ИНДОЕВРОПЕЙСКОЕ ЯЗЫКОЗНАНИЕ И КЛАССИЧЕСКАЯ ФИЛОЛОГИЯ – XI

Материалы чтений,
посвященных памяти
профессора Иосифа Моисеевича Тронского

18–20 июня 2007 г.

Утверждено к печати
Институтом лингвистических исследований РАН

Конференция проводится в рамках
постоянно действующей Школы индоевропейского
сравнительно-исторического языкознания при ИЛИ РАН

ISBN 978-594047-219-3
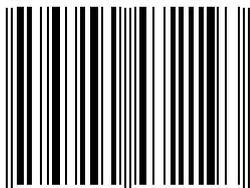
9 785940 472193

A. Keidan

Word boundaries in Pāṇini and Avesta: a linguistic view

I. *Introduction*

When the theories of the grammatical traditions of the Antiquity (such as Pāṇini's grammar in the Ancient India) and those of contemporary linguistic studies seem to match up, in most cases it is only a coincidence, due to the fact that a science of language has to answer the same questions in every period of time, so that some solutions could be found on the same logical ground, both in the traditional linguistics and in the modern studies. Indeed, some theoretic gaps between the ancient approach to the language and modern linguistics always remain strikingly evident, although, at first sight, the results may appear as similar. The most important difference between the two methods consists in a difference of their main purpose: Pāṇini wants to *teach* correct Sanskrit (but the question has been largely debated, cf. Cardona 1997, 543ff.), whereas modern linguists usually try to establish the laws governing the growth and the functioning of the Language *in general* or, as in the case of the IE studies, that of a highly abstract and fictitious linguistic entity, which could not have any practical use.

Thus, for example, as shown by Cardona (1992) in his investigation on Pāṇini's treatment of *Lex Bartholomae*, what Indian Grammar says about the underlying form of the double aspirate stems (like *bhudh-* 'to wake') is only apparently matching with the nowadays diachronic reconstruction of such stems. In fact, in Pāṇini's view, this is a logically deduced, abstract phonological form, which was no way conceived as a really existing form, but served as a starting point for further morphonological transformations bringing ultimately to the correct phonetic form found in the actual language. In modern historical linguistics, instead, the stems with two aspirates are reconstructed as belonging to a more or less "real" linguistic form ascribed to a reconstructed PIE stage.

In this paper I will investigate some aspects of Pāṇini's theory of the word boundaries (in terms of *external sandhi*), in comparison with what the modern Indo-European studies say on this topic and, furthermore, with the state of affairs in the orthography of the sacred books of *Avesta*. As I will try to show, the two ancient grammatical

traditions, the Pāṇinian one and that of native Avestan philologists, gave two very similar conceptions concerning this problem, in both cases sharply distinguishing from what the modern historical linguistic studies say about it.

II. *Case endings in* -bh- *and* -su

As acknowledged by many scholars, in the history of the Indo-Iranian languages (and, more in general, in the PIE) there was a period in which the paradigm of the singular number was more developed than that of the plural (not to speak about the dual). For instance, the plural was lacking of several oblique case-endings. For instance, the case endings in *-*bh*- and the locative suffix *-*su*, which appears in "standard" Vedic (and Avestan) as almost completely univerbated declensional terminations, in this earlier period may have been a kind of unbound morphemes, namely adpositions (i.e., pre- or postpositions). Furthermore, a few examples of the unbound use of such case-endings are to be found in the *Veda* as well. Thus, *ṚV* VIII.11.1 reads *devé mártyeṣu* 'among the gods [and] in the men': what we find here are two nouns terminating with -*e* followed by a unique loc.pl. ending -*ṣu* (see Bartholomae 1888, 583; however, his reading of this passage of *Ṛgveda* is debatable, cf. Oldenberg 1912, 87). This ending seems to act just like a postposition that governs an entire NP, similarly to what is found in *ṚV* VII.1.19 *dáme* […] *váne ā́* 'in the houses [and] the woods' (for the seek of clarity, some of the sandhi are undone), with postposition ā́ governing a conjunction of two locatives (still a Bartholomae's reading). Instead, in Classical Sanskrit the oblique endings in -*bh*- and the loc.pl. ending -*su* are no longer detachable from the noun and are to be considered as true case terminations.

However, even in Classical Sanskrit, in the neuter stems in sibilant and some of those in palatal the -*bh*- and -*su* endings were clearly attached to a nom.sg. form instead of the regular declensional stem. E.g., *samrāj*- 'king' followed by the ending -*bhyas* gives the form *samrāḍbhyas*, with *rāḍ*- virtually identical to the nom.sg. case-form of this word in a voiced context (e.g. *samrāḍ bhāvati* '[he] becomes a king'). Instead, if the ending were attached directly to the declensional stem as in the rest of the paradigm, the word form would have been *samrāgbhyas. Furthermore, *manas*- 'thought' gives *manobhyas* apparently from nom.sg. *mano* (so, with the external sandhi of -*as* in voiced context), instead of the regularly expected form *manadbhyas (stated by Brugmann and Delbrück

1897–1916, v. I, §830a; such a sandhi is attested e.g. in *ṚV* I.6.3 *uṣádbhir* from *uṣā́s* 'dawn'; however, Debrunner and Wackernagel 1929, §150 regard *-dbh-* < *-sbh-* as analogical and innovated). In Avestan we can observe some other instances of nom.sg. functioning as stem, e.g. *vaɣž̌biš* 'with words' from nom.sg. *vaxš* instead of the stem form *vač-*, see Hoffmann and Forssman (1996, 116).

In the other declensional types the phenomenon is more opaque, or completely non-existing. Thus, the feminine stems in *-ā* are ambiguous, being their nom.sg. the same as the bare stem: *senābhyas* from *senā-* 'army', nom.sg. *senā*. In many other declensional stems the *-bh-* and *-su* endings are regularly attached to the bare stem: *patibhyas* from *pati-* 'lord, master' (cf. nom.sg. *patiḥ*, *patir* in voiced context). Lastly, in the thematic declension we can observe an innovated stem in *-e-* influenced by the plural forms of the pronouns: *devebhyas* from *deva-* 'god', nom.sg. *devaḥ* (*devo* in voiced context), cf. *tebhyas* 'to/from them' from *sa-* 'he', cf. Brugmann and Delbrück (1897–1916, v. II, §277).

Some other facts from Sanskrit and other IE languages corroborate the hypothesis of a later origin of these case-endings from ancient unbound morphemes, as well.

1. To begin with, no such endings are found in Hittite, that is the most early attested of the IE languages. This led some scholars (see a summary in Szemerényi 1972, 150) to conclude that the endings in *-bh-* and *-su* entered the nominal paradigm during the so-called "fourth period" in the history of the PIE, when the Hittite branch was already separated from the rest of the *Stammbaum*.

2. Neither in the remaining IE languages these morphemes became always true case-endings. The endings in *-bh-* are found in: Indo-Aryan, Iranian, Armenian, Greek, Italic, and Celtic (besides some other minor languages). Balto-Slavic and Germanic languages show almost regularly endings beginning with *-m-* in the corresponding case forms (in Germanic, however, a preposition tracing back to the morpheme in *bh- is preserved), cf. Brugmann and Delbrück (1897–1916, v. II, 119).

3. In the Greek language the case-endings in *-bh-* never entered the nominal paradigm completely. The only form of the IE morpheme attested in Greek is -φι(ν), which is usually employed in what appears as a polysemic oblique case-form of the plural, e.g. Mycenaean *po-pi* 'with/on/by feet' (read πόπφι). Furthermore, the termination -φι occurs in Mycenaean exclusively with consonantal stems and feminine *ā*-stems (see Lejeune 1958, 167). In the Homeric Greek -φι(ν) ending has been expanded also to the thematic declension, and to the singular

(cf. Homer φ 295 Ἰλιόφι 'from Ilium', and so on). However, the singular interpretation is debatable: an adverbial reading is almost always admissible since in the post-Mycenaean period this formations were semi-adverbial, and not inflectional, in nature.

4. Moreover, in some declensional types the case-endings in *-bh-* and *-su* could bring no accent in Vedic, even if stressed flections were required, cf. Sanskrit *mahádbhis* but *mahatā́*. The Greek case-ending -φι does not bring the accent at all, cf. ναῦφι but νηός. So, they could be considered as having entered the paradigm in a later period, when the accentual pattern of the noun declension underwent a complete reorganization, cf. Kuryłowicz (1935, 209).

5. In addition, such an accentual feature is also shared by many derivatives: Greek λευκός 'white' but λεῦκος 'white fish', Vedic *kṛṣṇá* 'black' but *kŕ̥ṣṇa* 'Krishna', which leads us to the conclusion that the forms in -φι were perceived as derivatives, «[…] simplement des *adverbes dérivés de thèmes nominaux*, et non pas des cas au sens propre du mot» as says Kuryłowicz (1935, 167).

6. In Avestan, the initial *-b-* of these case-endings is almost always preserved from being lenited into *-β-*, as if it stood at the beginning of a separate word, and not in internal position, see De Vaan (2003, 351fn435).

7. No other case termination begins with a bilabial stop: this puts the *-bh-* terminations in a particular position, cf. Meillet (1937, 298).

8. There are other such semi-adverbial particles in the IE languages, which never became case-endings (as those found in Greek οὐρανόθεν 'from heaven', οὐρανόθι 'in heaven'), or entered the inflectional paradigm only in a later period (such as Sanskrit *-tas* in *mukhatás* 'from the mouth', cf. also the next point).

9. The history of the Indo-Aryan is a continuing rebirth of case-endings: the phonetic erosion of the end of the word was contrasted by creating new case inflections from former adverbial morphemes, cf. Pischel (1981, §369) for Prakrit abl.pl. ending *-himto* from *-bhi*(*s*) plus the already mentioned particle *-tas* (a very similar agglutination process had taken place also in Avestan: cf. the loc.pl. case ending *-huuā* going up to the original suffix plus the postposition *ā*). Eventually, this process generated the present day system of three different "layers" of case-endings clearly visible in the Modern Indo-Aryan languages, cf. Masica (1991, §8.4).

To sum up, the case forms in *-bh-* (and, to a certain degree, also those in *-su*) represented, in the earliest IE languages, not exactly flectional, but rather adverbial forms, very polysemic from the

functional point of view, and with survival of a free use. To say it with Bartholomae (1888, 581), «Es ist zweifellos, dass hier *bhiṣ*, [*bhyas*], *su*, *ṣu* vom sprechenden deutlich als bildungselemente empfunden wurden». Such originally adverbial morphemes entered the inflectional paradigm properly only in a later period, not in the whole IE domain, and not always to the same degree. In particular, in Vedic these terminations were definitively univerbated when the category of declensional stem was no longer alive, and thus were attached to the nom.sg. form (in its voiced sandhi variant) functioning as a new stem.

III. *Pāṇini's point of view*

Let us see now what Pāṇini's Grammar says about the -*bh*- and -*su* case endings, and which differences exist between the Ancient Indian point of view and that of modern linguistics. In the *sūtra* 1.4.14 Pāṇini gives the definition of the category of *pada* 'word', or, better, 'basic word form'. The rule says: *suptiṅantam padam* 'what ends in *sUP* or in *tiṄ* is *pada*'. This means that those elements which terminate with one of the declensional endings belonging to one of the two sets: *sUP* 'nominal endings' (listed in the *sūtra* 4.1.2) and *tiṄ* 'verbal endings' (listed in the *sūtra* 3.4.78) are considered as *pada*. The terms *sUP* and *tiṄ* are *pratyāhāra* 'artificial technical terms', i.e. they are abbreviations obtained by compounding the first and the last member (in their abstract shape) of the list itself.

What is unusual with Pāṇini's concept of *pada* is that he considers as *complete* a word-form not necessarily provided with the declensional ending. This is stated in the subsequent *sūtra* 1.4.17 *svādiṣv asarvanāmasthāne* 'what precedes *su* etc., excluding the *sarvanāmathāna* set, is also considered *pada*'. Here, by the term *svādi* 'beginning with *su*', Pāṇini defines a group of nominal endings, among which there are those beginning with -*bh*-, such as -*bhis* (instr.pl.), -*bhyas* (dat.pl. and abl.pl.), -*bhyām* (instr.du, dat.du and abl.du), and the loc.pl. ending -*su*. Therefore, Pāṇini considers as *pada* also the word-forms deprived of such case-endings. On the boundary between certain types of nominal stems and these terminations some apparently external sandhi phenomena (i.e. euphonic transformations depending on the position of the word in the utterance) can, indeed, be observed.

At first sight, such a twofold definition of 'word' would be perfectly fitting with the modern Proto-Indo-European reconstruction: a stem being considered as a word even without the case endings attached seems to imply that, in Pāṇini's view, the

terminations in *-bh-* and *-su* were free morphemes. However, the conception of *pada* does not entail any consequence on Pāṇini's theory of lexical semantics and syntax (*pace* Cardona 1997, 42): he does not in fact consider forms like *raḍbhyas* as a merging of two different semantic items, or two different syntactical units, two "words". For nowhere in his Grammar the morphemes in *-bh-* and *-su* are stated to be *pada*s themselves.

The concept of *pada* is purely phonological—if not completely abstract and procedural—in nature. It is only a "descriptive trick" used by Pāṇini in order to account, in a wholly synchronic manner, for the—real or apparent—external sandhi phenomena involving some consonantal stems and the following case-endings.

IV. *Evidence from* Avesta

In Avestan the loc.pl. ending *-huuā* is preceded by a period in writing, as if it were a compound; the compounds, indeed, were written with a period splitting the members thereof. This features of Avestan seem to prove the late origin of this case ending, which would be still a semi-autonomous lexical item in documented period, perceived by the speakers as a compounded stem, and not a grammatical morpheme. However, this interpretation is false.

In fact, the text of *Avesta* had undergone what was defined as a redactional compound split, which was often made on an etymologically wrong ground, cf. Narten (1986, 275). The ancient Iranian philologist, or the *diasceuasts*, as Hoffmann (1958, 8) calls them, reshaped the text of *Avesta* splitting graphically the members of the compound words. The split was marked by writing a period, but also by restoring artificially the nom.sg. case form in the first member of compound (*druxš.manah-* 'liar' instead of **druj.manah-*), or by introducing the "composition-vowel" (Bartholomae 1894-95, 150) *-ō* at the end of the first member, by analogy with the nom.sg. of stems in sibilant and those of the thematic declension, cf. De Vaan (2003, 434).

The actual criterion used by Avestan diasceuasts in order to individuate the compounded items was simply that of the possible autonomous occurrence, without any consideration of meaning. As a consequence, since loc.pl. *-huuā* coincided with pronoun *huua-* 'this', the decision was made to mark this case ending with a period and a preceding *-ō* as if it was a second member of a compound (e.g. *təmō.huua < təmah-* 'darkness', *xšapō.huua < xšapan-* 'night'). That this was no way based on a true etymological analysis is proven by the fact that many other nominal and verbal terminations (e.g. *aošō.tara-* comparative of *aoša-* 'harmful'), or even word fragments wrongly

considered as morphemes (e.g. instr.pl. *drəguuō.dəbīš < drəguuant-* 'deceitful', noticeably not °*.bīš*) were marked in the same way, see De Vaan (2003, §22.5).

V. Conclusion

We can see now that two different ancient philological traditions, that of *Veda*s and that of *Avesta*, reached the same results independently. They both managed with a sacred text that had to be preserved from the corruption of the time, being no more natively spoken the languages in which the two books were written. In both cases, this process brought to an etymologically wrong analysis, which however was necessary in order to account for the orthoepy of the ritual recitation.

The main difference is that Pāṇini limited himself in endorsing the results of the *padapāṭha* analysis (which predated him), when he introduced his merely artificial and descriptive concept of *pada*. The Avestan philologist, instead, had gone further in reshaping analogically the members of compounds, and in individuating wrong compounds. In any case, the two ancient tradition in consideration are completely different, as far as the general approach is concerned, from the modern diachronic analysis of the Indo-Iranian languages. Any coincidence is only fortuitous.

<div align="center">Reference</div>

Bartholomae, Ch. 1888: "Die arische flexion der adjektiva und partizipia auf nt-", *KZ* 29: 487–586

Bartholomae, Ch.1894–1895: "Vorgeschichte der Iranischen Sprachen", in Geiger, W. and Kuhn, E. (eds.), *Grundriss der Iranischen Philologie. I.1*, Strassburg

Brugmann, K. and Delbrück, B. 1897–1916[2]: *Grundriss der vergleichende Grammatik der indogermanischen Sprachen*, Strassburg

Cardona, G. 1992: "Indian grammatical traditions and historical linguistics", in Polomé, E.C. and Winter, W. (eds.), *Reconstructing Languages and Cultures*, Berlin, New York, 239–259

Cardona, G.1997[2]. *Pāṇini. His Work and its Traditions. 1. Background and Introduction*, Delhi

De Vaan, M. 2003: *The Avestan Vowels*, Amsterdam, New York

Debrunner, A. and Wackernagel J. 1929: *Altindische Grammatik. III.1 Deklination der Nomina*, *Zahlwörter—Pronomina*, Göttingen

Hoffmann, K. 1958: "Altiranisch", in *Handbuch der Orientalistik. I.IV Iranistik*, Leiden — Köln

Hoffmann K. and Forssman B. 1996: *Avestische Laut- und Flexionslehre*, Innsbruck

Kuryłowicz J. 1935: *Études indo-européennes. 1*, Cracow

Lejeune, M. 1958: "La désinence -φι en mycénien", in *Mémoires de Philologie Mycénienne*, 159–184

Narten, J. 1986: "Zum Vokalismus in der Gatha-Überlieferung", in Schmitt R. and Skjærø P.O. (eds.), *Studia grammatica iranica. Festschrift für Helmut Humbach*, München, 257–278

Masica, C. P. 1991: *The Indo-Aryan languages*, Cambridge

Meillet, A. 1937: *Introduction à l'étude comparative des langues indo-européennes*, Paris

Oldenberg H. 1912: *Ṛgveda. Textkritische und exegetische Noten. Siebentes bis zehntes Buch*, Berlin

Pischel, R. 1981[2]: *A Grammar of the Prākrit Languages*, transl. from German by S. Jhā, Delhi

Szemerényi, O. 1972: "Comparative Linguistics", in Haugen, E. and Winter, W. (eds.), *Linguistics in Western Europe* (CTL 9), The Hague, Paris, 119–195