

## Sprachwandel im Sanskrit? Eine Corpusstudie zum Einfluss P#inis auf die Lexik des Sanskrit

,  
hellwig7@gmx.de  
Universität Düsseldorf, Deutschland

,  
petersen@phil.uni-duesseldorf.de  
Universität Düsseldorf, Deutschland

### Problemstellung und Vorarbeiten

Die Sanskrit-Grammatik A####dhy#y# („[Buch mit] acht Kapitel[n]“) des Linguisten P####ini, der wahrscheinlich gegen 350 v. u. Z. in Nordwest-Indien lebte, ist eine der ältesten wissenschaftlichen Grammatiken einer Sprache (Cardona 1976). Die A####dhy#y# bildet die Grundlage für eine bis heute kontinuierlich fortgeführte Tradition wissenschaftlicher Sanskrit-Linguistik in Indien, und ihr Einfluss auf die indische, aber auch auf die europäische Geistes- und Wissenschaftsgeschichte lässt sich kaum hoch genug ansetzen. Die A####dhy#y# systematisiert linguistische Phänomene des – seinerzeit wohl noch nicht verschriftlichten – spätvedischen Sanskrit und legt so die Grundlage für das klassische Sanskrit, die *lingua franca* von Wissenschaft, Philosophie und Literatur, die eines der größten vormodernen Textcorpora hervorgebracht hat. Daneben wirken Inhalt und Beschreibungsmethoden der A####dhy#y# auf beinahe jeden Bereich der altindischen Geistesgeschichte ein. Aus Sicht der westlichen Sprachwissenschaft entwickelt P####ini in der A####dhy#y# formale Methoden wie eine Metasprache oder Ersetzungsregeln, die heute zu den zentralen Elementen moderner linguistischer Theorien gehören (Kiparsky 2009).

Während der Inhalt der A####dhy#y# und die auf ihr gründende altindische Sanskrit-Linguistik in der Indologie intensiv erforscht wurden, wurde ihr Einfluss auf die Textproduktion und den Sprachwandel in Sanskrit weniger systematisch untersucht. Zum Sprachwandel ist anzumerken, dass das klassische Sanskrit nahezu ausschließlich für Literatur, Wissenschaft und Religion, aber nicht für die alltägliche Kommunikation eingesetzt wurde. Zudem hat die grammatikalische Tradition das Sanskrit als unveränderliche Sprache interpretiert (Deshpande 1993). Daher ist zu erwarten, dass der Sprachwandel im Sanskrit weniger stark ausgeprägt ist als bei gesprochenen Sprachen. Eine kursorische Lektüre von

Sanskrit-Texten zeigt, dass das P####inäische Regelsystem auf den Ebenen von Phonetik und Morphologie fast uneingeschränkt verwendet wurde. Die A####dhy#y# hat sich hier von einer ursprünglich deskriptiven zu einer präskriptiven Grammatik gewandelt und setzt einen „Goldstandard“, von dem sich Sprachvarianten wie z. B. das epische Sanskrit (Oberlies 2003) unterscheiden lassen.

Kaum erforscht ist bisher die Frage, inwieweit die A####dhy#y# nicht nur grammatische Phänomene wie z. B. Wortbildung oder Phonetik, sondern auch die Lexik des Sanskrit beeinflusst hat. Sind Wörter, die in der A####dhy#y# vorkommen, durch die über Jahrhunderte anhaltende Rezitation der Grammatik vor Sprachwandel geschützt? Hier schließt sich die Frage an, ob Autoren die A####dhy#y# selbst oder eine vereinfachende Darstellung als Referenzwerk verwendet haben. Aufgrund der Komplexität der A####dhy#y# formieren sich v.a. seit dem 11. Jahrhundert u. Z. neue grammatikalische Schulen, die auf der A####dhy#y# aufbauen, ihr Regelsystem aber vereinfachen und teilweise auch erweitern (Coward / Raja 1990: 19-20). Diese Schulen werden erst im 16. Jahrhundert durch Bha####t#to#j# D#k#itas Grammatik Siddh#ntakaumud# verdrängt, die die P####inäische Tradition wiederherstellt, obwohl auch hier die Regeln neu angeordnet werden. Dazu passt Houbens Beobachtung, wonach sich die aktive Leserschaft der A####dhy#y# auf einen kleinen Zirkel von Spezialisten beschränkte (Houben 2008: 566), obwohl ihr die grammatikalische Tradition, wie Deshpande bemerkt, im Lauf der Jahrhunderte eine wachsende Wertschätzung entgegenbringt (Deshpande 1998).

Die vorliegende Studie schätzt den Einfluss des Vokabulars der A####dhy#y# auf die spätere Textproduktion mit einem corpusbasierten Ansatz ab. Dabei wird die zeitliche Verteilung P####inäischer Beispielnomina (s. Abschnitt ) in der Literatur des klassischen Sanskrit untersucht. Die Studie geht von Vorarbeiten in (Hellwig / Petersen 2015) aus, wo gezeigt wurde, dass das Beispielvokabular der A####dhy#y# im Lauf der Zeit immer seltener auftritt. Dieses Ergebnis widerspricht der These, dass die Verwendung von Worten in der A####dhy#y# diese vor dem Aussterben schützt. Allerdings ließ sich mit der in (Hellwig / Petersen 2015) verwendeten Datengrundlage nicht ermitteln, ob frühe Sanskrit-Nomina aus anderen Textklassen eine ähnliche diachrone Verteilung zeigen. Die vorliegende Studie erweitert den Referenzrahmen daher um eine zweite Gruppe von Nomina aus der frühen religiösen Literatur ( *#ruti*, „das Anhören [‘heiliger’ Texte]“), deren Texte (Br#hma#as, Upani#ads) in der Sanskrit-Tradition ebenso hoch geschätzt werden wie die A####dhy#y#. Die Verwendung von „religiösen Nomina“ wird mit derjenigen von Beispielnomina aus der A####dhy#y# kontrastiert.

### Daten

Als Corpus dient das Digital Corpus of Sanskrit (DCS) mit rund 3.570.000 Tokens, dessen Texte automatisch

tokenisiert und morphologisch-lexikalisch analysiert und danach manuell korrigiert wurden (Hellwig 2015). Alle Texte des DCS sind einer von fünf Hauptperioden der Sanskrit-Literatur zugeordnet. Die Struktur dieser Zeitachse stellt angesichts der Tatsache, dass zahlreiche wichtige Sanskrit-Texte anonym und ohne Datierung überliefert sind, gerade in den frühen Zeitschichten nur eine grobe Näherung dar (Hellwig 2010). Daneben sind alle Texte mit einem inhaltlichen Label versehen, das aus einem traditionellen Klassifikationsschema abgeleitet wurde.

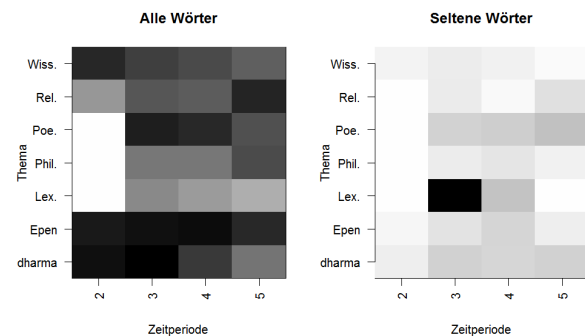
Der Einfluss der A####dhy#y# auf die spätere Literatur wird anhand der Verteilung von 1341 Nominaltypen untersucht, an denen P####ini grammatikalische Phänomene exemplifiziert. Diese Beispielnomina werden in P####ini's Metasprache von Nomina unterschieden, die auf ein externes Objekt referieren (s#tra 1.1.68 der A####dhy#y#, vgl. Séaghdha 2004: 19ff. für eine Übersicht über die Forschung). Während z. B. das Nomen n#sik# in s#tra 1.1.8 der A####dhy#y# auf die Nase referiert, beziehen sich die Nomina *deva* und *brahman* in s#tra 1.2.38, das das vedische Akzente diskutiert, nicht auf einen Gott und einen Brahmanen, sondern werden in ihrer „lautlichen Erscheinung“ verwendet: „Bei [den Nomina] *deva* und *brahman* tritt der *anud#tta* und *svarita* an die Stelle des *svarita*.“ (Übersetzung nach Böhlingk 1886; unsere Ergänzungen in eckigen Klammern). In der Datenbank A####dhy#y# 2.0, die den Text der A####dhy#y# mit zahlreichen grammatischen und semantischen Annotationen unterlegt, sind solche nicht-referierenden Beispielnomina explizit markiert (Petersen / Hellwig 2015). Dadurch lässt sich das P####inäische Beispielvokabular problemlos für corpuslinguistische Studien erschließen.

Das Vergleichsvokabular der frühen religiösen Literatur entstammt wie die A####dhy#y# der frühesten Zeitschicht des DCS, die von der weiteren Auswertung ausgeschlossen wird. Da für die #ruti keine äquivalenten wortsemantischen Annotationen vorliegen, werden aus dem #ruti-Untercorpus die 1268 Nominaltypen ausgewählt, die mehr als einmal auftreten. Diese „#ruti-Nomina“ enthalten Begriffe wie *loka* („(Menschen- / Götter-)Welt“), #tman („Selbst“) und die Namen spätvedischer Gottheiten und spiegeln damit die zentralen Inhalte der #ruti-Literatur wider.

Da das DCS weder thematisch noch zeitlich ausgewogen ist, beruhen die im folgenden dargestellten Verteilungen von Lexemen auf wiederholten Stichproben (*samplings*) fester Größe. Dazu wird das Corpus zuerst anhand von Zeitperioden und inhaltlichen Labels in Faktorstufen aufgeteilt. Anschließend werden aus jeder Faktorstufe 100 zufällige Stichproben von 500 Nomina gezogen. Für jede Stichprobe wird ausgezählt, wie viele der Referenzwörter vorkommen, und die resultierenden prozentualen Anteile werden für jede Faktorstufe gemittelt. Diese Mittelwerte bilden den Ausgangspunkt für die Auswertung im nächsten Abschnitt.

## Auswertung

Abbildung 1 zeigt die gemittelten relativen Häufigkeiten, mit denen Beispielwörter aus der A####dhy#y# im späteren Sanskrit auftreten, als *heatmap*, in der dunkle Farbtöne eine höhere durchschnittliche Häufigkeit in einer Faktorstufe anzeigen. Obwohl gerade seltene Beispielwörter in einigen Domänen wie der Sanskrit-Lexikographie (Zelle „Lex.“, Zeitstufe 3) gehäuft auftreten, nimmt die Verwendung des P####inäischen Vokabulars insgesamt mit der Zeit ab. Dieses Ergebnis deutet darauf hin, dass der direkte Einfluss der A####dhy#y# auf das Sanskrit-Vokabular im Lauf der Zeit geringer wird.

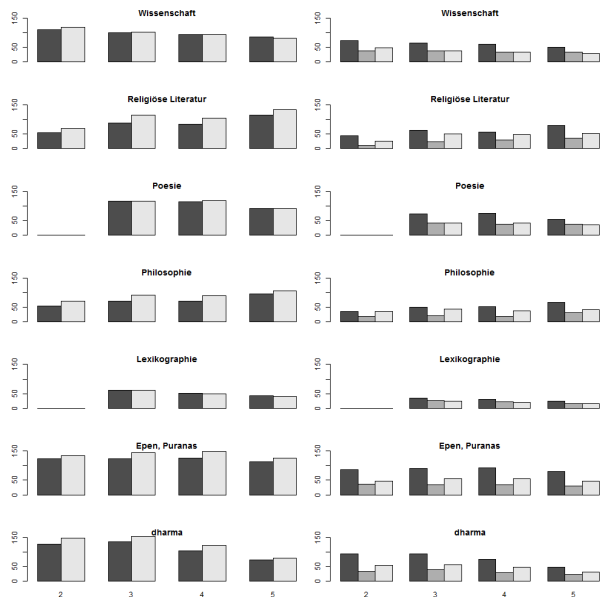


**Abb. 1:** Verteilung von Beispielnomina aus der A####dhy#y#

Abbildung 1 klärt allerdings nicht, ob die Verwendungshäufigkeit bei Lexemen aus der A####dhy#y# weniger stark abnimmt als bei Lexemen aus anderen frühen Texten. In Abbildung 2 wird deshalb die zeitliche Verteilung P####inäischer Beispielnomina derjenigen der #ruti-Nomina in fünf Konfigurationen gegenübergestellt. Der linke Teilbereich zeigt, dass sich die Nomina aus beiden Gruppen grundsätzlich ähnlich über die Sanskrit-Literatur verteilen. Während in der epischen, wissenschaftlichen und „Rechtswissenschaft“ eine allgemeine Abnahme früher Nomina zu beobachten ist, erleben sie in der religiösen und philosophischen Literatur eine Art *revival*. Die Analyse der daran beteiligten Begriffe zeigt, dass hierfür v.a. Begriffe wie *yoga*, *pr#na* („Atem“), #tman („Selbst“) oder *jñ#na* („Wissen“) verantwortlich sind, die z. B. in Yoga-Texten oder bei der damit verbundenen Mikro-Makrokosmos-Spekulation eine zentrale Rolle spielen.

Für die Plots auf der rechten Seite von Abbildung 2 wurde die Vereinigungsmenge des P####inäischen Beispielvokabulars (P) und der #ruti-Nomina (S) in die drei Teilmengen  $P \cap S$  (Schnittmenge),  $P \setminus S$  (alle aus P, die nicht in S sind) und  $S \setminus P$  (alle aus S, die nicht in P sind) aufgeteilt. Der allgemeine Trend dieser Detailverteilungen unterscheidet sich auf den ersten Blick nicht grundsätzlich von der linken Seite des Plots. Bemerkenswert ist allerdings der vergleichsweise hohe Anteil, den die Schnittmenge

P# S am später verwendeten frühen Nominalvokabular einnimmt. Dieser Befund deutet aus unserer Sicht weniger auf eine verstärkte Rezeption P##ini's hin als auf die Tatsache, dass die A###dhy#y# populäre Beispielwörter verwendet.

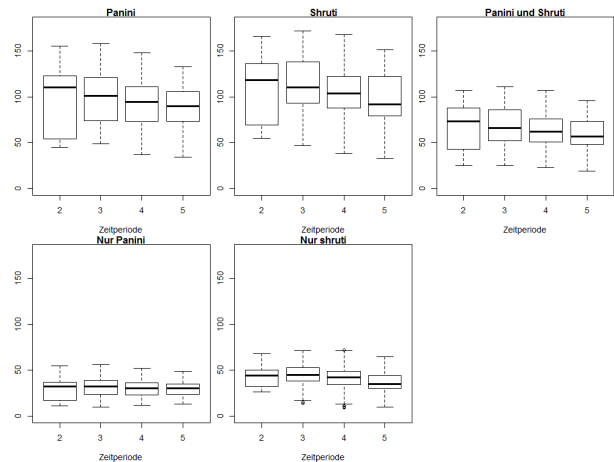


**Abb. 2:** Links: Verteilung von Nomina, die in der A###dhy#y# (dunkelgrau) und in der #ruti (hellgrau) vorkommen. Rechts: Verteilung von Nomina, die in #ruti und A###dhy#y# (dunkelgrau), nur in der A###dhy#y# (mittelgrau) und nur in der #ruti (hellgrau) vorkommen.

Abschließend werden die Häufigkeiten für alle fünf Konfigurationen pro Zeitperiode über alle Themengebiete gemittelt. Die drei oberen Plots in Abbildung 3 zeigen, dass der Kernbestand des spätvedischen Vokabulars im klassischen Sanskrit zunehmend seltener verwendet wird. Weniger eindeutig ist die Verteilung in den beiden unteren Plots, die eine geringe, aber recht konstante Verwendung der Nomina aus P\ S erkennen lassen.

Diese geringere Abnahme des P##inäischen Kernvokabulars scheint die These zu stützen, dass die A###dhy#y# dauerhaft rezipiert wurde und ihre Wörter durch diese regelmäßige Verwendung weniger stark dem allgemeinen Wandel der Lexik unterliegen. Allerdings muss diese Schlussfolgerung in zwei Punkten deutlich eingeschränkt werden. Erstens ist die Vergleichsgruppe der religiösen Texte nur deshalb bis heute überliefert, weil sie ebenfalls als wichtig angesehen und deshalb regelmäßig rezipiert wurden. Hier fehlt ein Corpus von Texten aus der Zeit P##ini's, die danach keine Rolle mehr spielten. Zweitens sollten die Vergleichswörter aus der #ruti genauer semantisch differenziert werden, da z. B. durch den hohen Anteil von Eigennamen für Gottheiten, die auch in späteren Texten eine wichtige Rolle spielten, die lexikalische Aussterberate in den #ruti-Texten unterschätzt werden

könnte. Dieser zweite Punkt wird in einer Folgestudie untersucht werden.



**Abb. 3:** Über Themengebiete gemittelte Häufigkeiten für die fünf Vergleichskonfigurationen

## Fußnoten

1. mukhan#sik#vacano 'nun#sika#, „Ein mit Mund und Nase ausgesprochener Laut heisst *anun#sika*(nasal)“.

## Bibliographie

A###dhy#y#: A###dhy#y# 2.0 <http://panini.phil.hhu.de/panini/panini/> [letzter Zugriff 10. Februar 2016].

**Böhtlingk, Otto** (1886): *P##ini's Grammatik*. Leipzig: Verlag von H. Haessel.

**Cardona, George** (1976): *P##ini. A Survey of Research*. The Hague / Paris: Mouton.

**Coward, Harold G. / Raja, K. Kunjunni** (1990): *The Encyclopedia of Indian Philosophies. 5: The Philosophy of the Grammarians*. Princeton: Princeton University Press.

**Deshpande, Madhav M.** (1993): "Historical Change and the Theology of Eternal (nitya) Sanskrit", in: Deshpande, Madhav M.: *Sanskrit and Prakrit. Sociolinguistic Issues*. Delhi: Motilal Banarsidass Publishers 53-74.

**Deshpande, Madhav M.** (1998): "Evolution of the Notion of Authority (Pr#m##ya) in the P##inian Tradition", in: *Histoire Épistémologie Langage* 20, 1: 5-28.

**Hellwig, Oliver** (2010): "Etymological Trends in the Sanskrit Vocabulary", in: *Literary and Linguistic Computing* 25, 1: 105-118.

**Hellwig, Oliver** (2015): "Morphological Disambiguation of Classical Sanskrit", in: Mahlow, Cerstin / Piotrowski, Michael (eds.): *Systems and Frameworks for Computational Morphology*. Fourth International Workshop, SFCM 2015, Stuttgart, Germany, September 17-18, 2015. Proceedings (= Communications in Computer and Information Science 537). Cham: Springer 41-59.

**Hellwig, Oliver / Petersen, Wiebke** (2015): What's P#ini got to do with it? The use of ga#a-headers from the A####dhy#y# in Sanskrit literature from the perspective of Corpus Linguistics. *Proceedings of the WSC 2015* (forthcoming).

**Houben, Jan** (2008): "Bha#t#toji D#k#ita's "Small Step" for a Grammarian and "Giant Leap" for Sanskrit Grammar", in: *Journal of Indian Philosophy* 36: 563-574.

**Kiparsky, Paul** (2009): "On the Architecture of P#ini's Grammar", in: Huet, Gérard / Kulkarni, Amba / Scharf, Peter (eds.): *Sanskrit Computational Linguistics*. Berlin / Heidelberg: Springer 33-94.

**Oberlies, Thomas** (2003): *A Grammar of Epic Sanskrit*. Berlin: De Gruyter.

**Petersen, Wiebke / Hellwig, Oliver** (im Druck): "Annotating and Analyzing the Ast#dhy#y#", in: Alonso Almeida, Francisco / Ortega Barrera, Ivalla / Quintana Toledo, Elena / Sánchez Cuervo, Margarita (eds.): *Input a Word. Analyse the World. Selected Approaches to Corpus Linguistics*. Cambridge: Cambridge Scholars Publishing.

**Petersen, Wiebke / Soubusta, Simone** (2013): "Structure and implementation of a digital edition of the A####dhy#y#", in: Kulkarni, Malhar (ed.): *Recent Researches in Sanskrit Computational Linguistics*. Delhi: D.K. Printworld 84-103.

**Séaghdha, Diarmuid** (2004): *Object-Language and Metalanguage in Sanskrit Grammatical Texts*. Diplomarbeit. Cambridge: University of Cambridge.