Research Article

Jan E.M. Houben*

Linguistic Paradox and Diglossia: the emergence of Sanskrit and Sanskritic language in Ancient India[1]

https://doi.org/10.1515/opli-2018-0001
Received December 1, 2017; accepted February 20, 2018

Abstract: "We know that Middle Indian (Middle Indo-Aryan) makes its appearance in epigraphy prior to Sanskrit: this is the great linguistic paradox of India." In these words Louis Renou (1956: 84) referred to a problem in Sanskrit studies for which so far no satisfactory solution had been found. I will here propose that the perceived "paradox" derives from the lack of acknowledgement of certain parameters in the linguistic situation of Ancient India which were insufficiently appreciated in Renou's time, but which are at present open to systematic exploration with the help of by now well established sociolinguistic concepts, notably the concept of "diglossia". Three issues will here be addressed in the light of references to ancient and classical Indian texts, Sanskrit and Sanskritic. A simple genetic model is indadequate, especially when the 'linguistic area' applies also to what can be reconstructed for earlier periods. The so-called Sanskrit "Hybrids" in the first millennium CE, including the Prakrits and Epics, are rather to be regarded as emerging "Ausbau" languages of Indo-Aryan with hardly any significant mutual "Abstand" before they will be succesfully "roofed," in the second half of the first millennium CE, by "classical" Sanskrit.

Keywords: Sanskrit, Prakrit, sociolinguistics, "Hybrid" Sanskrit, Old Persian, diglossia, emerging languages

1 Introduction

"We know that Middle Indian (Middle Indo-Aryan) makes its appearance in epigraphy prior to Sanskrit: this is the great linguistic paradox of India." In these words Louis Renou (1956: 84) referred to a problem in Sanskrit studies for which so far no satisfactory solution had been found. I will here propose that the perceived "paradox" derives from the lack of acknowledgement of certain parameters in the linguistic situation of Ancient India which were insufficiently appreciated in Renou's time, but which are at present open to systematic exploration with the help of by now well established sociolinguistic concepts, notably the

1 For their invitation to participate in a fascinating and unique event of comparative historical sociolinguistic research, the Conference "Strategies of Language Variation: Transcultural Perspectives" (Vienna, 24-25 April 2015), the author thanks Vincent Eltschinger and the organizers of the Conference ; he thanks Chiara Barbati and Christian Gastgeber for their invitation to contribute to the proceedings. Wouter Henkelman he thanks for a discussion on the Behistun (Bīsutūn) inscriptions and on Old Persian. He is grateful to two anonymous reviewers for their suggestions for improvement. Representation of characters: R and r with subscript circle have been replaced by R and r with subscript dot; these and a few other obvious replacements by typographically less problematic diacritics were necessary to produce a readable pdf.

Article note: this paper belongs to the Topical Issue on Historical Sociolinguistic Philology, ed. by Chiara Barbati and Christian Gastgeber

***Corresponding author: Jan E.M. Houben,** EPHE, PSL – Ecole Pratique des Hautes Etudes, Paris Sciences & Lettres, Université Paris, E-mail:johannes.houben@ephe.sorbonne.fr

concept of "diglossia" (Deshpande 1985; Houben 1996a). This I will do by addressing three issues specified by the organizers of the Conference "Strategies of Language Variation: Transcultural Perspectives." These will be addressed in the light of references to ancient and classical Indian texts, Sanskrit and Sanskritic, and in continuation of my earlier studies in this domain (esp. Houben 1996b, 2011, 2014, 2016). No completeness can be claimed in dealing with these issues: the main points are to be developed at other occasions. In a broad sense, Sanskrit can be taken to include its predecessors, Vedic or "the older dialects of Veda and Brāhmaṇa" (Whitney 1888) and Pāṇini's and Patañjali's *bhāṣā* 'conversational language'. In a strict sense, Sanskrit refers to classical Sanskrit which arose in the first centuries CE and flourished throughout the first millennium in South Asia and beyond, until the beginning of the second, "vernacular" millennium (Pollock 2006), when its niche became more restricted (Houben 2008).

Just as structural and generative linguistics have a predecessor in Ferdinand de Saussure (1857-1913) – who has several times been explicitly mentioned as such[2] – sociolinguistics has an important predecessor in Antoine Meillet (1866-1936), for whom language was "éminemment un fait social" (Meillet 1921: 230). However, he has been rarely recognized as such.[3] Meillet added that language enters exactly into the definition of a "social fact" as given by Émile Durkheim in his 1895 essay on the method of sociology.[4] Although Meillet intended the "social" nature of language to be applicable both to current and historical forms of language use, sociolinguistics is at present mainly developed with regard to current language use. To systematically apply a sociolinguistic approach to historical texts and languages is therefore an innovative move from which both sociolinguistics and the study of ancient languages and literatures can be expected to profit and progress in new ways.

Contributors to the conference had been invited to address "issues resulting from taking up a current sociolinguistic perspective on the phenomena of variation / heterogeneity in conceptually *written* language in *historical* texts" and, more specifically, to address the following six specific questions:

a. In addition to the 'matrix language' (the most used one), which other languages or varieties can be found in your treated texts (e.g. in the form of style shifts and code-switching)?
b. How frequently do these languages/varieties occur? Are the different linguistic systems easily distinguishable, or are some words difficult to assign to a specific language/variety?
c. What social significance or 'meaning' (social associations) do these different languages/varieties bear?
d. What rhetorical effects can the use of these languages/varieties be linked with?
e. To what extent can be assumed that the audience (the addressees / readers) of the texts understood such rhetorical effects?
f. Are there any contemporary or historical meta-linguistic materials commenting on the use of different languages/varieties in the surviving texts? What comments are made?

2 The importance of Saussure's work for structural and generative linguistics was frequently pointed out in the latter half of the previous century and has become so trite that it now normally remains implicit. As early as in 1957, N.C.W. Spence observed "it can be said that 'we are all Saussureans now'." Noam Chomsky placed himself explicitly in the tradition of saussurean linguistics. A study and analysis on the importance of Saussure's work for 20th century linguistics (till the mid-eighties) is Kaldewaij 1986.

3 I am only aware of L.-J. Calvet 1998 who briefly highlighted the importance of Antoine Meillet for modern sociolinguistics in the beginning of his book and draws attention to mild theoretical divergences with Saussure. For Meillet, however, his own conviction that the "faits de langue" should be explained through the "life of man in society" ("la vie de l'homme en société") (Meillet 1936: 226) was rather supported by the position that each language is "a rigorously arranged system where everything coheres" ("un système rigoureusement agencé, où tout se tient") (Meillet 1936: 158), the position of which Saussure his held to be the earliest spokesman although the exact statement is nowhere found in his writings or in the Cours de linguistique générale (Koerner 1997). In the time of Meillet, a sharp and detailed awareness of the dynamics in the relationship between "man in society" and linguistic variation and evolution can already be found, for instance in Jakob Wackernagel's "Einleitung" to his Altindische Grammatik (1896); however, concepts and theories to deal with this dynamism were still entirely lacking: these started to be developed together with the emergence of the discipline of sociolinguistics, most importantly, in the work of, on the one hand, William Labov (who explicitly presents his work as relevant, also, to understand and reconstruct the past: Labov 1994: 9-27), and, on the other hand, in the work of Heinz Kloss (Kloss 1967, Muljačič 1986) regarding the emergence and development of dialects and languages.

4 First chapter of Durkheim 1895, entitled "Qu'est-ce qu'un fait social?"

Answering questions (c) to (f) would amount to writing a new history of Sanskrit and Indian literature from a sociolinguistic perspective such as explored, for instance, in Madhav Deshpande's *Sanskrit and Prakrit: Sociolinguistic Issues* (1993). Question (b) consists of two related ones, which will be addressed here separately.

2 Sanskrit: matrix, or outcome of related idioms?

The first question touches on a fundamental issue in the study of Sanskrit of which the earliest Western students of Sanskrit were well aware, but which has till now not been addressed properly, partly on account of the extension of the subject, partly on account of the lack of adequate concepts.

> a. In addition to the 'matrix language' (the most used one), which other languages or varieties can be found in your treated texts (e.g. in the form of style shifts and code-switching)?

The question as formulated presupposes that we have sufficient access both to a "'matrix language' (the most used one)" and to "other languages or varieties" which are its outcome. In our case, the language to which we have access quite extensively is (classical) Sanskrit, but it is not demonstrated, and a priori unlikely, that Sanskrit would be the "'matrix language' (the most used one)" for languages and idioms with which it had a dynamic relationship. Throughout its long history, Sanskrit, as the language that is well formed or well prepared (*saṁ-s-kṛta*), presupposes varieties of language that are less well formed – either because the linguistic norms are imperfectly realized, or because different linguistic norms are followed. As the name of a language or of a variety of language, the term Sanskrit (*saṁ-s-kṛta*) is relatively late (from the first centuries CE onwards), but it is linguistically more or less identical with the language used and discussed in Patañjali's Mahābhāṣya ('Great Commentary', 2nd century BCE) as the *bhāṣā* 'conversational language' described in Pāṇini's grammar, the Aṣṭādhyāyī (AA), 4th century BCE.[5] The language used and described in the Mahābhāṣya will become exemplary in the period of "classical" Sanskrit. In addition, a more archaic variety is described by Pāṇini: the language of the ancient Vedic hymns. The oldest, very extensive collection of Vedic hymns is the Ṛgveda, rich in poetic eulogies of Vedic gods such as Agni, Indra, Varuṇa, but also containing "philosophical" reflections on, for instance, the place of man in the world and in the universe (Renou 1957b). The verbal root *kṛ* 'to do' has the present stem *kṛ-ṇu* in the entire Ṛgveda, with three exceptions, all in the tenth and last Maṇḍala, generally regarded as the latest one (Whitney 1888: 260). In verse 2 of ṚV 10.145, for instance, the imperative *kur-u* is used by a woman "conjuring against her co-wife for the affections of their joint husband" (Jamison & Brereton 2014: 1630). This does not represent a systematic and generally accepted style shift in the Ṛgveda, as in other sentences in ṚV 10 attributed to women we find that only the older stem *kṛ-ṇu* is used (as in ṚV 10.95, the dialogue between Purūravas and Urvaśī). In his extensive study under the title "Tracing the Vedic dialects," Michael Witzel (1989: 101) refers to this rare use of *kuru* in the Ṛgveda and to several other indications of "social levels of language" in Vedic texts. Subsequently, however, Witzel's study (see also Witzel 1987, 1997) is focused on the parameters "geography" and "time" and the parameter "social levels of language" is no more taken into account. Forms derived from *kur/kur-u* rather than *kṛ-nu* become more prominent in the Atharvaveda and it is the normal present stem in the prose of the Brāhmaṇas and in classical Sanskrit. While *kuru* may be regarded, formally, as a "later" form in the tenth Maṇḍala, invoking a later stage of the language, viz. classical Sanskrit, cannot explain the form synchronically. It has been proposed that it derives from a "Vedic Prakrit", a "Middle Indo-Aryan" otherwise unattested form **kuṇu*, from Vedic *kṛṇu* (Mayrhofer 1951: 136, with a reference to Wackernagel's *Altindische Grammatik* I, 1896). This amounts to the contemporaneous availability to the users of the language, at least at the time of the tenth Maṇḍala of the Ṛgveda, of two levels of speech, one

5 For Pāṇini, even particularities regarding the placement of the accent are important in the characterization of this conversational use of language (in contradistinction to *chandas,* the Veda, esp. or originally the Vedic Saṁhitās Thieme 1935 : 67ff), as in AA 6.1.181 *vibhāṣā bhāṣāyām* [… 159 *antodāttaḥ*] "In the domain of conversational speech [… has the elevated pitch accent on the last syllable] optionally (preferably not)." Just before the "creolization" of Indo-Aryan in the form of the development of "classical" Sanskrit and the complete loss of accent, this still was a living feature in the conversational language (Thieme 1985).

level of high prestige normally used to address the gods, and a middle or lower, "Prakritic" one, used, for instance, by women. In the other hymn containing *kuru* (ṚV 10.19, verse 2b), which may also represent a more popular register as it deals with the returning home of cattle in the evening, this form co-occurs with another linguistic form, the simple nominative plural ending in *dev'āḥ … yajñíyās* (10.19.7c). This form could be regarded as simply "later" if we arrange the linguistic forms exclusively according to a "time-line" of predominant usage, but it points, contemporaneously, to an apparently widespread distinction between levels of language known not only in Vedic but also in old Iranian: the distinction between the simple nominative plural ending of *o/a* stems (Skt. *āḥ*, Av *å/a*, OP *āḥ*) and the double ending *āsaḥ* (Skt. *āsaḥ*, Av *åŋhō*, OP *āḥ/āhaḥ*). To the discussion by Witzel (1989: 212) is to be added that the double ending *āhaḥ* in OP is found only very exceptionally, namely in the expression *aniyāha bagāha* 'the other gods', which "seems to come from the language of religion" (Kent 1950: 9), in other words, from a "higher prestige" level of language. On the Indian side, the simple ending *āḥ* is the one that, in the formulation of Witzel, "has gained prominence in all Prākṛts (*āḥ* > *ā*), except for *āse* in Pāli verses" (with a reference to von Hinüber 1986 : 144, §312). Again, the double ending is found in the "higher level" context of poetry. In view of the variation attested in Vedic and in Avestan and OP, the simple ending *āḥ* probably never "gained in prominence" but had always remained available next to the "higher level" use of language in poetry and in religious contexts. Traces of an actual "matrix language" are rare in the transmitted texts, but they are sufficiently attested to infer that it was current, including in "Prakritic" or "Middle Indo-Aryan" language use contemporaneous with the composition of the Veda. The double ending Skt. *āsaḥ*, Av *åŋhō*, OP *āḥ/āhaḥ*) was, on the other hand, a pre-Vedic and pre-Avestan "hyper prestigious" form, not to say a "hypersanskritized" form, if we allow ourselves to take the term "sanskritization" in a generic, linguistic sense and apply it to a situation long before the emergence of classical Sanskrit or Sanskrit in the strict sense of the word.

In fact, as early as in 1896, Jakob Wackernagel was well aware of a distinction in language according to what he called in his time "Volksklassen". Ca. fifty years later his statement to this effect was rendered as follows by Louis Renou (1957a: 7):

> Ainsi la scission du langage d'après les classes sociales, qui s'observe partout mais n'est nulle part plus forte que dans l'Inde, se laisse attester dès l'époque védique.
> (Thus the division of language according to social classes, which is observed everywhere but is nowhere stronger than in India, can be witnessed from the Vedic period.)[6]

Renou was able to add to this statement a new note 89 on "linguistic stratification with social origin (stratification linguistique d'origine sociale)" with several bibliographical references, at a time that it would still take around a decade before sociolinguistics would emerge as an academic discipline. Renou's most recent reference was to Marcel Cohen's *Pour une sociologie du langage* (1956) which explored the possibilities for sociolinguistic studies and for a sociology of language.

In the next period for which language use is sufficiently accessible, the one to which the grammarians Pāṇini (4th century BCE) and Patañjali (2nd century BCE) belonged, the role of "'matrix language' (the most used one)" accrued, again, not to "Sanskrit", referred to as *bhāṣā*, the 'conversational language', but to some form of Prakrit, a continuation of the "Prakritic" language use infered for the Vedic period, and a language variety on which we have, for Pāṇini's and Patañjali's period, still only very limited direct information, mainly in the inscriptions of king Aśoka (3rd century BCE). Patañjali's commentary the Mahābhāṣya or Vyākaraṇa-Mahābhāṣya is itself an excellent example of conversational Sanskrit, although the term *saṁ-s-kṛta* is still nowhere used to refer to this language or idiom. In addition, there are the extensive texts of early Buddhism, which, however, have been fixed in writing a few centuries later, long after the discourses and discussions of the Buddha which are supposed to be reported in many of these texts.

6 Wackernagel 1896, "Einleitung" p. XIX: "Somit hat die überall vorkommende, aber in Indien am schärfsten ausgeprägte Scheidung der Sprache nach Volksklassen hier schon in v. [vedischen] Zeit geherrscht."

Although clapping is never done with a single hand, from the current perspective we perceive for over around two millennia, starting with Patañjali's Mahābhāṣya, a single "Sanskrit" hand clapping.[7] From the clapping itself we have to infer that there was, according to time and circumstances, another "proto-Prakrit" or "Prakrit" or "approximative Sanskrit"[8] hand clapping of which we often have no direct information at all, sometimes only a limited amount of evidence (as in Aśoka's inscriptions), and only for later periods in the course of the second millennium CE somewhat detailed information – but by that time Prakrit (and Pali) no more represent a Prakritic language in current use but have developed into codified, mostly literary languages in their own right. The extrapolations to which this uneven distribution of the evidence and its frequent distortion through transmission continuously invites us, are unavoidably informed by our understanding of linguistic processes in better documented areas and periods. Hence, we cannot afford to neglect either the exploration of primary sources, or the reflection on fundamental theoretical issues connected to their interpretation.

The language described by Pāṇini and Patañjali was limited to the "high prestige" form of linguistic usage current in their time. Those current idioms contained a whole range of linguistic forms in contemporaneous use, from "Prakritic" to various degrees of approximation of the high standard Sanskrit, next to *bhavati* 'he is', for instance, both *bhoti* and *hoti*. This can be inferred from, *inter alia*, the language attested in Aśokan inscriptions found throughout the Indian subcontinent and dated in the 3rd century BCE, in between Pāṇini and Patañjali. Although the grammarians decided to describe only the desirable "high prestige" forms of the language, and not to bother about indicating all possible lower forms (*apaśabda*), sporadic references in Patañjali's commentary give an idea of these forms regarded as having a lower prestige (see below).

An important domain of sociolinguistic variation is ancient Indian theatre. A number of "classical" Indian dramas have been transmitted over the centuries and are available, the most important ones dating from the middle of the first millennium CE onwards. The dramas follow patterns and rules which have been set forth in texts such as the Nāṭyaśāstra (2nd or 3rd century CE? Kane 1971: 43-47; S.K. De 1960: 18). The rules also concern which language is to be used by which character. In larger classical dramas, "Sanskrit is spoken mainly by the educated, upper-class male protagonists, while various types of Prākrits are used by most women and by males of lower rank and education" (Hock & Pandharipande, 1976: 113). The earliest dramas that are fragmentarily preserved are those by a Buddhist author, Aśvaghoṣa (ca. 100 CE), otherwise known as author of a poetic biography of the Buddha in Sanskrit, the Buddhacarita. Of Aśvaghoṣa's play Śāriputraprakaraṇa only fragments of the last two Acts (out of nine in total) are preserved. The story of the play concerns the conversion to the Buddhist doctrine of Maudgalyāyana and Śāriputra. Sanskrit, in prose and in verse, is spoken in this drama by the Buddha and his disciples, Maudgalyāyana and Śāriputra, and a Śramaṇa; the Vidūṣaka, who is a Brahmin, speaks Prakrit (Keith 1924: 82). The use of Sanskrit by one group of characters in classical drama and the use of Prakrit by another group has been taken as one of the reasons to accept the presence of diglossia in ancient India (Hock & Pandharipande 1976; Lee 1986, for whom, however, the hypersanskritisms are a stronger reason to accept diglossia). At this stage, several centuries after Aśoka, both the Sanskrit and the various Prakrits are literary languages, one having or symbolizing the highest prestige, the various Prakrits with somewhat lower or at least "different" prestige, and all at some distance from what was by that time, depending on the area, the widely spoken regional language or vernacular.

---

7 In the final discussion of the ISS seminar in 1994 it was, as far as I remember, Professor H.H. Hock who used this metaphor for the relationship between Sanskrit and a not always easily recognizable other language or other form of linguistic usage with which it interacts. That we are justified in distinguishing a dynamic interaction over time of a limited number of languages was recently demonstrated by Andrew Ollet, who further observed that "a dichotomy between Sanskrit and Prakrit" was "[a]t the foundation of this language order" of three literary languages in India mentioned by Mīrzā Khān in the 17th century (Ollet 2017: 1-4).

8 The expression sanskrit approximatif des bouddhistes was proposed by Helmer Smith (1954 : 3) as equivalent, or rather as a gentle, terminological corrective, to the title Buddhist Hybrid Sanskrit which F. Edgerton gave to his extensive study published in three volumes (Grammar, Dictionary, Reader) in 1953.

3 Sanskrit and the frequency of its varieties

The next question consists of two closely related ones, of which we will address here first the following.

b-A. How frequently do these languages/varieties occur?

Ancient India had very sophisticated techniques for memorization (Scharfe 2002, Houben & Rath 2012) and was very late in accepting writing for the transmission of its sacred texts, in comparison with its neighbours, China and Mesopotamia. After an initial, predominantly oral period, the various religious and philosophical systems accepted writing in an environment in which orality nevertheless remained predominant for a long time. Due to the conditions of the Indian climate and the properties of Indian manuscripts – mostly prepared from palm leafs or, in the north, of birch bark – they deteriorate after a relatively short period of two to three centuries and are to be copied if a subsequent generations considers their content sufficiently important. This situation has led to a very uneven distribution of quantitative manuscript survival (for some Vedic texts: oral tradition plus manuscript tradition). For older periods there are therefore no direct data available that allow us to answer the question "How frequently do these languages/varieties occur?" with precision, specifying place and time. In a limited domain such as epigraphy, some quantitative observations can be made. The oldest inscriptions, starting with those of Aśoka, are in an early Prakrit. They remain to be written in Prakrit, until the middle of the second century C.E., when the Śaka ruler Rudradāman had a text inscribed in perfect Sanskrit in which he "celebrates his own cultural and political achievements" (Pollock 2006: 68). In subsequent centuries, we find not only inscriptions in Sanskrit, but also in Prakrit, and in intermediate forms, for which the term Epigraphical Hybrid Sanskrit has been coined (Damsteegt 1978). The element "Hybrid" in this term would suggest that the language use reflected in the inscriptions is generated from separate and disparate linguistic entities or processes – for which, however, there is no evidence. "Hybridity" is as questionable here as it is in "Buddhist Hybrid Sanskrit" (Edgerton 1953), the Sanskritic language of a large body of Buddhist texts, and in "Gāndhārī Hybrid Sanskrit" (Salomon 2001). The tenacity of the "Hybrid" in reflection on ancient Indian language use is parallel to the hardly less tenacious, and hardly less problematic, conceptually underlying biological metaphor of languages as living organisms.[9] The inscriptions of both "Buddhist" and "non-Buddhist" character of the first till the fifth century in what appears to us as approximative (see above) or intermediate Sanskrit, or even approximative or intermediate Prakrit, show, however, that religious affiliation was not a decisive parameter in the choice of idiom. If different idioms are combined in a single inscription, a neat division is seen: the one of higher prestige is used in the *praśasti*, the part in which the king or another donor is praised, his genealogy given, his achievements celebrated; and the idiom of lower prestige is used in the management part which records in widely understandable terms the donation etc. Another, equally problematic, employment of the term "Hybrid Sanskrit" concerns the language used in a mathematical text – again in a context where religious affiliation is insignificant – which is fragmentarily available in a single manuscript, the Bakhshali manuscript, so called after the village where it was found in what is now north-west Pakistan.

More certainty in the establishment of the date of this and other undated manuscripts would be a great help in contextualizing the linguistic evidence for specific idioms and registers of language use in pre-modern South Asia, which would be a prerequisite for judging "how frequently" a language or register, including Sanskritic language use, occurs. The language is a quite particular one in the case of the Bakhshali manuscript. Scholars have recently again referred to this language uncritically as "Hybrid Sanskrit" in the recent article "The Bakhshālī Manuscript: A Response to the Bodleian Library's Radiocarbon Dating" (Plofker et al. 2017). The authors of this article respond to the dates presented by a research team of the

9 Aware of the problems surrounding the concept of "dialect" with reference to the "linguistic variability" of post-Vedic Sanskritic language, Salomon adopted the *contradictio in terminis* "vernacular Sanskrit" (Salomon 1989: 278; 1986; Deshpande 1993: 33ff) where "conversational Sanskrit" (cp. Śūdraka's *vyāvahārikā vāk*, referred to by Salomon 1989: 289) would have been more appropriate.

Bodleian Library as the results of radiocarbon dating of a few samples of the manuscript.[10] Plofker et al. rightly oppose the conclusion of the Bodleian Library research team that different parts of the manuscript are to be dated several centuries apart, one folio to the 3rd-4th century, another to the 7th-8th century, another to the 9th-10th century. Instead, they defend the results of earlier research which has shown that the handwriting is identical throughout the manuscript with a few exceptions (not affecting the selected samples for the radiocarbon dating) and that the script corresponds to that in use in the 12th (Kaye 1927), or the 7th to 12th century (Hayashi 1995). Plofker et al. explain and eloquently defend the conclusions reached by Hayashi in 1995, but do not reflect on how the divergent results may have emerged from the radiocarbon dating, so that their refutation remains incomplete. If the finding that samples of the same manuscript would be centuries apart is not based on mistakes in the procedure of sampling etc., or if the manuscript was at the moment it was written upon not partly consisting of older, recycled pages, there are still some factors that have evidently been overlooked by the Bodleian research team: the well-known divergence in exposure to cosmic radiation at different altitudes[11] and the possible variation in background radiation due to the presence of certain minerals in exposed, mountainous rock have nowhere been taken into account.[12] Among the variables of carbon dates, variation in script and linguistic variation, the first is the most objective but still much in need of calibration for relatively recent, historical dates. In view of the strong normativity of linguistic usage within the dimension "sanskrit - approximative sanskrit" it is difficult to derive a linear chronological difference from the observed linguistic variation. Also writing is a normative activity and moreover dependent on some amount of individual variation from scribe to scribe. However, writing has been much less subject either to the intensive study of early scripts by later generation scribes[13] or to the conscious reintroduction of archaisms in later forms of writing (something we see in language, most famously the studied archaizing "Vedic" language use in parts of the Mahābhārata and in the Bhāgavatapurāṇa). We therefore have to take quite seriously the judgement of palaeographists such as Richard Salomon who observed that, what he teleologically called "Proto-Śāradā," "first emerged around the middle of the seventh century" (Salomon 1998: 40). This excludes the earlier dates attributed to manuscript folio's on which a fully developed form of Śāradā appears. The "hardest" evidence to judge the date of a manuscript such as the Bakhshali and its sections would therefore be the palaeographic evidence.[14] Other evidence, including the laboratory results of radiocarbon dating, is to be interpreted in the light of the results reached by careful palaeographic study.

With regard to the question "How frequently do these languages/varieties occur?" we can conclude that much relevant material is available but that quantification is not obvious and rendered difficult in the absence of sufficiently reliable dating and contextualization of texts. For the period of India's pre-literary (pre-Aśokan) orality, only estimates can be proposed on the basis of indications found in early Vedic texts and in early Buddhist and Jaina texts which originated in that period. Another impediment is the unreflective use by modern scholars of outdated concepts, and, more generally, conceptual poverty in linguistic reflection about Sanskrit which was still excusable a century ago or even fifty years ago but not in modern times.

10 The results and rash interpretations were, prematurely, widely publicized in an article that appeared in the newspaper the Guardian (Devlin 2017), and on the website of the Bodleian Library.

11 Cp. the well-known high natural radiation level at Denver, at an altitude of around 1609 m (one mile) above sea level.

12 Since C-14 has a half-life of 5730 years, a minute deviation from average conditions (on account of a variation in background radiation) can be expected to lead to a considerable deviation in the resulting value. For instance, if one manuscript folio is produced from the bark of birch tree A found at altitude X and another folio from a different birch tree B at altitude Y, the two folios can be expected to show significantly divergent C-14 values even if they were produced and written upon contemporaneously.

13 By the 14th century, Indian pandits invited by Fīrūz Shāh Tughluq, were unable to decipher Aśokan inscriptions written a millennium earlier, even though the then current scripts had all descended from the Brāhmī script used in Aśoka's time (Salomon 1998 : 199f). But Vedic pandits, up to the present day, make an effort to recite Vedic hymns in the same manner as the earliest transmitters.

14 Where the dates are a matter of plus or minus one or more centuries according to different types of evidence and estimates, palaeographic evidence of a relatively current script is strong. However, it cannot give more precision than a period of one or more centuries, cf. Salomon 1998: 169.

4 Sanskrit and related varieties: wide apart or nearly indistinguishable?

The query regarding frequency of languages or language varieties (b-A) presupposes that these can be distinguished. Hence, it cannot be dissociated from the following question:

> b-B. Are the different linguistic systems easily distinguishable, or are some words difficult to assign to a specific language/variety?

Sanskrit and forms of Prakrit such as Aśokan Prakrit and Pali, in spite of numerous well-defined differences, are nevertheless still to be regarded as very close. They were no doubt to a large extent mutually understandable. When I made, on my way to the World Sanskrit Conference in Melbourne (January 1994), an intermediate stop in Bangkok, I visited a fellow-student from Pune who was a Buddhist monk in Bangkok and he insisted that I would meet his professor who was teaching Pali to his students. Unfortunately, this professor did not speak English and I did not speak Thai. We succeeded nevertheless to communicate, in front of his class, quite well, with him speaking Pali and me speaking Sanskrit.

How (a) Sanskrit and (b) its counterparts, Prakrit and numerous gradations of approximative Sanskrit, are different but also very close is clear from brief narrative passages in Patañjali's Mahābhāṣya, the extensive commentary on Pāṇini's grammar. We select here a few of these passages, first of all Bhāṣya (Section) 23 of the introductory chapter in the Mahābhāṣya which gives one out of several reasons why grammar should be studied:

> Those (rival) Asuras uttering (the words) *he'layo he'layaḥ* have been defeated. Therefore a brahmin must not speak barbaric language (*mlech-*), that is, he must not use corrupt words. *Mleccha* 'barbaric language' indeed is (the same as) *apaśabda* 'corrupt speech'. So that we should not become *mleccha* 'barbarians, users of barbaric language': that is (also a reason) why one should study grammar.[15]

The passage is roughly parallel to the Śatapatha-Brāhmaṇa in the Mādhyandina recension, 3.2.1.23-24, where it is part of a more extensive narrative starting at Śatapatha-Brāhmaṇa (Mādh.) 3.2.1.18. The Devas and Asuras,[16] the "divine counterparts of the vedic Aryans and their rivals" (Parpola & Parpola 1975: 212), are in fierce competition in the context of a ritual. After having lost the adherence of the goddess Speech (Vāc) who was initially at their side, the Asuras shout something. Instead of *he'layo he'layaḥ* of the grammarians, Śatapatha-Brāhmaṇa (Mādh) has *he'lavo he'lavaḥ*, which the medieval commentator Sāyaṇa glosses as *he'rayo he'rayaḥ*. Whether the exclamation was *he'layo he'layaḥ* or *he'lavo he'lavaḥ*, in both cases it would corresponds to Prakrit versions of Sanskrit *he'rayo he'rayaḥ* "hey, enemies!" (Thieme 1938).[17] Either way, the "barbaric language" of the Asuras would be very close to the high standard required by the gods, both in the narrative of the Śatapatha-Brāhmaṇa and in that of the grammarians. This remains valid even if the word *mleccha* has no Indo-European etymology, and is perhaps, together with Pali *milakkha* 'barbarian', a continuation of the toponym Meluḫḫa found in Akkadian and Old Babylonian cuneiform sources where it refers to a distant foreign country engaged in sea trade.[18] It also remains valid irrespective of whether the undeciphered symbols of the Indus Civilization are taken as representing a language,

15 MBh I:2.7-9 *te'surā he'layo he'laya iti kurvantaḥ parābabhūvuḥ / tasmād brāhmaṇena na mlecchitavai nāpabhāṣitavai / mleccho ha vā eṣa yad apaśabdaḥ / mlecchā mā bhūmety adhyeyaṁ vyākaraṇam*; cf. translation, analysis and discussion Joshi & Roodbergen 1986 : 38f. Cf. Joshi 1989.

16 An earlier stage in which neither *deva* nor *asura* is intrinsically "bad" or demonized is attested in the Ṛgveda but neither attested nor directly inferable in the Avesta (Herrenschmitt & Kellens 1993). A formula associating devas and mortals, however, was apparently Indo-Iranian (Swennen 2015).

17 In fact, as noted by Parpola & Parpola (1975: 212f), the formulation of the narrative in the Mahābhāṣya is closer to the one in the Kāṇva recension (4.2.1), except that it has the Asuras exclaim *hailo haila iti*, which they propose is corrupt for the exclamation he'layo he'layaḥ as known to the grammarians.

18 Neither can Skt. *mleccha* be directly derived from Pali *milakkha* nor the inverse, but both may represent a non Indo-european Meluḫḫa: Parpola 1994: 170; 2015: 215ff; Parpola & Parpola 1975.

usually either a (proto) Dravidian one (Parpola 1994, 2015),[19] or an early Indo-Aryan one,[20] or as not being in any way linguistic (Farmer et al. 2004). Whatever the linguistic reality in terms of languages from very different language families, the Brāhmaṇa and Mahābhāṣya authors perceived only the variation between Sanskrit and gradations of approximative Sanskrit. If we assume that some (proto-) Dravidian language was, in their time, somehow geographically near, its presence apparently remained largely unperceived. That the distinction between correct speech and reprehensible, incorrect speech was minute is also clear from the discussion of the Mahābhāṣya passage by the early sub-commentator Bhartṛhari (MBhD I:8.14-18). Bhartṛhari asks: what is in fact faulty in the exclamation of the Asuras? Apparently, this is not evident. He provides four grammatical answers, each of them implying a minute deviation from the correct form, one of them being that an *-r-* has become an *-l-* (which will in later times characterize the literary Prakrit Māgadhī). The main point that emerges is "that the terms *mleccha* and *apaśabda* are used with reference to Prakrit words" (Joshi 1989: 268) and that the correct and incorrect words are really very close, and do not involve any other language than what we understand as early varieties of Indo-Aryan, even if languages of other linguistic families must have been spoken by minorities or communities near by.

The next passage in the Mahābhāṣya, Bhāṣya 24, gives another example of incorrect use of language:

> A word which is defective on account of the accent or on account of a phoneme, which is incorrectly used, does not convey that (the intended) meaning. (This defective word) as a thunderbolt in the form of a word, brings damage to the sacrificer, just as the word *índra-śatru* (with accent on the first syllable which makes it an exocentric compound, instead of *indra-śatrú* which would have conveyed the desired meaning of "enemy or killer of Indra").[21]

The first detailed explanation of this verse and of the example is neither in the MBh nor in any of the earlier Vedic texts[22] which tell the story of the conflict between Asura Tvaṣṭar and god Indra, but in Bhartṛhari's commentary on the Mahābhāṣya, MBhD I:9.1-15. Bhartṛhari makes clear that *índra-śatru* was pronounced with udātta on the first syllable and is hence an exocentric compound (bahuvrīhi) meaning "he whose enemy or killer is Indra"; however, the speaker obviously intended a nominal compound of the tatpuruṣa type, which should have had the udātta on the final syllable (*antodātta*), meaning: "enemy or killer of Indra". Asura Tvaṣṭar apparently pronounced all formulas in accented Sanskrit correctly, except for the accent of one syllable, which led to an opposite result: the son for the birth of whom he prayed would be killed by Indra, rather than becoming the killer of Indra.

Still another example of incorrect use of language is given in the Mahābhāṣya, Bhāṣya 119:

> There were sages (a group of sages) (nick-) named *yarvāṇas-tarvāṇas*. Their perception of *dharma* was direct, they knew the far and the near, they knew what could be known and they had come to realize ultimate reality. These worthy persons used the expressions *yarvāṇastarvāṇas* when they should have used *yad vā naḥ tad vā naḥ* "whatever (happens) to us, (let) that (happen) to us." Still, they did not use incorrect words at the time of sacrificial ritual. But the Asuras did use incorrect words at the time of sacrificial ritual. That is why they were defeated.[23]

In this example the situation is the inverse of the preceding two: a group of sages uses here correct language within the ritual and wrong language in daily life with distortion or wrong euphonic combination of a few syllables (*r* instead of *d*, *ṇ* instead of *n*: an excess of cerebralisation).

---

19 Parpola interprets a number of signs but does not claim to have deciphered the script.

20 Various "decipherments" of the Indus script as representing either a Dravidian or an Indo-Aryan language have been reviewed by G. Possehl (1996) and I. Mahadevan (2002); Kalyanaraman (2016), through an ingenious and extremely flexible "rebus" interpretation, tries to read the Indus symbol sequences as multilingual messages mainly of artisans.

21 MBh I:2.11-12: *duṣṭaḥ śabdaḥ svarato varṇato vā mithyā prayukto na tam artham āha / sa vāgvajro yajamānaṁ hinasti yathendraśatruḥ svarato 'parādhāt*. Cf. translation, analysis and discussion Joshi & Roodbergen 1986 : 39ff.

22 ŚB 1.6.3.8; other versions of the story in TS 2.4.12.1, 2.5.2.1; MaitrS 2.4.3c: 40.5; JB 2.155.

23 MBh I:11.11-14: *yarvāṇastarvāṇo nāmarṣayo babhūvuḥ / pratyakṣadharmāṇaḥ parāparajñā viditaveditavyā adhigatayāthātathyāḥ / te tatrabhavanto yad vā nas tad vā na iti prayoktavye yarvāṇastarvāṇa iti prayuñjate / yājñe karmaṇi punar nāpabhāṣante / taiḥ punar asurair yājñe karmaṇy apabhāṣitam / atas te parābhūtāḥ*. Cf. Joshi & Roodbergen 1986 : 156ff. Cf. Joshi 1989.

Since the ancient Indians had a highly developed system of grammatical analysis and their reflection on language and grammar was on a high level, it is legitimate to ask what was their own view on the distinguishability or difference or closeness of the two linguistic structures, Sanskrit (or its predecessor, Patañjali's *bhāṣā*) and Prakrit.

At the end of book 1 of the Vākyapadīya (VP 1.175-183) the relation between correct and substandard words and their capacity to express meaning are discussed. The VP-verses of this passage envisage two situations:

I. The speaker sincerely tries to speak correct language (*śabda*), but produces substandard words (*apabhraṁśa*).

II. The speaker is in a community in which the substandard *apabhraṁśa* words have become generally known and accepted on account of a (non-Sanskrit, Prakrit) tradition.

Under (I), the correct word, *śabda*, is *vācaka* 'expressive of meaning'; the substandard *apabhraṁśa* word is not itself *vācaka* 'expressive of meaning', but it brings to mind the intended correct word, *śabda*. Under (II), *śabda*, the correct word, is not or no more expressive of meaning: in the (non-Sanskrit, Prakrit) community it is the Prakrit word that has become directly expressive of meaning. In this regard, Patañjali's Mahābhāṣya and Bhartṛhari's Vākyapadīya explicitly accept that *śabda* and *apaśabda* or *apabhraṁśa*, the "correct" and the "incorrect" word, can be equally expressive (MBh 1:8.21 *samānāyām arthagatau śabdena cāpaśabdena ca*; VP 1.27 *arthapratyāyayanābhede*; 3.3.30, *asādhur … vācakatvāviśeṣe vā*). The relevant passages have been discussed in detail in Houben 1995: 237-242 and 1997: 336-341, where a difference in orientation was demonstrated between the verses of the Vākyapadīya (by Bhartṛhari) and the ancient Vṛtti, or, more precisely, the longer ancient Vṛtti (*bṛhatī* to distinguish it from the, in significant respects different, *laghuvṛtti*).[24] This longer ancient Vṛtti holds that substandard words can never directly express their meaning, not even in communities where these substandard words have become well-established. The *laghuvṛtti*, however, emphatically accepts what is evidently also the view set forth in the Mahābhāṣya and in Bhartṛhari's verses of the Vākyapadīya: the more liberal view that the correct and the substandard incorrect word can be equally expressive. Where the longer Vṛtti entirely neglects the statement in the kārikā that there is no difference between correct and incorrect words in expressing their meaning (*arthapratyāyayanābhede*), the *laghuvṛtti* provides at this place an explicit explanation. The main point is here that in the case of Prakrit, or, in Bhartṛhari's terms, in the case of substandard words (*apabhraṁśa*, corresponding with what we would call Prakrit words), the boundary with what we would call Sanskrit is extremely fluid: it is the individual words that are substandard, there is no systemic or structural change from the language "Sanskrit" to a language "Prakrit" as is the consistent perspective (and aim of reconstruction) of modern scholars of Sanskrit and Middle Indo-Aryan.

On a theoretical level, Bhartṛhari's position, according to his own statements as found in his *magnum opus*, the Vākyapadīya (VP), corresponds to the "hocus-pocus" position rather than to the "God's truth" position.[25] On the basis of the oft-cited words of Sir William Jones (1786): "The Sanskrit language, whatever may be its antiquity, is of a wonderful structure", and on the basis of Saussure's view that Sanskrit is an

24 The ancient Vṛtti is sometimes directly attributed to Bhartṛhari but is rather the work of a student or close follower of his, perhaps named, as colophons of several VP-manuscripts indicate, Harivṛṣabha, "(Bhartṛ-)Hari's bull" (i.e., the favourite one or best among his students). Where the long and short version are different it is fair to assume that the one corresponding more closely to the bare statements in the kārikās contains the original, unmodified or less modified version of the commentary.

25 On the "God's truth" vs. "hocus pocus" controversy in linguistics: Householder 1952: 260 (review of Harris 1951); Burling 1964; Houben 1993; Peeters 2001. For Peeters, the 'God's truth' vs. 'hocus-pocus' controversy acquires an entirely new meaning in the light of unprecedented possibilities to take into account data of neuroscience. It is hence feasible and, for Cognitive Linguistics as Peeters wants to see it, desirable, to adopt a 'God's truth' position in a new, extended sense of the word, as dealing with psychological *and* with biological reality, the mind as well as the brain. Burling's position as expressed in his 1964 article, on the contrary, was that structural semantics does not of itself reveal the cognitive system of the speakers (which would correspond to a 'God's truth' position), but that it can provide a set of rules (corresponding to a 'hocus-pocus' position) which account for the way terms are used by speakers of the language : a pragmatic position closer to the one adopted by the ancient Indian grammarians.

"ultra-grammatical" (1916: 183), that is, for Saussure, an extremely systematic, language, one should rather have expected that Bhartṛhari would accept and deal with the presence of "real structures" in Sanskrit. The presence of such structures, however, is precisely what Bhartṛhari emphatically denies (Houben 1993, 2009). His position on *apabhraṁśa* or our "Prakrit" is part and parcel of this denial of a given structure in language: it is the individual words that are substandard, incorrect or simply different. We may add that the differences between Sanskritic and Prakrtic words referred to by the grammarians were relatively small, as they are often a matter of changing one or two phonemes. Closeness between an early Prakrit and (later, classical) Sanskrit (or, contemporaneously, conversational old Indo Aryan, *bhāṣā*) is also what we see in Aśoka's inscriptions. This is in a period which is still predominantly oral, with Aśoka pioneering royal inscriptions, and with some kind of writing probably in use in ephemeral contexts but not for sacred or philosophical texts (Houben & Rath 2012). In view of the structural and lexical closeness briefly illustrated here, it is hardly justified to speak of different "languages" or even of different "dialects" with regard to the idioms of "old" and "middle" Indo-Aryan which were contemporaneously in active use at the time of the early grammarians and Aśoka.

Several centuries later, when Sanskrit and several Prakrits have become literary languages, skilled poets are able to write verses which can be read in Sanskrit and one or more Prakrits at the same time, as in the Devīśataka of the 9th century poet Ānandavardhana, verse 74 of which is as follows: *alolakamale cittalalāmakamalālaye / pāhi caṇḍi mahāmohabhaṅgabhīmabalāmale //* "O you whose prosperity is not unstable, residing in the eminent lotus which is the mind, protect (us), o Caṇḍī (passionate one), you who are pure on account of your formidable power to destroy the grand delusion." The commentator Kayyaṭa explains that in this verse there is co-occurrence of six "languages": Sanskrit, Mahārāṣṭrī, Paiśācī, Māgadhī, Śaurasenī, Apabhraṁśa (*saṁskṛta-mahārāṣṭra-piśāca-māgadha-sūrasenāpabhraṁśātmikāḥ ṣaḍ api bhāṣā atra samāviṣṭā yadā tadā bhāṣāṣaṭkasamāveśo 'yam*).[26] This remarkable feat can be achieved precisely because of the closeness of Sanskrit and the various Prakrits, and next by a clever choice of words and by leaving out "difficult" consonant clusters in Sanskrit such as *rm-*, which corresponds to *mm-* in Aśokan and other Prakrits, and *kt-* and *tr-* which both correspond to *tt-*. In another employment of the poetic figure of *bhāṣāśleṣa* 'embrace of languages' the two statements in the intertwined languages are different, and require a different division of words in the unitary expression (Lienhard 1984: 137; Hahn 2012). In view of the structural and lexical closeness briefly illustrated here, it is hardly justified to speak of different "languages" or even of different "dialects" with regard to the idioms of literary Sanskrit and the literary Prakrits.

With this result we may go back to the "great linguistic paradox" of Louis Renou: the fact that "Middle Indian (Middle Indo-Aryan) makes its appearance in epigraphy prior to Sanskrit" (Renou 1956: 84). Renou links this to the choice of the Buddha, two centuries before Aśoka, to impart his teaching, at the basis of all later Buddhist doctrine, in Middle Indo-Aryan, "a Māgadhī or pre-Māgadhī dialect" (ib.). Similarly, Mahāvīra had decided to impart his teaching, at the basis of all later Jaina doctrine, in Middle Indo-Aryan. King Aśoka, inspired by and converted to Buddhism, would therefore have ordered his inscriptions to be in Prakrit dialects as well: this would have remained the habit for inscriptions for centuries to come. For the Buddha's choice to teach in a Middle Indo-Aryan dialect, Renou refers, in the next paragraph, to a well-known narrative found in various Buddhist canonical texts according to which two monks, converted Brahmins, propose to put the discourses of the Buddha into *chandas*. In terms of linguistic knowledge available in the Buddha's time this can only mean: to transpose them into a text with Vedic metre and Vedic accents (in accordance with the phonetic, grammatical and metrical rules of some early Prātiśākhya-treatise). The Buddha rejects the proposal, and encourages the monks, on the contrary, to transmit the speech of the

26 Devīśataka by Ānandavardhana (9th cent.) in Śivadatta & Paṇaśīkar 1916 (*Kāvyamālā* pt. 9) : 20-21. I take the compound *alolakamale* as a bahuvrīhi-compound, as Hahn (2012: 81: "deine Schönheit ist nicht schwankend"), but the commentator Kayyaṭa takes it as karmadhāraya, *alolā acapalāpi tvaṁ kamalā lakṣmīḥ alolakamalā*; the compound *citta-lalāma-kamala-ālaye* is taken by Hahn (ibid.) as "du Wohnsitz für den Zierdelotus 'Geist'," whereas I prefer to follow here commentator Kayyaṭa's more likely interpretation as exocentric compound (bahuvrīhi): *cittam eva lalāmakamalaṁ pradhānapadmaṁ tad evālayaḥ kulāyo yasyāḥ*. On Ānandavardhana's Devīśataka also Ingalls 1989.

Buddha *sakāya niruttiyā*, i.e., "in one's own mode of expression" (ib.).[27] Retrospectively, scholars have read in this story the rejection by the Buddha of the use of Sanskrit for his teaching. However, as no Sanskrit, in the strict sense of the term, can have been available in the Buddha's time as an identifiable linguistic option for communication, there can have been no rejection of this not yet existing Sanskrit by the Buddha. The passage is, moreover, clear in specifying that the rejection concerned *chandas*, which was at that time indeed an identifiable linguistic option, not so much for colloquial communication but for perpetuating a teaching. In a recent, extensive and brilliant analysis of several versions of the *sakāya niruttiyā* passage according to canonical texts of various Buddhist schools, Vincent Eltschinger has justly drawn again attention to the interpretation of this passage in two schools whose canons are not in Pali but in Sanskrit: the Sarvāstivādins and the Mūlasarvāstivādins (Eltschinger 2017: 315f[28]). The relevant passages of these schools, unfortunately available only in Chinese translation, clearly imply the rejection by the Buddha not of Sanskrit but of the adoption of metrical chanting and intonation of *chandas* for the transmission of the Buddha's teaching. This interpretation, which equally suits the well-known Pali version quoted and discussed for instance by Edgerton in 1953,[29] is not the result of an adjustment to the Sarvāstivādins' and the Mūlasarvāstivādins' use of Sanskrit as the language of their canons: it reflects the generosity of the Buddha's allowance to his monks to teach *sakāya niruttiyā*, "in one's own mode of expression," which should have included a whole range of Sanskritic and Prakritic language use, comprising also any predecessors of classical Sanskrit available in his time, which, as we have seen, were anyway very close to each other and to a very large extent mutually understandable.

If, however, there is no indication that the Buddha would ever have rejected Sanskrit as an available, linguistic option, the apparent "adoption" of Sanskrit by later generations of Buddhists necessarily appears in an entirely different light as well: this was then rather a matter of relative strength and growth of Buddhist communities or sects that were prone to accepting and developing Sanskritic language (grammar) and literature. The important and foundational contributions to the development of Sanskrit literature and grammar in the early centuries CE – e.g. by Aśvaghoṣa, mentioned above, by the Buddhist grammarian Candrācārya, referred to in Bhartṛhari's Vākyapadīya, and by the lexicographer Amarasiṁha – are then no more betrayals to a linguistic choice of the Buddha, but legitimate explorations of one of the available options of language use, originally perhaps a minority option, left open by the Buddha. Nor can the choice of words by King Aśoka in his inscriptions, around two centuries after the Buddha, be understood as the choice of a language, Prakrit, against another language, Sanskrit or a predecessor, when the concept of language as an identifiable entity that can be chosen or rejected did not even exist. In the linguistic situation of the time, neither the Buddha nor the Mahāvīra nor Aśoka had any other choice when they wanted to address a large public in the Indian, Indo-Aryan realm. Even Vedic ritualists were speaking some Middle Indo-Aryan or Prakritic idiom outside the ritual sphere, as is clear from the story of the *yarvāṇastarvāṇa* sages in Patañjali's Mahābhāṣya (see above). The language form which we know as "classical" Sanskrit was not yet existing because it still had to be co-produced by the Buddhists.

Renou's linguistic paradox is therefore to a large extent based on an optical illusion, a *trompe-l'œil*, as Renou himself to some extent realized (Renou 1956: 84) if we accept, on the one hand, that language options in the Indo-Aryan realm, from pre-Vedic times onwards, included a range of contemporaneous linguistic forms to which different levels of prestige were attached; and, on the other hand, that the different varieties were actually extremely close and to a large extent mutually comprehensible. The linguistic situation in ancient India evoked, to Louis Renou, German Switzerland, "where the normal means of communication is the dialect, and nevertheless German has the position of a spoken language" (Renou 1956: 87). A few years later, Swiss German would be one of the defining languages in Ferguson's definition of diglossia (1959), next to Arabic/Egyptian Arabic, Haitian Creole and Greek. The situation in India as reflected in literature

27 Renou (1956: 84) translates first, appropriately, "dans son mode d'expression propre," but introduces next the concept which is precisely most problematic in this context when, in a footnote, he further explains it as "dans son dialecte propre."

28 See also the earlier discussions of these passages by Sylvain Lévi 1915 and Lin Li-Kouang 1949 referred to by Eltschinger.

29 Edgerton, 1953 :1, correctly renders *chandaso āropema* quite literally as "put into Vedic." Both Chinese Buddhist interpreters and modern interpreters have tried to read in this statement a proposal, rejected by the Buddha, to express the word of the Buddha in Sanskrit: see further references and detailed discussion by Eltschinger.

and in the grammarians' examples and analyses, fully applicable at least in the area defined as Āryāvarta (according to the Manusmṛti between the Himalayas and the Vindhya mountains and between the eastern and western sea[30]), has indeed a remarkable parallel in the coexistence and interpenetration of High German and Schweizerdeutsch in Switzerland and matches the classical definition of diglossia, as demonstrated in Houben 1996a. The non-Indo-Aryan languages that must have been spoken by some communities in that realm apparently remained under the threshold of perception. The "extreme superposition" perceived by Pollock (2006: 50) and, in different terms, by Robert (2012), refer to a clearly distinct situation where Sanskrit, or, in Japan, Chinese, is incorporated into cultural and linguistic life in Karnataka, resp. in Japan, while all actors are and remain sharply aware of their otherness and distant origin.

5 Conclusion

A new exploration of the language situation in ancient India at the time of the Buddha and the early grammarians is required in terms of the emergence of dialects, sociolects, and new languages, with as contrasting parallel the contemporaneous development in ancient Persia as reflected in old Persian inscriptions and in the Avesta. The following methodological observations have been made regarding fundamentally different ways of seeing languages.

> Scholars are recognizing that languages are not always easily nor best treated as discrete, identifiable, and countable units with clearly defined boundaries between them .... Rather, a language is more often comprised of continua of features that extend across time, geography, and social space. There is growing attention being given to the roles or functions that language varieties play within the linguistic ecology of a region or a speech community. ... Languages can be viewed, then, simultaneously as discrete units (particles) amenable to being listed and counted, as continua of features across time and space (waves) that are best studied in terms of variational tendencies as examples of 'change in progress', and as parts of a larger ecological matrix (field), where functional roles and usage of the linguistic codes for a wide range of purposes are more in focus. (Lewis 2009)

Madhav Deshpande (2006: 141) rightly explained that

> [t]he notion of language family implies that languages B and C are branchings of a common ancestor A, and this fact of a genetic connection accounts for certain features. On the other hand, the notion of a linguistic area implies that languages A and B, though belonging to different language families and originally possessing different linguistic features come to share some of each others features over a long period of time through intense contact.

Here too, the 'linguistic area' model (in which languages appear in a 'field'[31]) is superimposed on a 'family' model (in which languages are discrete units generating new units over time). However, the latter's priority cannot always and everywhere be taken for granted.

Extensive researches since the 19th century suggest that within the period that interests us, from 1000 BCE to 1000 CE, Old Persian, Avestan, Vedic, Middle Indo-Aryan and classical Sanskrit evolved within a large area of Indo-Iranian dialect continuity (Meillet 1908: 24-30), from 'linguistic area' to 'linguistic area', with several shifts of the geographical point of gravity, from Persepolis to Gandhara and from the northwest of the Indian subcontinent to the central Gangetic plain, and to India's southern states (the Deccan and further south). Apart from "time" and "geography", it is indispensable to take into account a third parameter throughout this period and throughout the large area of the partly overlapping Iranian and Indian "worlds": the parameter of sociolinguistic variation between a pole of high prestige characterized by elaboration and

30 Subsequent descriptions of this area (esp. those in the Mahābhāṣya and in the Manusmṛti) point to an ecological transformation (from still largely forested area suitable to agro-pastoralism to an urbanized environment), which goes hand in hand with major transitions in ritual and religion (from Vedic to Buddhism): Houben 2011.

31 For India as a linguistic area see Emeneau 1956, Kuiper 1968 ; also Ollet 2017 can be regarded as a contribution to this domain of research.

sophistication (for instance in a "Dichtersprache"[32]), and a pole of lower prestige characterized by easiness of access and solidarity. The term *āriya / ārya* 'noble' as a qualifier of speech is occasionally attested in this large area, in a multilinguistic context (in the multilingual Behistun inscription[33]) in the Iranian part, and in a diglossic context (ŚĀ, passage on the language spoken in assemblies[34]) in the Indian part. Almost contemporaneously with these employments but far in the east, the Buddha proposed an ethical focus or reinterpretation of *ārya* 'noble' with 'nobility' being dependent on behaviour and effort, and independent of acceptance of a hereditary 'nobility'.

Under some conditions it may be appropriate to attribute primary status to a model of "family" relationships between languages as "particles" or as discrete units, for instance with regard to languages that survive and remain relatively stable in mountainous areas.[35] For languages that flourish in areas of intensive contacts a simple genetic model may be entirely inadequate, especially when the 'linguistic area' applies also to what can be reconstructed for earlier periods (cf. Pinault 2002). When studying emerging languages such as the early stages of classical Sanskrit and literary Prakrit, these should obviously not be posited as discrete units. Invoking the concept of "hybridity" in connection with the name of a well-defined language (as in "Buddhist Hybrid Sanskrit," "Epigraphical Hybrid Sanskrit") was only a stopgap solution when dealing with languages or idioms that were emerging, successfully or without lasting success, as standards or as roofing language, or that were disappearing. The study of the emergence and disappearance of new standard languages is currently a large field of study to which sociolinguistics has contributed significantly in recent decades. Concepts used with regard to the evolution of new standards in Germanic (Goossens 1985) or Romance languages (Muljačič 1986, 1989, 1993) can and, for a better scientific grasp on the subject, should be applied and tested with regard to Indo-Iranian, Indo-Aryan and Sanskrit. The so-called Sanskrit "Hybrids" in the first millennium CE, including the Prakrits and Epic Sanskrit from the time of Aśoka onwards, are then rather to be regarded as emerging "Ausbau" languages of Indo-Aryan with hardly any significant mutual "Abstand" before they will be successfully "roofed," in the second half of the first millennium CE, by "classical" Sanskrit, for which Pāṇini and Patañjali, filtered by the work of Buddhist grammarians such as Candrācārya (contributing, *inter alia*, to a definitive abandonment of linguistic accent and of the subjunctive), will become authoritative. The appropriate question to ask with regard to Pāṇini "as a variationist" and his period would then not just be: what was "actual Sanskrit usage" giving the "best possible fit" with the rules (Kiparsky 1979: 5-6), but rather what was the diglossic range within which he and the intended public of his grammar were functioning.

Abbreviations

AA = Aṣṭādhyāyī of Pāṇini. (a) ed. and tr. : Otto Böhtlingk, *Pāṇini's Grammatik, herausgegeben, übersetzt, erläutert und mit verschiedenen Indices versehen*, Leipzig: Haessel, 1887. (Réim. Hildesheim, Olms, 1964, etc.) (b) ed. and tr.: Sumitra M. Katre, *Aṣṭādhyāyī of Pāṇini*, Delhi: Motilal Banarsidass, 1989; (c) ed. and tr. (into French by Louis Renou, complete; partial tr. into English), crossref. and research tools: *Ganakakastadhyayi: A Software on Sanskrit Grammar based on Pāṇini's Sūtras* by Shivamurthy Swamiji, downloadable at www.taralabalu.org.
AiĀ = Aitareya-Āraṇyaka: ed. and tr. Keith 1909.

32 For poetic aspects and aspects of world view and philosophy of an Indo-european poets' language or Dichtersprache, see: Schmitt 1967, Watkins 1994, Pinchard 2009 ; the subject awaits further exploration from a sociolinguistic perspective.

33 OP part of Darius' Behistun (Bīsutūn) inscription, column 4 line 89: *ariyā*, apparently used with reference to the language of this (Old Persian) part of the inscription : Schmitt 2009 : 87.

34 ŚĀ 8.9 explains the success accruing to someone having a certain esoteric knowledge: *śuśrūṣante* [read thus] *hāsya parṣatsu bhāṣyamāṇasyedam astu yad ayam īhate yatrāryā vāg vadati vidur enaṃ tatra* "men want to listen to him when he speaks in the assemblies; (they say) 'this should be done when he desires it'; where *āryā* speech is uttered (or: where speech is sounding noble) they know him there." (Keith 1909: 314-315; 1908: 55).

35 As for instance the Himalayan languages investigated by G. van Driem and his team: van Driem 2001; on the Basque language surviving in a mountainous region at the foot of the Pyrenees in France and Spain: Morvan 1996, Allieres 1998, Vennemann 2003.

Av = Avestan
JB = Jaiminīya-Brāhmaṇa: ed. Raghu Vira and Lokesh Chandra, Sarasvati Vihara Series 31, 1954; Second rev. edition: Delhi-Varanasi-Patna, 1986; partly ed. and tr. W. Caland, *Das Jaiminīya-Brāhmaṇa in Auswahl*, Amsterdam Academy 1919.
MaitrS = Maitrāyaṇī-saṁhitā: ed. L. von Schroeder, Leipzig 1881-1886.
MBh = Patañjali's Vyākaraṇa-Mahābhāṣya, ed. F. Kielhorn, third ed. rev. by K.V.Abhyankar, Pune 1962-1972: ref. to volume, page, line.
MBhD = Bhartṛhari's Mahābhāṣya-Dīpikā, fasc. 1: ed. and transl. J. Bronkhorst, Pune 1985: ref. to Āhnika, page and line.
OP = Old Persian
ṚV = Ṛgveda: Ṛg-Veda-Saṁhitā: ed. F.M. Müller, saṁhitā and pada texts (2 vols). Third ed. Varanasi, 1965; tr. Jamison & Brereton 2014.
ŚĀ = Śāṅkhāyana-Āraṇyaka: ed. (ŚĀ VII-XV) Keith 1909, Appendix; tr. Keith 1908.
ŚB = Śatapatha-Brāhmaṇa: Śatapatha-brāhmaṇa with (Sāyaṇa's) Vedarthaprakash commentary, ed. by several learned persons. Kalyan-Bombay : Laxmi Venkateshwar Steam press, Saṁvat 1997 / San 1940. (The Bhāṣika-sūtra in vol. 5: 300-320); transl. Julius Eggeling, Sacred Books of the East 12, 26, 41, 43, 44 (Oxford, 1882-1900).
Skt. = Sanskrit
TS = Taittirīya-Saṁhitā: ed. A. Weber in *Indische Studien* vol. 11-12, 1871-1872; tr. Keith 1914.
VP = Bhartṛhari's Vākyapadīya: ed. of the mūla-kārikās by W. Rau, Wiesbaden, 1977; ed. (Kāṇḍa I with Vṛtti and Paddhati) by K.A. Subramania Iyer, Pune, 1966; ed. (Kāṇḍas I and II with commentaries) by Gangadhara Sastri Manavalli, Benares 1887.

References

Allieres, Jacques. 1998. Review of Morvan 1996. *Lapurdum: Basque studies review, 1998 Numéro III.IV recensions, informations*: 315-317.
Bloch, Jules. 1950. *Les Inscriptions d'Aśoka, traduites et commentées*. Paris: Société d'édition Les Belles Lettres.
Burling, Robbins. 1964. "Cognition and Componential Analysis: God's Truth or Hocus-Pocus." *American Anthropologist* 66: 20-28.
Calvet, Louis-Jean. 1998. *La sociolinguistique*. 3e édition corrigée. (Série: Que sais-je?) Paris: Presses Universitaires de France.
Cohen, Marcel. 1956. *Pour une sociologie du langage*. Paris: Albin Michel.
Damsteegt, Theo. 1978. *Epigraphical Hybrid Sanskrit: its Rise, Spread, Characteristics and Relationship to Buddhist Hybrid Sanskrit*. Leiden: E.J. Brill.
De, Sushil Kumar. 1960. *History of Sanskrit Poetics*. Second rev. edition. Calcutta: Mukhopadhyay.
Deshpande, Madhav M. 1985. "Sanskrit grammarians on Diglossia." In: *South Asian Language: Structure, Convergence and Diglossia* (ed. by Bh. Krishnamurti): 312-321. Delhi: Motilal Banarsidass.
Deshpande, Madhav M. 1993. *Sanskrit and Prakrit*. Delhi: Motilal Banarsidass.
Deshpande, Madhav M. 2006. "Aryan origins: Brief History of the Linguistic Arguments." In: *India: Historical Beginnings and the Concept of the Aryan* (by R. Thapar, J.M. Kenoyer, M. Deshpande, S. Ratnagar): 98-156. New Delhi: National Book Trust.
Devlin, Hannah. 2017. "Much ado about Nothing: Ancient Indian Text Contains Earliest Zero Symbol." *The Guardian*.
Driem, George van. 2001. *Languages of the Himalayas* (2 vols.). Leiden: E.J. Brill.
Durkheim, Émile. 1895. *Les règles de la méthode sociologique*. (First appeared in *Revue philosophique* in the form of four articles in 1893 vols. XXXVII May and June ; and in 1894 XXXVIII July and August) Paris: Félix Alcan.
Edgerton, Franklin. 1953. *Buddhist Hybrid Sanskrit, Grammar and Dictionary*. Vol. I (Grammar), Vol. II (Dictionary). New Haven: Yale University Press.
Eltschinger, Vincent. 2017. "Why did the Buddhists adopt Sanskrit?" *Open Linguistics*; 3: 308-326.
Emeneau, Murray. 1956. "India as a linguistic area." *Language* 32: 3-16.
Farmer, Steve, Richard Sproat, Michael Witzel. 2004. "The Collapse of the Indus-Script Thesis: The Myth of a Literate Harappan Civilization." *Electronic Journal of Vedic Studies*, 11-2: 19-57.
Ferguson, Charles A. 1959. "Diglossia." *Word* 15: 325-340.
Goossens, Jan. 1985. "Herauslösung und Herausbildung des Niederländischen." *Entstehung von Sprachen und Völkern: Glotto- und Ethnogenetische Aspekte europäischer Sprachen* (ed. P.S. Ureland): 287-304. Tübingen: Max Niemeyer.

Hahn, Michael. 2012. "Der *Bhāṣāśleṣa* – eine Besonderheit kaschmirischer Dichter und Poetiker?" In: *Highland Philology: Results of a Text-Related Kashmir Panel at the 31st DOT, Marburg 2010* (ed. R. Steiner): 77-105.

Harris, Zellig S. 1951. *Methods in Structural Linguistics.* Chicago: Chicago Univ. Press.

Hayashi, Takao. 1995. *The Bakhshālī Manuscript: an Ancient Indian Mathematical Treatise*. Groningen: Egbert Forsten.

Herrenschmitt, Clarisse & Jean Kellens. 1993. "*Daiva." *Encyclopaedia Iranica*, 6: 599-602.

Hinüber, Oskar von. 1986. *Das ältere Mittelindisch im Überblick.* Wien: Österreichische Akademie der Wissenschaften. [See now also: id., 2. erweiterte Auflage. Wien: Österreichische Akademie der Wissenschaften, 2001.]

Hock, Hans Heinrich, and Rajeshwari Pandharipande. 1976. "The sociolinguistic position of Sanskrit in Pre-Muslim South-Asia." *Studies in Language Learning*, vol. 1, no. II: 105-138.

Houben, J.E.M. 1993. "Bhartṛhari en de structuur van het Sanskrit: 'God's truth' of 'hocus pocus'?" [in Dutch] In: *Nieuwe wegen in taal- en literatuurwetenschap: Handelingen van het Eenenveertigste Filologencongres* (Proceedings of the 41st Congress of Philologists, Tilburg, 15 December 1990): 143-158. Tilburg: Tilburg University Press.

Houben, Jan E.M. 1995. *The Saṁbandha-samuddeśa (chapter on relation) and Bhartṛhari's philosophy of language*. Groningen: Egbert Forsten.

Houben, Jan E.M. 1996a. "Socio-linguistic attitudes reflected in the work of Bhartṛhari and later Grammarians." In: *Ideology and Status of Sanskrit* (ed. J.E.M. Houben): 157-193. Leiden: E.J. Brill, 1996.

Houben, Jan E.M. 1996b. "Introduction: Towards a Social History of the Sanskrit Language." In: *Ideology and Status of Sanskrit* (ed. J.E.M. Houben): 1-13. Leiden: E.J. Brill, 1996.

Houben, Jan E.M. 1997. "Bhartṛhari's Perspectivism (1): The Vṛtti and Bhartṛhari's perspectivism in the first Kāṇḍa of the Vākyapadīya." In: *Beyond Orientalism: the impact of the work of W. Halbfass on Indian and cross-cultural studies* (ed. by K. Preisendanz and E. Franco): 317-358. Amsterdam: Rodopi, 1997.

Houben, Jan E.M. 2008. "Bhaṭṭoji Dīkṣita's 'small step' for a Grammarian and 'Giant Leap' for Sanskrit Grammar." *Journal of Indian Philosophy* 36 (2008) : 563-574.

Houben, Jan E.M. 2009. "Bhartṛhari as a 'cognitive linguist'." In: *Bhartṛhari: Language, Thought and Reality* (Proceedings of the International Seminar, Delhi, December 12-14, 2003) (ed. by M. Chaturvedi) : 523-543. Delhi: Motilal Banarsidass.

Houben, Jan E.M. 2011. "Vedic ritual as medium in ancient and pre-colonial South Asia: its expansion and survival between orality and writing." In: *Travaux de Symposium International "Le Livre. La Roumanie. L'Europe." Troisième édition – 20 à 24 septembre 2010, Section IIIA: Veda-Vedāṅga et Avesta entre Oralité et Écriture* (ed. by J.E.M. Houben et J. Rotaru): 147-183. Bucarest: Bibliothèque de Bucarest.

Houben, Jan E.M. 2014. "A Tradição Sânscrita entre Memética Védica e Cultura Literária." (in Portuguese) *Linguagem & Ensino*, v.17, n.2, p. 441-469, maio/ago.

Houben, Jan E.M. 2016. "From Fuzzy-Edged 'Family-Veda' to the Canonical Śākhas of the Catur-Veda: Structures and Tangible Traces." In: *Vedic Śākhās: Past, Present, Future. Proceedings of the Fifth International Vedic Workshop, Bucharest 2011*, sous la dir. de J.E.M. Houben, J. Rotaru and M. Witzel, p. 159-192. Cambridge, Mass. : Harvard University.

Houben, J.E.M., and Saraju Rath. 2012. "Manuscript Culture and its impact in 'India': Contours and Parameters." In: *Aspects of Manuscript Culture in South India*, ed. by S. Rath: 1-53. Leiden: E.J. Brill.

Householder, Fred W. Jr. Review of Z.S. Harris 1951. *International Journal of American Linguistics* 18 (1952). 260–268.

Ingalls, Daniel H.H. 1989. "Ānandavardhana's *Devīśataka*." *Journal of the American Oriental Society* 109.4: 565-575.

Jamison, Stephanie, and Joel Brereton. 2014. *The Rigveda: the earliest religious poetry of India.* Oxford: Oxford University Press.

Joshi, S.D. 1989. "Patañjali's views on *Apaśabdas*." In: *Dialectes dans les littératures indo-aryennes* (ed. C. Caillat): 267-274. Paris: Institut de Civilisation Indienne.

Joshi, S.D., and J.A.F. Roodbergen. 1986. *Patañjali's Vyākaraṇa-Mahābhāṣya. Paspaśāhnika. Edited with Translation and Explanatory Notes*. Pune: University of Poona.

Kaldewaij, Jelle. 1986. "Structuralisme en transformationeel generatieve grammatica: continuiteiten en discontinuiteiten in de twintigste eeuwse taalkunde – Structuralism and Transformational Generative Grammar: Continuities and discontinuities in twentieth century linguistics (in Dutch; summary in English)." Dissertation Rijksuniversiteit Utrecht.

Kalyanaraman, S. 2016. *Harappa Script & Language: Data mining of corpora, tantra yukti and knowledge discovery of a civilization*. Herndon, VA: Sarasvati Research Center.

Kane, Pandurang Vaman. 1971. *History of Sanskrit Poetics.* 4th rev. edition. Delhi: Motilal Banarsidass.

Kaye, G.R. 1927. *The Bakhshālī Manuscript: A Study in Mediaeval Mathematics*. Calcutta: Government of India, Central Publication Branch.

Keith, Arthur Berriedale. 1908. *Śāṅkhāyana-Āraṇyaka, with an appendix on the Mahāvrata.* London: The Royal Asiatic Society.

Keith, Arthur Berriedale. 1909. *Aitareya Āraṇyaka: Introduction, Translation, Notes, Indexes and an appendix containing the portion hitherto unpublished of the Śāṅkhāyana Āraṇyaka.* Oxford: Clarendon Press.

Keith, Arthur Berriedale. 1914. *The Veda of the Black Yajus School entitled Taittirīya Sanhitā,* Pts. I-II. Cambridge: Harvard University Press.

Keith, Arthur Berriedale. 1924. *Sanskrit Drama in its Origin, Development, Theory and Practice.* London: Oxford University Press.

Kent, Roland G. 1950. *Old Persian. Grammar, Texts, Lexicon.* New Haven: American Oriental Society.

Kiparsky, Paul. 1979. *Pāṇini as a Variationist.* Cambridge, Mass.: MIT Press.
Koerner, E.F.K. 1997. "Notes on the history of the concept of language as a system 'où tout se tient'." *Linguistica atlantica* 18/19 (1996/1997): 1-20.
Kloss, Heinz. 1967. "'Abstand languages' and 'Ausbau languages'." *Anthropological Linguistics* 9, no. 7: 29-41.
Kuiper, F.B.J. 1968. "The genesis of a linguistic area." *Indo-Iranian Journal* 10: 81-102.
Labov, William. 1994. *Principles of Linguistic Change. Volume I. Internal Factors*. Oxford: Blackwell.
Lévi, Sylvain. 1915. "Sur la récitation primitive des textes bouddhiques." *Journal Asiatique*, 11e série, tome 5: 401-447.
Lewis, M. Paul. 2009. *Ethnologue: Languages of the World*. Sixteenth edition. Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/.
Lienhard, Siegfried. 1984. *A History of Classical Poetry: Sanskrit – Pali – Prakrit.* (A History of Indian Literature, 3, 1.) Wiesbaden: Otto Harrassowitz.
Lin, Li-Kouang. 1949. *Introduction au Compendium de la Loi (Dharma-Samuccaya): L'Aide-mémoire de la Vraie Loi* (*Saddharma-smṛty-upasthāna-sūtra*). Paris: Adrien Maisonneuve.
Mahadevan, Iravatham. 2002. "Aryan or Dravidian or Neither? A Study of Recent Attempts to Decipher the Indus Script." *Electronic Journal of Vedic Studies*, 8-1: 1-19.
Mayrhofer, Manfred. 1951. *Handbuch des Pāli, mit Texten und Glossar: eine Einleitung in das sprachwissenschaftliche Studium des mittelindischen.* Heidelberg: Carl Winter.
Meillet, Antoine. 1908. *Les Dialectes Indo-européens*. Paris: Librairie Ancienne Honoré Champion.
Meillet, Antoine. 1921. *Linguistique historique et linguistique générale*. Paris: Librairie Ancienne Honoré Champion.
Meillet, Antoine. 1936. *Linguistique historique et linguistique générale,* Tome II. Paris: Librairie C. Klincksieck.
Morvan, Michel. 1996. *Les Origines Linguistiques du Basque*. Bordeau: Presses Universitaires.
Muljačič, Žarko. 1986. "L'enseignement de Heinz Kloss (modifications, implications, perspectives)." *Langages, No. 83, Glotto-politique* (Septembre 86): 53-63.
Muljačič, Žarko. 1989. "The emergence of the Florentine > Italian language." *Synchronic and diachronic approaches to linguistic variation and change; Georgetown University round table on languages and linguistics, 1988* (ed. J. Walsh): 221-226. Washington DC: Georgetown University Press.
Muljačič, Žarko. 1993. "Standardization in Romance." *Bilingualism and Linguistic Conflict in Romance (Trends in Romance Linguistics and Philology*, vol. 5) (ed. R. Posner & J. N. Green): 77-116.
Ollet, Andrew. 2017. *Prakrit, Sanskrit and the Language Order of Premodern India.* Oakland: University of California Press.
Parpola, Asko. 1994. *Deciphering the Indus Script.* Cambridge: Cambridge University Press.
Parpola, Asko. 2015. *The Roots of Hinduism: The Early Aryans and the Indus Civilization*. Oxford: Oxford University Press.
Parpola, Asko and Simo. 1975. "On the relationship of the Sumerian Toponym Meluḫḫa and Sanskrit Mleccha." *Studia Orientalia,* 46: 205–238.
Peeters, Bert. 2001. "Does Cognitive Linguistics Live up to its Name?" *Language and Ideology: vol. 1: Theoretical cognitive approaches* (eds. R. Dirven, B. Hawkins, E. Sandikcioglu): 83-106. Amsterdam: John Benjamins.
Pinchard, Alexis. 2009. *Les langues de sagesse dans la Grèce et l'Inde anciennes*. (École Pratique des Hautes Études, Sciences historiques et philologiques - III. Hautes Etudes du monde gréco-romain, 43.) Genève-Paris: Droz.
Pinault, Georges-Jean. 2002. "Tocharian and Indo-Iranian: Relations between two linguistic areas." In: *Indo-Iranian Languages and Peoples* (ed. by N. Sims-Williams): 243-284. Oxford: Oxford University Press.
Plofker, Kim, Agathe Keller, Takao Hayashi, Clemency Montelle, and Dominik Wujastyk. 2017. "The Bakhshālī Manuscript: A Response to the Bodleian Library's Radiocarbon Dating." *History of Science in South Asia*, 5.1: 134-150.
Pollock, Sheldon. 2006. *The Language of the Gods in the World of Men: Sanskrit, Culture and Power in Premodern India.* Berkeley: Univ. of California Press.
Possehl, Gregory. 1996. *Indus Age: The Writing System.* Philadelphia: University of Pennsylvania Press.
Renou, Louis. 1956. *Histoire de la Langue Sanskrite.* Lyon: Editions IAC.
Renou, Louis. 1957a. "Introduction générale" (trad. du "Einleitung" de J. Wackernagel, mise à jour, annotée). J. Wackernagel, *Altindische Grammatik*. Band I. Lautlehre. New Edition. Göttingen: Vandenhoeck und Ruprecht.
Renou, Louis. 1957b. *Hymnes spéculatifs du Veda*. Paris: Gallimard.
Robert, Jean-Noël. 2012. *Japanese Hieroglossia.* (Trad.: Liz Libbrecht) Paris: Collège de France. (http://books.openedition.org/cdf/3101, acc. 30-11-2016)
Salomon, Richard. 1989. "Linguistic variability in post-Vedic Sanskrit." 275-. In: *Dialectes dans les littératures indo-aryennes* (ed. C. Caillat): 275-294. Paris: Institut de Civilisation Indienne.
Salomon, Richard. 1998. *Indian Epigraphy: A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and the other Indo-Aryan Languages.* New York: Oxford University Press.
Salomon, Richard. 2001. " 'Gāndhārī Hybrid Sanskrit': New Sources for the Study of the Sanskritization of Buddhist Literature." *Indo-Iranian Journal* 44 (3): 241–252.
Saussure, Ferdinand de. 1916. *Cours de Linguistique Generale*. Geneva.
Scharfe, Hartmut. 1977. *Grammatical Literature.* (A History of Indian Literature, 5, 2.) Wiesbaden: Otto Harrassowitz.
Scharfe, Hartmut. 2002. *Education in Ancient India.* Leiden: E.J. Brill.
Schmitt, Rüdiger. 1967. *Dichtung und Dichtersprache in indogermanischer Zeit.* Wiesbaden: Harrassowitz.

Schmitt, Rüdiger. 2009. *Die altpersischen Inschriften der Achaimeniden: editio minor mit deutscher Übersetzung.* Wiesbaden: Reichert.

Śivadatta and Vāsudeva Lakṣmaṇ Śāstrī Paṇaśīkar. 1916. *Kāvyamālā* (*Collection of old and rare Kāvyas ...etc.*), Pt. 9. Bombay: Nirnaya Sagar.

Smith, Helmer. 1954. *Analecta Rhythmica*. *Studia Orientalia* (edidit Societas Orientalis Fennica) XIX: 7, 3-17.

Spence, N.C.W. 1957. "A hardy perennial: the problem of *la langue* and *la parole*." *Archivum Linguisticum* 9: 1-27.

Swennen, Philippe. 2015. "Hommes et dieux dans le formulaire métrique indo-iranien ancien." In: *Démons iraniens – Actes du colloque international organisé à l'Université de Liège les 5 et 6 février 2009 à l'occasion des 65 ans de Jean Kellens* (ed. Ph. Swennen): 61-72. Liège: Presses Universitaires.

Thieme, Paul. 1935. *Pāṇini and the Veda : Studies in the early history of linguistic science in India.* Allahabad : Globe Press.

Thieme, Paul. 1985. "Nennformen aus Anrede und Anruf im Sanskrit." *Münchener Studien zur Sprachwissenschaft* 44: 239-258. (1995. *Kleine Schriften* II: 1054-1073.)

Vennemann Gen. Nierfeld, Theo. 2003. *Europa Vasconica – Europa Semitica.* Berlin: Mouton De Gruyter.

Wackernagel, Jakob. 1896. *Altindische Grammatik. Band I. Lautlehre.* Göttingen: Vandenhoeck und Ruprecht. (Repr. with "Introduction générale", the translated and profusely annotated "Einleitung" of Wackernagel, by Louis Renou, 1957.)

Watkins, Calvert. 1994. "Language of Gods and Language of Men : Remarks on Some Indo-European Metalinguistic Traditions." in *Selected Writings* II: 456-472. Innsbruk: Institut für Sprachwissenschaft der Universität Innsbruck.

Whitney, William Dwight. 1888. *Sanskrit Grammar. Including both the classical language and the older Dialects of Veda and Brāhmaṇa.* 2nd ed. (First edition: 1879.) Leipzig. [Repr.: Delhi: Motilal Banarsidass, 1977.]

Witzel, Michael. 1987. "The Localisation of Vedic Texts and Schools." (Materials on Vedic śākhās, 7)." In: *India and the Ancient world. History, Trade and Culture before A.D. 650. P.H.L. Eggermont Jubilee Volume.* (ed. by G. Pollet): 173-213. Leuven 1987.

Witzel, Michael. 1989. "Tracing the Vedic Dialects." In: *Dialectes dans les littératures indo-aryennes* (ed. C. Caillat): 97-265. Paris: Institut de Civilisation Indienne.

Witzel, Michael. 1997. "The Development of the Vedic Canon and its Schools: The Social and Political Milieu." *Inside the Texts – Beyond the Texts: new approaches to the study of the Vedas* (ed. M. Witzel): 257-345.