sandhifreien Version) einem *y* bzw. *v* + Vokal im metrisch restaurierten Text? Es sei vorweg bemerkt, daß eine solche Analyse natürlich von den Interpretationen der einzelnen Bearbeiter abhängt, insbesondere im Falle des metrisch-restaurierten Textes, mit dessen Ausgestaltung nicht jeder in jedem Fall einverstanden sein wird⁸; jedoch ist zu erwarten, daß sich gerade durch Fragestellungen der genannten Art Widersprüche und Inkonsistenzen in den einzelnen Textfassungen aufdecken lassen und Verbesserungsvorschläge ergeben (für Beispiele s. im folgenden). Auf der Basis des derzeitigen Bearbeitungsstandes ergibt die Untersuchung nun folgende Ergebnisse:

Für von sprachwirklichem Kṣaipra-Sandhi betroffenes *i/ī* finden sich 108 Belege. Davon entfallen:

- 80 Belege auf Präpositionen bzw. Präverbien (*ab^hi*: 34; *ád^hi*: 3; *ápi*: 5; *áti*: 4; *ní*: 1; *pári*: 12; *práti*: 21)
- 7 Belege auf Konjunktionen bzw. Adverbien (*nahí*: 2; *ítí*: 4; *yádi*: 1);
- 12 Belege auf die Formel *jánitry ájījanat* in 10,134,1-6ef in dem Refrain
deví jánitry ájījanat b^hadrá jánitry ájījanat
- 9 Belege auf sonstige (verbale oder nominale) Formen (im einzelnen: *ásāmi*: 10,22(848),3a; *ásnoti*: 3,59(293),2d; *gávi*: 10,116(942),4c; *hamsi*: 8,90(699),5c; *mávi*: 10,128(954),3b; *pradísi*: 10,110(936),11c; *váṣṭi*: 10,95(921),4b; *viđ^hāti*: 1,120,1c; *yāti*: 9,15(727),1a)

Schon eine einfache Frequenzanalyse im Hinblick auf die zehn Bücher des RV (s. Abb. 9) ergibt dabei ein interessantes Bild: Gegenüber der aufgrund des Umfangs zu erwartenden Häufigkeit bei gleicher Verteilung (s. Abb. 8) zeigt sich eine starke Präferenz für vom Kṣaipra-Sandhi betroffene Formen im 10. Buch (+27), ansonsten allenfalls noch im 9. Buch (+6) während das 8. Buch eine erstaunliche Unterfrequenz aufzuweisen hat (-10; s.

Book	Frequency
-1	1
-2	2
-1	3
-5	4
-1	5
-6	6
-5	7
-10	8
6	9
27	10

Abb. 7

Book	Frequency
108	1
5x	2
6x	3
6x	4
7x	5
8x	6
9x	7
14x	8
9x	9
19x	10

Abb. 8

Book	Frequency
22	1
3	2
5	3
1	4
6	5
2	6
4	7
4	8
15	9
46	10

Abb. 9

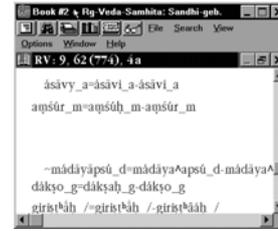


Abb. 5: Sandhianalyse (I)

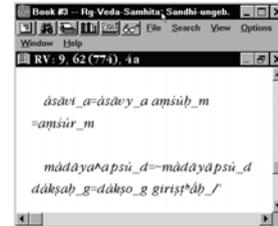


Abb. 6: Sandhianalyse (J)

⁸ Für einige Problemfälle sei auf den o.g. Beitrag zur Gedenkschrift Schindler verwiesen.

Abb. 7). Eine Gegenprobe, die sich auf das am häufigsten betroffene Wort, die Präposition $ab^h i$, bezieht, bestätigt dies allerdings nur zum Teil: hiernach ist die Kṣaipravariante $ab^h y$ zwar immer noch im 9. und 10. Buch überrepräsentiert (+6 / +3), doch kommt gerade im 9. Buch auch die nicht vom Sandhi erfaßte Variante ($ab^h i$) überdurchschnittlich häufig vor (+21; s. Tfl. 1). Zu berücksichtigen ist bei einer statistischen Auswertung weiter, daß die 12mal im 10. Buch erscheinende Formel *devī jānītry ājjanat b^hadrā jānītry ājjanat* strenggenommen nur einmal gezählt werden sollte.

	$ab^h y + V$			$ab^h i + V$		
	Belege ⁹	erwartet	Diff.	Belege ¹⁰	erwartet	Diff.
insges.	34	34x	0	99	99x	0
1	5	6x	1-	12	19x	7-
2	2	1x	1+	5	4x	1+
3	1	2x	1-	2	6x	4-
4	1	2x	1-	6	5x	1+
5	2	2x	0	4	7x	3-
6	2	2x	0	10	7x	3+
7	2	2x	0	5	8x	3-
8	2	4x	2-	15	13x	2+
9	8	2x	6+	29	8x	21+
10	9	6x	3+	11	17x	6-

Tfl. 1

Für von metrisch relevantem Kṣaipra-Sandhi betroffenes u/\bar{u} lassen sich 36 Belege anführen. Davon entfallen

⁹ Die Belege für zu lesendes $ab^h y$ vor Vokal: RV 1,80,12c; 1,120,8a; 1,135,6d; 1,156,2d; 1,179,3d; 2,17 (208),4b; 2,26 (217),1b; 3,51 (285),1b; 4,58 (354),10a; 5,19 (373),1a; 5,31 (385),5d; 6,38 (479),3b; 6,47 (488),25b; 7,103 (619),3a; 7,104 (620),2a; 8,59 (1028),6d; 8,69 (678),11d; 9,26 (738),2a; 9,47 (759),1b; 9,68 (780),8b; 9,80 (792),2a; 9,86 (798),17c; 9,99 (811),4b; 9,101 (813),16d; 9,110 (822),6b; 10,4 (830),6b; 10,8 (834),8b; 10,18 (844),11d; 10,86 (912),8d; 10,121 (947),6b; 10,123 (949),2d; 10,123 (949),6b; 10,135 (961),2c; 10,159 (985),1d.

¹⁰ RV 1,84,4c; 1,94,8b; 1,101,7c; 1,108,6c; 1,115,2b; 1,123,7a; 1,124,9b; 1,139,3c; 1,140,6b; 1,164,27b; 1,179,3b; 1,189,3b; 2,4 (195),2c; 2,16 (207),8a; 2,22 (213),2a; 2,22 (213),4f; 2,28 (219),1b; 3,1 (235),4c; 3,46 (280),4a; 4,1 (297),3a; 4,1 (297),16c; 4,6 (302),1c; 4,12 (308),1c; 4,21 (317),2d; 4,24 (320),8b; 5,3 (357),7a; 5,5 (359),4b; 5,29 (383),12b; 5,33 (387),3a; 6,15 (456),12c; 6,17 (458),8c; 6,19 (460),3d; 6,21 (462),10b; 6,22 (463),1b; 6,25 (466),5d; 6,44 (485),17c; 6,49 (490),8b; 6,49 (490),15e; 6,50 (491),15b; 7,4 (520),9c; 7,23 (539),6b; 7,83 (599),5a; 7,93 (609),8b; 7,104 (620),21b; 8,1 (621),32c; 8,2 (622),19b; 8,21 (641),16c; 8,24 (644),19c; 8,24 (644),21c; 8,25 (645),21c; 8,40 (660),4a; 8,51 (1020),8a; 8,55 (1024),1b; 8,72 (681),2a; 8,75 (684),15b; 8,79 (688),2a; 8,92 (701),31a; 8,96 (705),15c; 8,100 (709),4b; 9,1 (713),4a; 9,4 (716),7a; 9,4 (716),8a; 9,20 (732),4a; 9,21 (733),3b; 9,51 (763),5a; 9,62 (774),3b; 9,63 (775),6b; 9,63 (775),12a; 9,63 (775),29b; 9,64 (776),8b; 9,66 (778),22b; 9,67 (779),3b; 9,77 (789),4d; 9,85 (797),2c; 9,85 (797),4c; 9,85 (797),7c; 9,85 (797),8a; 9,86 (798),14c; 9,86 (798),35a; 9,86 (798),43b; 9,94 (806),5a; 9,96 (808),16b; 9,97 (809),4a; 9,97 (809),7d; 9,97 (809),50d; 9,97 (809),51d; 9,107 (819),21d; 9,108 (820),2c; 10,1 (827),3d; 10,3 (829),3b; 10,5 (831),6b; 10,29 (855),3b; 10,31 (857),3c; 10,64 (890),1d; 10,64 (890),3b; 10,68 (894),1d; 10,75 (901),2b; 10,117 (943),8b; 10,148 (974),3a.

- 27 Belege auf die Präposition *ānu*, wobei zwei in identischen Versen auftreten (1,23,23a ≈ 10,9(835),9a);
- 2 Belege (je 1) auf die Konjunktionen *nū* (10,61(887),13a) und *tū* (8,51(1020),4c);
- 3 Belege (je 1) auf die Lokativ-Plural-Formen *devēṣu* (10,121(947),8c), *mānuṣeṣu* (1,58,6a) und *vṛtrēṣu* (6,26(467),2c);
- 2 Belege (je 1) auf die Neutra *dāru* (10,146(972),4b) und *urū* (9,96(808),15d) sowie
- 2 Belege (je 1) auf die Imperativformen *sāstu* (7,55(571),5d) und *sīvyatu* (2,32(223),4c).

Die Frequenzanalyse ergibt insofern ein ähnliches Bild wie bei *iī*, als wiederum eine überdurchschnittliche Häufigkeit für das 10. Buch zu konstatieren ist (+6); für das 9. Buch gilt dies jedoch nicht (-1, s. Abb. 10). Das Ergebnis wird weiter dadurch relativiert, daß zumindest einer der Belege entfallen dürfte. Dies ist der Vers RV 6,26 (467), 2c

tvām vṛtrēṣv indra sātpatim tarutram (*tvām caṣṭe muṣṭihā gōṣu yūd^hyan //*),

der bei van Nooten / Holland wie folgt „restauriert“ erscheint:

tvām vṛtrēṣv indra sātpatim tarutram.

Richtig erscheint demgegenüber der in Oldenbergs Noten zitierte Vorschlag Grassmanns¹¹: „Eher mag *tuām vṛtrēṣu* zu lesen, *indra* zu tilgen sein“. Dies ergibt sich aus einem Vergleich mit RV 6,46 (487), 1c:

tvām vṛtrēṣv indra sātpatim nāras (*tvām kāṣṭ^hāsv ārvataḥ //*);

aus diesem regulären 12-Silbler (metrisch: *tuvām vṛtrēṣu indra sātpatim nāras*) kann der Vokativ *indra* in 6,26,2c eingedrungen sein. Die wahrscheinlichere Restauration von 6,26,2c würde dann lauten: **tuvām vṛtrēṣu {indra} sātpatim tarutram.*

Unwahrscheinlich ist auch die Auffassung von RV 1,58,6 bei van Nooten / Holland, der der Beleg für *mānuṣeṣv* zu verdanken ist. Die Strophe lautet:

dad^hūṣ tvā b^hṛgavo mānuṣeṣv ā rayim nā cāruṃ suhāvaṃ jāneb^hyaḥ /
hōtāram agne āti^him vāreṇyam mitrām nā śevaṃ divyāya jānmane //

In der Harvard-Ausgabe werden die Verse a-c als 11-silbig und lediglich d als 12-silbig aufgefaßt:

dad^hūṣ tvā · b^hṛgavo mānuṣeṣv ā 11
rayim nā cāruṃ suhāvaṃ jāneb^hyaḥ / 11
hōtāram agne āti^him vāreṇyam 11
mitrām nā śevaṃ divyāya jānmane // 12

Statt dessen lassen sich alle vier Verse problemlos als 12-Silbler herstellen, womit Übereinstimmung mit der vorhergehenden Jagatīstrophe 1,58,5 erzielt wird:

dad^hūṣ tvā · b^hṛgavo mānuṣeṣu ā 12
rayim nā cāruṃ suhāvaṃ jāneb^hyaḥ / 12
hōtāram agne āti^him vāreṇyam 12
mitrām nā śevaṃ divyāya jānmane // 12

Für diese Analyse spricht nicht zuletzt, daß *jāneb^hyaḥ* und *vāreṇyam* auch in allen ihren sonstigen Belegen „zerdehnt“ zu lesen sind (*jāneb^hyaḥ* 3mal¹², *vāreṇyam* 20mal¹³) und so auch bei van Nooten / Holland behandelt sind.

Count	Item
0+	RV
-2-	1
1+	2
-1-	3
2+	4
1+	5
-1-	6
-1-	7
0+	8
-1-	9
6+	10

Abb. 10

Count	Item
36+	RV
7+	1
1+	2
2+	3
2+	4
2+	5
2+	6
3+	7
4+	8
3+	9
6+	10

Abb. 11

Count	Item
36+	RV
5+	1
2+	2
1+	3
4+	4
3+	5
1+	6
2+	7
4+	8
2+	9
12+	10

Abb. 12

¹¹ In den Prolegomena faßte Oldenberg den Vers noch als 13-silbig auf: *tuām vṛtrēṣu || indra sātpatim tarutram.*

¹² RV: 1,7,10b; 1,55,5b; 10,156(982),4c.

¹³ RV: 1,9,5b; 1,79,8b; 1,159,5a; 3,2(236),4a; 3,12(246),1b; 3,27(261),10a; 3,34(268),8a; 3,40(274),5b; 3,62(296),6c; 3,62(296),10a; 5,8(362),1d; 5,35(389),3a; 5,39(393),2a; 6,16(457),33c; 8,1(621),19b; 8,15(635),7c; 8,27(647),1d; 8,102(711),18b; 9,65(777),29a; 10,35(861),7a

Als Beispiel für die Auswertungsmöglichkeiten im Hinblick auf metrische Strukturen¹⁴ mag eine Analyse 13silbiger Versstrukturen dienen. Hierfür finden sich auf der Grundlage der vorliegenden Bearbeitungen insgesamt 34 Belege¹⁵, wobei die Frequenzanalyse ein Überrepräsentanz im 8. Buch andeutet (+8, s. Abb. 14). Auch hier bedarf jedoch jeder Einzelfall einer genauen Überprüfung.

So ist z.B. RV 1,161,8a (s. Abb. 13) nur unter der Bedingung als 13silbig aufzufassen, daß hier die Partikel *iti* dem Kṣaipra-Sandhi unterliegt, wie in der Harvard-Ausgabe angedeutet: *idám udakám pibatéty abravītana*. Betrachtet man jedoch den weiteren Kontext, so liegt eine andere Vermutung nahe. Der Doppelvers 1,161,8ab lautet:

*idám udakám pibatéty abravītanedám vā
gḥā pibatā muñjanéjanam /.*

View	Options	Help
0+ RV		
2+ 1		
0+ 2		
-1- 3		
0+ 4		
-2- 5		
0+ 6		
0+ 7		
8+ 8		
-2- 9		
1+ 10		

Abb. 14

View	Options	Help
34x RV		
6x 1		
1x 2		
1x 3		
1x 4		
2x 5		
2x 6		
2x 7		
5x 8		
3x 9		
5x 10		

Abb. 15

View	Options	Help
34x RV		
8x 1		
1x 2		
0x 3		
1x 4		
0x 5		
2x 6		
2x 7		
13x 8		
1x 9		
6x 10		

Abb. 16

Abb. 13: Metrische Analyse von RV 1,161,8a in synchronisierten Textfenstern

¹⁴ Vgl. hierzu ausführlich den o.g. Aufsatz in der Gedenkschrift Schindler sowie „Laryngeals and Vedic Metre“ in: *Sound Law and Analogy. Papers in honor of Robert S.P. Beekes on the occasion of his 60th birthday*, ed. A. Lubotsky, Amsterdam / Atlanta, GA 1997, 63-80.

¹⁵ RV 1,88,1b; 1,120,4c; 1,127,1f; 1,127,4b; 1,138,3f; 1,155,4d; 1,161,8a; 1,164,23d; 2,43(234),2c; 4,1(297),2a; 6,10(451),1b; 6,15(456),14a; 7,32(548),12c; 7,104(620),15a; 8,1(621),26c; 8,3(623),24c; 8,21(641),4c; 8,22(642),14c; 8,35(655),13a; 8,61(670),8c; 8,97(706),13a; 8,97(706),13b; 8,97(706),13c; 8,97(706),13d; 8,97(706),14b; 8,97(706),15b; 8,100(709),6d; 9,68(780),7b; 10,50(876),2c; 10,51(877),9a; 10,66(892),1a; 10,87(913),12b; 10,114(940),4a; 10,121(947),7c.

In a könnte nun *iti* als ein sekundärer, auf die Redaktoren zurückgehender Zusatz zu tilgen sein, womit sich ein regulärer 12-Silbler ergäbe, der in seiner Struktur Vers b genau entspräche: **idám udakám píbatābravī-ana*. Die angenommene Parallelität kommt auch in Geldners Übersetzung zum Ausdruck: „Trinket ihr dies Wasser“, so sprachet ihr, „oder trinket diese Schilfschlempel!“ Diese Annahme ist sicher leichter als die von Oldenberg (Noten a.l.) vertretene, die jeder sprachwissenschaftlichen Grundlage entbehrt: „*idám udakám* wohl mit Verschleifung über das *-m*“. Erforderlich bleibt aber eine Untersuchung zur Gestaltung wörtlicher Rede im RV, die sich nicht nur auf *iti* und seinen Gebrauch, sondern gerade auch auf fehlendes *iti* erstrecken muß.

Die derzeit anstehende Etappe beim weiteren Ausbau der Datenbasis für den R̥gveda besteht in der Eingabe grammatischer Informationen (sog. „Tagging“), die ein Retrieval im Hinblick auf morphologische und syntaktische Fragestellungen (z.B. des Typs „wo verbinden sich die Formen eines gegebenen Verbums mit einem Dativ Plural?) erlauben sollen. Ziel der Bearbeitung soll zunächst die Gewinnung neuer Aufschlüsse über die relative Chronologie der Lieder, letztlich dann die Rekonstruktion eines „Urtexts“ bilden; angestrebt wird dabei eine maximale Annäherung an diejenige Gestalt der r̥gvedischen Liedtexte, die bei der Abfassung durch die Dichter bestanden haben könnte.

Auch für das Avestacorpus liegen inzwischen verschiedene Bearbeitungen vor, von denen im Zusammenspiel miteinander ein Erkenntnisgewinn zu erwarten ist. Im einzelnen handelt es sich um die folgenden Komponenten:

- der in Transkription umgesetzte Wortlaut der in der Geldnerschen Ausgabe enthaltenen Texte (Eingabe von 1985 bis 1988 durch S. Gippert-Fritz im Rahmen des DFG-Projekts Avesta-Wörterbuch; Überarbeitungen und Korrekturen durch M. de Vaan, J. Gippert u.a.);
- die nicht in Geldners Ausgabe enthaltenen Texte in vereinheitlichter Transkription (Eingabe durch S. Gippert-Fritz bzw., für das Nirangistān nach der Ausgabe Waag, durch H. Kumamoto).

Im Rahmen des AUREA-Projekts (s.o. Anm.) sind seit 1996 hinzugekommen:

- der Variantenapparat der altavestischen Textteile nach den über die Ausgaben von Geldner und Westergaard verfügbaren Informationen (s. Abb. 17);
- nicht bei Geldner bzw. Westergaard erfaßte Varianten der über Faksimile verfügbaren Handschriften J2, K5, Mf4 und P1 sowie der von Geldner nachkollationierten Hs. Br2 (verfügbar über Geldners Handexemplar seiner Ausgabe in der UB Marburg);
- der Variantenapparat der nicht-altavestischen Teile des Yasna nach den über die Ausgaben von Geldner und Westergaard verfügbaren Angaben (sämtlich durch M. de Vaan, weiterverarbeitet durch R. Gehrke).

Ein eigener, bereits abgeschlossener Bearbeitungsschritt bestand in

- der Rekonstruktion der „liturgischen“ Textgestalt; dies betrifft zum einen die Anordnung des Textes nach metrischen oder Sinnabschnitten (s. Abb. 19), zum anderen
- die Auffüllung von in den Ausgaben (und teilweise auch den Handschriften) ausgelassenen oder abgekürzt niedergeschriebenen Formeln etc. bei unterschiedlicher Markierung und getrennter Indizierung der verschiedenen

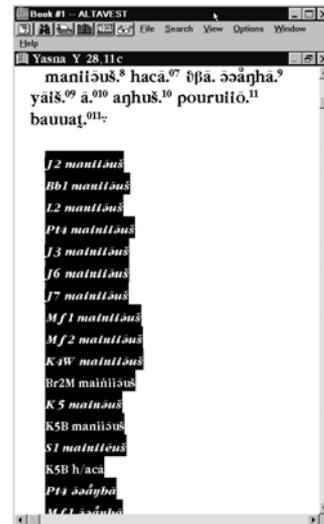


Abb. 17: Variantenapparat (altavest.)

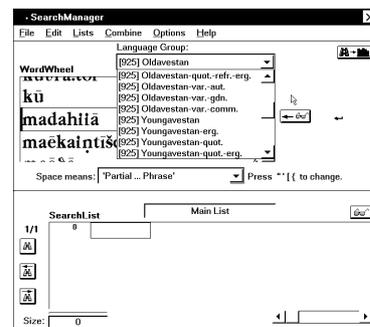


Abb. 18: Sprachliche Diversifikation

Textelemente nach Sprachstufen (Altavestisch, Jungavestisch, Pseudoaltavestisch) und Textsorten (Haupttext, Repetition, Zitat, Formel etc.; s. Abb. 18 und 20).

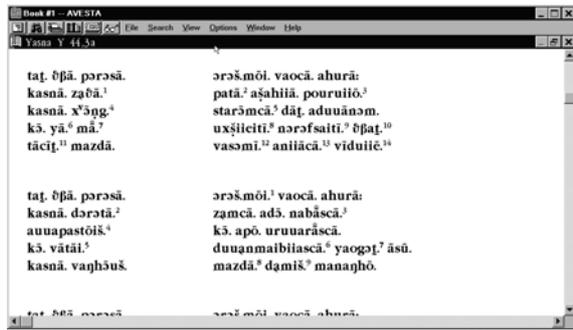


Abb. 19: Altavestischer Text in metrischer Darstellung



Abb. 20: „Liturgische“ Restitution

Weitere Bearbeitungsschritte, die zur Zeit unternommen werden bzw. für die nähere Zukunft geplant sind, betreffen

- H) die Eingabe der Pahlavī-Version des Vidēvdād (durch A. Cantera);
- I) die Eingabe der Pahlavī- und Sanskrit-Version des Yasna (durch D. Durkin-Meisterernst; vgl. das Beispiel aus dem Hōm-Yašt in Abb. 21) und
- J) die digitale Erfassung relevanter Handschriften (vollständig digitalisiert wurde bereits die Hs. J2 nach der Faksimile-Ausgabe von Mills) und deren Verknüpfung mit dem Text über sog. „Hyperlinkverfahren“ (s. Abb. 22).

Eine unmittelbare Zielsetzung, die sich mit den Weiterbearbeitungen im Rahmen des AUREA-Projekts verbindet, besteht in der Erweiterung der Informationsbasis hinsichtlich der handschriftlichen Überlieferung. Dies betrifft im wesentlichen die in Geldners Ausgabe noch nicht verwerteten Handschriften Mf4 und Br2, von denen zumindest die erstere (als Schwesterhandschrift von Pt4) gemäß Geldners eigener Einschätzung von hohem Wert ist (Prolegomena, xxiv-xxvi).

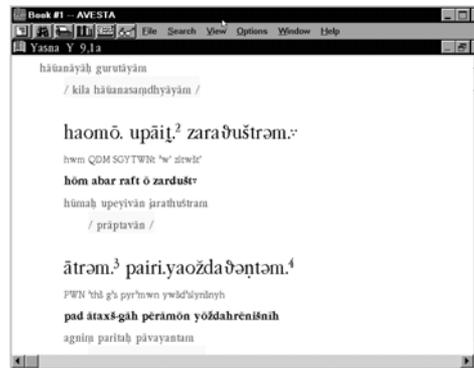


Abb. 21: Pahlavī- und Sanskrit-Version

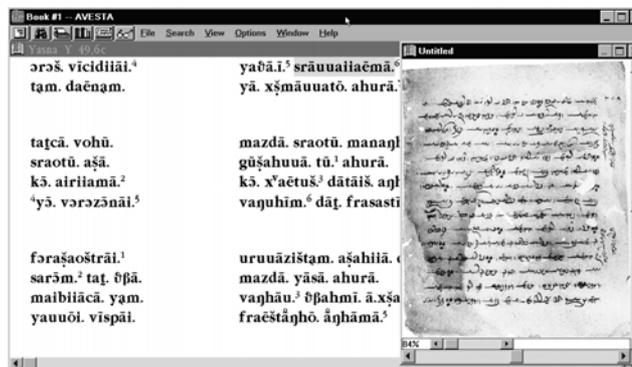


Abb. 22: Verknüpfung von Text und Grafik (Handschrift J2)

Darüber hinaus soll eine weitgehende Überprüfung und, soweit erforderlich, Ergänzung oder Korrektur der Geldnerschen Lesungen erfolgen. Daß dies trotz der allgemein anerkannten und bewunderten Akribie, mit der Geldner gearbeitet hat, sinnvoll ist, läßt sich an verschiedenen Beispielen belegen.

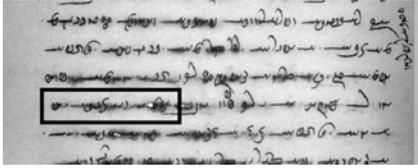


Abb. 23: Y. 28,11bc in der Hs. K5

So erfaßt Geldner für Y. 28,11c an der Stelle des edierten *maniūduš* für die Hs. K5 eine Lesart *maināuš* (Fn. 8), die erheblich von der z.B. durch J2 vertretenen Hauptlesart abweicht. Lesbar ist an der gegebenen Stelle (fol. 139r der Faksimile-Ausgabe von K5) jedoch ganz eindeutig *maniūduš* (s. Abb. 23). Wie Geldners Lesung zustande gekommen ist, läßt sich nicht mehr eruieren; eine Variante *maināuš* zeigt keine der sonstigen verfügbaren Hss. (s. Abb. 17).

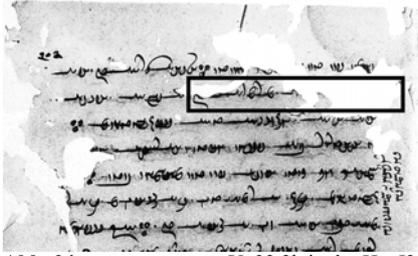


Abb. 24: Y. 32,2b in der Hs. J2

Anders gelagert ist Y. 32,2b, wo Geldner an der Stelle von ediertem *prati.mraoγ* für die Hs. J2 verkürzend die Lesart *prati* verzeichnet (Fn. 2). Legt man den durch die Faksimile-Ausgabe von Mills (Oxford 1893, S. 397) repräsentierten Erhaltungszustand der Handschrift zugrunde, so ist an der gegebenen Stelle gerade das Präverb aufgrund von Beschädigungen nicht mehr erhalten (s. Abb. 24). Auch wenn man mit Geldner als das entscheidende Merkmal das auslautende *-i* noch erkennen mag, wäre doch eine Notation angebracht, die die Aussagefähigkeit des Zeugen einschätzbar machen würde. Eine solche Notation könnte etwa wie folgt aussehen: $\{V \dots V\} i.mr \{NmN\} ra \{V \cdot V\} t$ (mit $\{V \cdot V\}$ für fehlende und $\{NxN\}$ für zu tilgende Elemente).

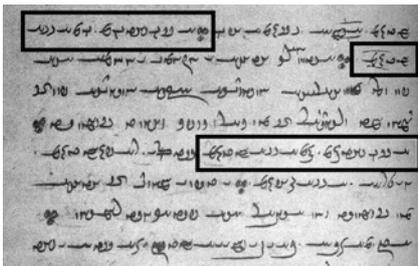


Abb. 25: Y. 34,4b in der Hs. K5

Bisweilen ist auch einfach eine Verdeutlichung Geldnerscher Angaben wünschenswert. Dies gilt etwa für Y. 34,4b, wo es im Apparat zu *asīstīm. āmauuantəm* (Fn. 5) heißt: „K5, in which this and the following word are written twice over by an oversight, has first *asīstīm*, then *asīstām* ... K5 *īm°* and *ām°*.“ Dabei wird die Tatsache außer Acht gelassen, daß die Wiederholung nicht, wie suggeriert, unmittelbar auf die erste Niederschrift folgt, sondern dergestalt, daß der Schreiber das Wortpaar zunächst irrig an das letzten Wort von 4a (*usmahī*) anschließen ließ, dann aber die für a geltende Pahlavī-Übersetzung nachschob, bevor er mit der zweiten Niederschrift des Wortpaars den Vers b neu begann (s. Abb. 25: fol. 171r). Dies könnte etwa durch folgende Notation wiedergegeben werden:

$\therefore \{Nasi\dot{s}t\ddot{i}m. \dot{i}mauuant\ddot{a}m. \cdot N\} \setminus \dots \setminus asi\dot{s}t\ddot{a}m. \dot{a}mauuant\ddot{a}m.$

Wie im Falle der *Ṛgveda-Saṃhitā* ist als nächster Bearbeitungsschritt im Rahmen des AUREA-Projekts auch für das Avestacorpus die Eingabe morphologischer und syntaktischer Zusatzinformationen (grammatikalisches „Tagging“) vorgesehen. Unmittelbares Ziel des Projekts ist darüber hinaus die Erstellung eines Index der Lesarten (zunächst für die altavestischen Texte), der für die Handschriften J2, K5, Mf4 und P1 vollständig sein und, über einen Vergleich mit den bei Geldner erfaßten Lesarten aus diesen Handschriften, Aufschluß über die Zuverlässigkeit des Geldnerschen Apparats geben soll.

Mittelbares Ziel der Weiterbearbeitung ist eine Rekonstruktion der archetypischen Gestalt altavestischer Textteile, wobei zunächst eine Rekonstruktion archetypischer Lesarten, dann eine Rekonstruktion der dahinter zu vermutenden lautlichen Gegebenheiten vorauszugehen hat. Unterstützend soll hierbei ein programmgesteuerter Vergleich altavestischer mit vedischen Wortformen eingesetzt werden, der nicht zuletzt metrische Strukturen berücksichtigen soll.