e-ISSN 2785-5953

Bhasha
Vol. 3 – Num. 1 – April 2024

Final Syllable Reduction in Middle Indic and Iranic

Andrew Ollett
University of Chicago, USA

Abstract The mutual influence of Indic and Iranic languages in the 'contact zone' of the upper Indus valley and the adjoining mountain ranges has long been known to scholarship. Previous scholarship has observed lexical borrowings in the ancient period, and phonological and morphological features in the modern period. Outside of this area, the Indic and Iranic languages developed in different ways, largely in accordance with the prominence of the syllable and the word, respectively, in their phonologies. The important Middle Indic languages of Gandhari and Apabhramsha, however, appear to have participated in certain structural changes that affected all of the contemporary Middle Iranic languages, above all the reduction of final syllables. The fact that the earliest attestation of these changes in Middle Indic occurs in languages strongly associated with the Indo-Iranian contact zone suggests that contact may have played a critical role. The chronology of these changes in Middle Indic also suggests that they occurred in the 'Saka-Kuṣāṇa Age' in the early centuries of the common era.

Keywords Middle Indic. Gandhari. Apabhramsha. Middle Iranic. Historical phonology. Language contact.

Edizioni Ca'Foscari

Peer review

Submitted 2023-12-14
Accepted 2024-04-22
Published 2024-05-24

Open access

Citation Ollett, A. (2024). "Final Syllable Reduction in Middle Indic and Iranic". *Bhasha*, 3(1), 53-106.

DOI 10.30687/bhasha/2785-5953/2024/01/004

1 Introduction

This paper makes three connected arguments. First, it identifies a sound change that affected all of the Middle Iranic languages and some of the Middle Indic languages, which I call Final Syllable Reduction (FSR).[1] FSR itself is a well-known feature of many Middle Iranic languages and of the Middle Indic languages Gandhari and Apabhramsha, but with one exception, scholars have not attempted to bring these changes under a single unified analysis. Second, it argues that the application of FSR to the Middle Indic languages in question, Gandhari and Apabhramsha, is due to contact with speakers of Iranic languages. Scholars have previously taken these changes to be internal developments within the Indic language family. Given the long history of contact between Indic and Iranic languages in the 'contact zone' where Gandhari and Apabhramsha appear to have originated, a finer-grained account of the sound changes in question is desirable for evaluating the plausibility of contact as an explanation. In that connection, the third argument of this paper is that FSR in Gandhari is due to its use as a *lingua franca*, starting in the first century BCE, by speakers of Iranic languages, and that similar developments explain the characteristic features of Apabhramsha. The long-term impact of these changes was profound, since these features of Apabhramsha subsequently spread to most of the North Indian languages, and resulted in the North Indian languages having phonological systems that differed greatly from those of their ancestors (Common Middle Indic) and their neighbours (the Dravidian languages).

2 The Long History of Indo-Iranian Contact

The contact between representatives of the Indic and the Iranic language families is full of both historical and theoretical significance. To start with its theoretical significance, it has long been recognised that the contact zone between these families forms part of larger 'linguistic areas'. A linguistic area is a zone in which certain linguistic features are found across the boundaries of language 'families' as constituted by historical linguistics. Retroflex consonants, for example, are found throughout the South Asian linguistic area, in the Indic, Iranic,

1 I follow Kümmel (2018) in using the term 'Iranic' to refer to the language family in distinction to 'Iranian' (used in a wide variety of other contexts). Similarly, I use 'Indic' to refer to the language family that other scholars call 'Indo-Aryan', especially since it is extremely difficulty to pry the word 'Aryan' away from the racial (and racist) ideas with which it has long been associated.

e-ISSN 2785-5953

Dravidian, and Munda language families.[2] The boundary between the Indic and Iranic language families is ostensibly the Indus river, with Iranic languages (Balochi and Pashto) on the western side and Indic languages (Sindhi, Seraiki, and Punjabi) on the eastern side, terminating in a dense tangle of languages where the Indus descends from the Hindu Kush. Yet many phonological, morphological, and syntactic features integrate these languages into a larger South Asian linguistic area.

The Indic and Iranic languages are, of course, related by common descent from Proto-Indo-Iranian. In contrast to two clearly unrelated languages, in the case of neighbouring Indic and Iranic languages, it can be difficult to determine whether a common linguistic feature has been maintained by both from the protolanguage or transmitted from one daughter language to another (cf. Morgenstierne 1975). A term that may prove helpful for language in the contact zone is 'linkage', which François (2015, 170-1) defines as "a network of dialects which remained in contact with each other for an extended period of time".

Linkages are actually predicted by the traditional 'family tree' model of language relationships, given two premises. The first is that the speakers of the languages in question stay more or less in one place – in other words, that the increasing distance between languages is due to the gradual but differential accumulation of sound changes by different social groups, rather than a sudden dislocation of one group of speakers (due, for example, to migration). A second and related premise is that the languages, and social groups, that have separated in this way can and often do remain in contact. It is a well-known weakness of the 'family tree' model that the situation of contact that it predicts as the general rule cannot be represented in the model itself. This is one of the reasons why the family tree model has fallen out of favour for accounting for and representing the internal differentiation of language families (cf. François 2015; Korn 2019).

Whatever model of language change we ultimately adopt, the Indo-Iranian contact zone is characterised by 'differentiation plus contact', i.e. a situation of contact between genetically related languages. At the time of our oldest historical documentation of the Iranic and Indic languages, they were already very similar to each other: "There must, in the sixth and fifth centuries B.C., have been hundreds of the most commonly used words which were practically identical on both sides of the linguistic border" (Morgenstierne 1974, 271; cf. also Chatterji 1960, 126). As these languages gradually diverged from each other due to sound changes, contact phenomena may have brought them back together.

2 Cf. Emeneau 1956; 1980 (focused on phonological features) and Masica 1976 (focused on syntax) for the South Asian linguistic area; Tikkanen 2008 focuses specifically on the 'transit zone' between South Asia and Central Asia.

One major obstacle to the study of language contact over the long term is the lack of documentation. In the case of Indo-Iranian, we are exceptionally lucky to have representatives of the language families in question attested for thousands of years. But even so there are long breaks in the historical record. Even if Yaghnobi might continue Sogdian in some sense, or Khowar might continue Gandhari, there is nevertheless a gap of many centuries. Sanskrit has played an important role in the reconstruction of Indo-Iranian contact in part because it provides a much greater quantity of evidence, thanks to its continuous use over a wide geographic area for thousands of years; but for precisely the same reason, this evidence needs to be used with care, and the conservatism of Sanskrit has meant that the only types of contact phenomena visible from this perspective are loanwords (cf. Chatterji 1960; Morgenstierne 1974). As a somewhat crude generalisation, we may say that scholars have approached Indo-Iranian contact using two very different sets of evidence: the modern languages (cf. Bashir 2016), each of which presents, synchronically, the outcome of various contact phenomena; and the ancient languages, which could in principle attest to the before-and-after stages of contact phenomena, but which have nevertheless been treated in much the same way as the modern languages, viz. as a 'snapshot' of contact phenomena that happens to have been taken earlier in time.

Ironically, although the modern evidence provides strong evidence for contact-induced sound changes, in the form of phonological isoglosses between neighbouring languages, I am not aware of any scholarship that attributes sound changes in the ancient languages to contact. That is due partly to the 'lexical bias' of the Sanskrit evidence, noted above, but also to the fundamental impossibility of ascertaining the causes of sound change outside of relatively broad parameters; the areal isoglosses that make contact a possible or likely explanation are often precisely what is lacking in the case of ancient languages.

Let us now turn to the historical significance of Indo-Iranian contact. Anyone who wants to travel between South Asia and the west by land must go through the Indo-Iranian contact zone. This region, and especially the mountain passes in the Karakorum range, have for this reason been considered one of the 'crossroads of Asia'. It has been the primary corridor within which new peoples, languages, cultures, religions, and technologies have entered South Asia, and through which South Asia, in turn, has been integrated into other Asian political and cultural formations.

The mountain ranges of Eastern Afghanistan, Northern Pakistan, and Northwest India are known for their linguistic diversity as well as the density and complexity of language contact. Bashir (2016) calls the region PHKKK (Pamir-Hindukush-Karakorum-Kohistan-Kashmir). Two anthropologists, Augusto and Alberto Cacopardo, have

argued that the same region, which they call Peristan, is culturally distinct from its lowland neighbours (Cacopardo 2016). In Zoller's interpretation, "these differences must be explained for the greater part in terms of earlier and later Indo-Aryan immigrations" (2018, 178; cf. also 248). Zoller appears to favour the hypothesis that the Indic languages in the PHKKK range represent an "earlier" wave of migration of Indic speakers that maintained its cultural identity in the face of subsequent waves of migration.

We turn now from prehistory to history. Gandhāra – the region corresponding to modern Peshawar and the adjoining valleys – was added to the Achaemenid Empire during the reign of Darius I (r. 522-486 BCE). The Achaemenid kings ruled the region through satraps until Alexander's victory over Darius III in 331 BCE. It was probably during the latter part of this period that the Sanskrit grammarian Pāṇini produced his *Aṣṭādhyāyī* (Eight Chapters), in Śālātura, now Chota Lahor in northern Pakistan. Sanskritists can easily forget that Pāṇini was a subject of a vast empire that stretched to the Mediterranean Sea. The representatives of the Achaemenid state in Gandhāra – satraps, administrators, bureaucrats, engineers, and soldiers – probably reflected the multiethnic and multilingual composition of the empire, and likely included speakers of Iranic languages (Persian, Median, etc.). It is to this "first period" of Indo-Iranian contact (according to Chatterji 1960) that we can trace the earliest secure examples of borrowings from one language into another, for example *dipi-*/*lipi-* for 'writing' (see below). In fact the technology of writing itself was one of the borrowings: the Kharōṣṭhī script, used in the former Achaemenid provinces of the Northwest, was based on the Achaemenid Aramaic script.

After Alexander, the Northwest came under the control of the Mauryas, and after the Mauryas, the Greeks. It is notable that Iranians continued to play a role in the administration of the Maurya and Indo-Greek states during this period: one *yavanarāja* (Western King) Tuṣāspha, evidently with the Iranic onomastic element *-aspa*, is credited with restoring a tank during the reign of Aśōka.[3]

Iranic-speaking peoples returned to the centre stage of political history in the Northwest starting in the first century BCE, which starts the "second period" of borrowing, according to Chatterji (1960). Groups began to move into South Asia from across the mountains, driven in part by large-scale population movements in Central Asia and the Iranian plateau. The most important group of newcomers were the Sakas, who spoke Eastern Iranian languages and who had lived in the plains of Central Asia. The Sakas under Maues (Moga) took power in the Punjab in around 80 BCE. Maues' successors,

3 Junāgaṛh inscription of Rudradāman, l. 8 (Kielhorn 1905-06).

including Azes (Aya), held the area until the Parthian ruler Gondophares took it over around 20 CE. By this time, however, Saka rulers had established themselves further east, around Mathura, and another group of Sakas (known by the family name Kṣaharāta) had established themselves in coastal Gujarat. The Mathura-based Sakas were absorbed into the Kuṣāṇa empire in the later first century CE, and around the same time, the Kṣaharātas of Gujarat were succeeded by another Saka family, the Kārdamakas. The Kārdamakas would continue to rule in Gujarat and Western Madhya Pradesh until their defeat by the Guptas in the later fourth century.

The reign of the Sakas and especially the Kuṣāṇas represents a period in which Northern India was not just ruled by Iranic-speaking peoples but closely connected to the predominantly Iranic-speaking regions to the east. The Sakas must have spoken an Eastern Iranic language, closely related to Khotanese and Tumshuqese, which are attested later; all of their coins and inscriptions, however, made use of other languages (Greek, Gandhari, and Sanskrit, in that order). The ethnicity and language of the early Kuṣāṇas is still under discussion, but it is clear that Kaniṣka (r. 126-150 CE) adopted Bactrian, an Eastern Iranic language, for public inscriptions and coinage. During the 'Saka-Kuṣāṇa Age', we have dozens of inscriptional attestations of Iranic names in Gandhari and Sanskrit, especially in the Indo-Iranian contact zone. The borrowing of political and military terminology from Iranic into Indic languages continued during this period as well (see below).

Even after the disappearance of Iranic-speaking dynasties from the plains of North India in the third century, considerable Iranic influence can be seen in the colophons of the Sanskrit manuscripts produced during the reign of the Palola Ṣāhis in Gilgit (von Hinüber 1980).

The Saka-Kuṣāṇa Age has a historical significance beyond its role, in this article, as the backdrop for contact between Indic and Iranic languages. Max Müller is only the most prominent of a number of scholars who posited a radical break in Indian literary and cultural history around the turn of the common era. Müller's chronology was quite different from the one that scholars generally accept today, but if we update his argument, he would state that there was a properly 'ancient' period that extended until about the fourth or third century BCE, which includes the entirety of the Vedas as well as some version of the canonical texts of the Buddhist and the Jains, and a 'revivalist' period that starts from the second century CE or so, which includes classical Sanskrit literature (Aśvaghōṣa, Kālidāsa, Bhāravi, etc.). Müller himself attributed the break to a "Turanian invasion" (1883, 85).

Müller's theory was embedded in an orientalist narrative, with its fetishisation of the very distant past (a time of advanced humanity which has been on the decline ever since) and its racialisation of

culture. These tropes must be resisted and refuted with vigilance. But the idea of an important articulation in the linguistic and cultural history of South Asia, falling within a century or two of the year zero, continues to find acceptance, explicitly or implicitly, even if Müller's name is hardly ever associated with it. The distinction between 'Vedic' and 'Classical' Sanskrit is one instance; we will see below that it articulated Mayrhofer's etymological dictionary of Sanskrit into two parts. And the evidence for migration into the subcontinent from Central Asia on the part of groups who identified as Sakas (Müller's "Turanians") between the first century BCE and the second century CE has only accumulated since then.

Some version of Müller's hypothesis may help to account for two important phenomena of which Müller himself had almost no knowledge whatsoever. One of these is the literature in the language that Bailey called "Gandhari", now known to be represented by hundreds of inscriptions, wooden and leather documents, and birch-bark manuscripts dating from the first century BCE to about the third century CE (Bailey 1946; Salomon 1999b). This literature was only discovered in the early twentieth century, with a large cache of additional materials coming to light in the final decade of that century. Another is the literary language called Apabhramsha (literally 'degradation') in premodern India, which was used mostly from the eight to the twelfth century CE. This language was known from references in Sanskrit texts already in Müller's time, but actual literature in the language was not published until the turn of the twentieth century. What Müller called the "Turanian invasion", which we might call the 'Saka-Kuṣāṇa Age', is of decisive importance to the development of both languages. Its importance to the development of Classical Sanskrit literature was already outlined, albeit controversially, by Lévi (1903), who similarly had no knowledge of Gandhari, and little knowledge of Apabhramsha, at the time.

3 Key Isoglosses between Middle Indic and Middle Iranic

A sound change that I will call Final Syllable Reduction (FSR) affected all of the Middle Iranic languages and two Middle Indic languages originating in the Indo-Iranian contact zone. FSR is a two-stage process. First, the quantity of the final syllable is reduced, such that all heavy final syllables become light. Secondly, the quality of the final syllable is lost. This might mean that they were pronounced similarly or identically, or it might mean that they were not pronounced at all. This second stage might itself have unfolded in two separate moments, i.e. the loss of final vowel quality followed by the loss of final vowels, or the loss of final vowels might have followed directly from the loss of final vowel quantity. The evidence in general does

not allow us to decide between these two steps or scenarios; what we observe, instead, in languages wherein final vowels are written, is the systematic confusion of forms of a word that historically had different final vowels.

The oldest-attested languages of the Indo-Iranian language family, such as Vedic Sanskrit, Avestan, and Old Persian, were all highly inflected. Nouns, adjectives, pronouns, and verbs were inflected according to their grammatical category, and most of these inflections were found at the end of the word. The final syllable of a word therefore contained indispensable information about its part of speech and grammatical category. Hence, although FSR is a 'mere' sound change, by targeting the final syllable, it resulted in the collapse and partial reconstruction of the inherited system of inflection. FSR is therefore implicated in a wide range of structural changes in the Middle Indic and Middle Iranic languages, including the loss of distinctions in gender, number, and case.

One important corollary to FSR is the development of 'extended' endings by attaching a pleonastic suffix (deriving from Proto-Indo-Iranian *-*ka*-) to nominal forms, i.e. nouns, adjectives, and pronouns. Although the conditions for the addition of this suffix differ from language to language, all the languages discussed in this paper ended up with secondary inflections.

A secondary feature of Middle Indo-Iranian that I identify here, although much more tentatively than FSR, is the fricativisation of consonants. Already in Proto-Iranian, a number of inherited stops had taken on a fricative pronunciation, specifically: (1) voiceless aspirated stops in all positions (/k^{h}/, /t^{h}/, /p^{h}/ → /x/, /θ/, /f/), and (2) voiceless unaspirated stops before other consonants (/k/, /t/, /p/ → /x/, /θ/, /f/ / _/C/). In the analysis below I connect both FSR and fricativisation to the shift from 'syllable' phonology to 'word' phonology in the relevant languages, which has been perceptively noted already by Kümmel (2014).

4 Overview of the Data

The languages that I argue were affected by the changes outlined above include:[4]

- Middle Iranic:
  - 'Eastern':
    - Sogdian: attested from the fourth to about the tenth century, in Sogdia (the area between the Zeravshan and Oxus rivers) and generally along the Silk Road. Written in a consonantal script derived from Aramaic, with a few documents also in the Brāhmī script.
    - Bactrian: attested in coins and inscriptions of the Kushana period (late first to early third century), and in leather documents up until the ninth century, in Bactria (northern Afghanistan) and neighbouring areas. Written in the Greek script.
    - Khotanese: attested in Khotan, in modern Xinjiang UAR, from the fourth through the eight century. Written in the Brāhmī script.
  - 'Western':
    - Middle Persian: the language of southwestern Iran, used as a literary language by Zoroastrians, Christians, and Manichaeans, attested in documents from Iran and Turfan (in Xinjiang UAR), from roughly the third century; I do not separately discuss Parthian, a 'Western' Iranic language that is in many ways similar to Middle Persian. Both are written in consonantal scripts derived from Aramaic.
- Middle Indic:
  - Gandhari: attested in the 'greater Gandhara' region (around modern Peshawar) from the first century BCE to the late third century CE, with some earlier inscriptions written in the same script (Kharōṣṭhī), as well as a collection of third century documents from Niya, in the Tarim basin.
  - Apabhramsha: a literary language first associated with certain communities of northwestern India (Gujjars and Ahirs), attested from the eight century onwards.

This list includes all of the attested Middle Iranic languages, but only two of the Middle Indic languages. In other words, the features that I take to be diagnostic of contact between the Middle Indic and Middle Iranic languages are general features of Middle Iranic, but their presence in Middle Indic is exceptional and requires explanation. None

4 Cf. Kümmel 2018; Korn 2019 for the conventional but controversial distinction between 'Western' and 'Eastern' Iranic languages.

of the other attested Middle Indic languages, including Prakrit, Ardhamagadhi, Pali, or Epigraphic Middle Indic exhibit either FSR or fricativisation. (Of course, final syllables in all these languages were reduced in the case that they originally consisted of superheavy syllables, but all superheavy syllables, not just those at the end of a word, were eliminated in Middle Indic; indeed this is one of the defining characteristics of Middle Indic). The loss of single intervocalic consonants in these languages, however, is sometimes said to have passed through a fricative phase; this will be discussed below in reference to Gandhari.

In the conclusion I suggest that the changes associated with FSR are also represented in modern languages. I am not a specialist of the modern Indic or Iranic languages, so I cannot speak with any authority about the range of languages affected by these changes. I will note, however, that most of the Indic languages of North India, including Hindi, Gujarati, Marathi, and Bengali, exhibit the loss of final vowels, and show clear evidence for FSR in their early records. Hence it appears that the changes that were generalised within the Iranic family of languages already by the beginning of the common era came to be generalised within the Indic family, too, albeit centuries later.

4.1 Sogdian and the Rhythmic Law

FSR is corroborated by the following features of Sogdian:

- Inherited final vowels are shortened. This is a postulate, because the quantity of vowels is not marked in Sogdian writing. I note, however, that all instances of final vowels can be interpreted as short, and this accords with the following generalisation about inherited final consonants.
- Inherited final consonants are lost.
  - The sequence */am#/ becomes /u#/ (Gershevitch 1954, §§ 349, 1171, 1194).
  - The sequence */aːm#/ becomes /u#/ (Gershevitch 1954, §§ 1173, 1192) or /a#/ (Gershevitch 1954, § 1175).
  - The sequence */iːm#/ becomes /i#/ (Gershevitch 1954, §§ 350. 1174, 1197).
  - The sequence */ah#/ becomes /i#/ (Gershevitch 1954, §§ 402, 1168, 1191).
  - The sequence */aːt#/ becomes /a#/ (Gershevitch 1954, § 1179).
  - The sequence */aːh#/ becomes /a#/ (Gershevitch 1954, § 1185).

An important exception to the above rules is monosyllables like /xoː/ ‘he’ and /xaː/ ‘she’.

The foregoing sound changes operate throughout Sogdian. Their results serve as input to another set of sound changes which,

however, are limited to certain types of words. I refer to the well-known 'Rhythmic Law' of Sogdian. Paul Tedesco was the first to notice that the inflectional endings of a word differ according to some prosodic feature of their base. Some bases are 'light', and some are 'heavy'. The endings of 'heavy' bases are reduced or elided versions of the endings of 'light' bases. See table 1 for an example.

Table 1 Examples of the Sogdian Rhythmic Law

	Light stem	**Heavy stem**
written	⟨βγy⟩	⟨β'γ⟩
pronounced	/βaɣi/	/βaːɣ/
category	nom.sg.	nom.sg.
meaning	"god"	"garden"

The nominative masculine singular ending /-i/ has been deleted after the heavy base /βaːɣ-∅/ but not after the light base /βaɣ-i/. This difference has been accounted for, by Gershevitch and Sims-Williams, by formulating a rule that deleted final unstressed vowels. This rule should target the final syllable of */βaːɣ-i/ but not of */βaɣ-i/. Hence the final vowel of the former should be unstressed (*/'βaːɣ-i/), whereas the final vowel of the latter should be stressed (*/βaɣ-'i/).

Sims-Williams explained the assignment of stress in the following terms:[5]

- stress a syllable containing a long vowel;
- otherwise, stress the final syllable.

Hence in Sims-Williams' account what makes a stem heavy is simply the occurrence of a long vowel anywhere within the stem. (According to him, a following consonant in the same syllable – which, in many languages, make the syllable prosodically heavy – does not make a stem heavy for the purposes of the Rhythmic Law, except in those cases where a consonant, such as 'ṁ' or 'r', can be interpreted as a nasalised or rhotacised vowel segment. I am not in a position to disagree with Sims-Williams about the patterns found in the language, but I do find the phonetic characterisation of preconsonantal 'ṁ' and 'r' to be idiosyncratic). Sims-Williams later (1996, 312) noted a tendency to omit "short vowels in initial open syllables" when Sogdian is written in the Brāhmī script (e.g. *mdhu* for [məðu]), which he took to support his idea that final syllables are stressed as a rule.

5 "[A] heavy syllable may therefore be defined very simply as a syllable which contains a long vowel or diphthong" and "it should be possible to define a heavy stem as one which contains at least one heavy syllable" (Sims-Williams 1984, 213).

The stress rule allegedly underlying the Rhythmic Law is unique to Sogdian. In fact, according to Novák (2013, 74), it did not operate in Yaghnobi, Sogdian's closest ancestor, at all. Novák, who accepts the Gershevitch-Sims-Williams account of the Rhythmic Law, argues for four distinct types of stress pattern in Sogdian:

- Stress I: the Proto-Iranian (and Proto-Indo-Iranian) 'free stress' system, where the location of prominence within a word is determined by lexical and morphological factors.
- Stress II: stress a heavy penultimate syllable, or if the penultimate is light, the antepenultimate syllable. Novák (2013) dates this shift prior to the divergence of Sogdian and Yaghnobi, hence it operated at least in the 'Proto-Sogdic' period.
- Stress III: the Gershevitch-Sims-Williams system described above, i.e. stress a syllable containing a long vowel, or if there is none, stress the final syllable. Novák notes that this system operated only in Sogdian, and not in Yaghnobi.
- Stress IV: stress the final syllable.

In the following I will propose a reanalysis of the Rhythmic Law that completely dispenses with Stress III. In this analysis, the Rhythmic Law of Sogdian does not presuppose any change in stress vis-à-vis other Middle Iranic languages. Rather, I consider it to be a deletion rule that operated in Sogdian but presupposes the same stress pattern shared by most of the other Middle Iranic languages, namely, Novák's Stress II. There are four motivations for this analysis.

First, Stress III is posited only in order to explain the Sogdian Rhythmic Law. There is, to my knowledge, no other evidence for it, and in fact the evidence for Sogdian written in Brāhmī (Sims-Williams 1996) supports Stress IV rather than Stress III. As Novák's summary shows, Stress III was not a feature of Proto-Sogdic, either. I also know of no other Middle Iranic languages that exhibited such a stress pattern, although possible parallels to the Rhythmic Law in Middle Persian will be discussed below.

Second, the phonological basis of this stress pattern remains unclear. Phonologists now tend to see stress as a form of prominence assigned cumulatively from lower to higher levels of phonological structure (cf. e.g. Nespor, Vogel 1986). In Hayes' (1995) influential account, stress patterns arise from the parsing of phonological units (syllables or moras) into prosodic feet, which are then grouped and assigned prominence according to language-specific parameters. Hayes has argued that all attested stress patterns can be accounted for with an inventory of only three types of prosodic feet: syllabic trochees, moraic trochees, and iambs. In order to motivate Stress III, we must specify how Sogdian parses its syllables into prosodic feet. Once we do so, however, we no longer need Stress III to account for the Rhythmic Law, as shown below.

Third, the traditional analysis of the Rhythmic Law understands stress assignment to work primarily at the level of the stem, i.e. prior to the addition of inflectional endings; hence there are 'heavy' stems and 'light' stems. But this analysis is complicated by the fact that some inflectional endings can make a stem heavy for the purposes of the Rhythmic Law, as we will see below. The fact that the same stem can be either light or heavy (for the purposes of the Rhythmic Law) depending on the ending strongly suggests that the Rhythmic Law is not sensitive to the vowel quantities within a stem but rather the syllable quantities within a prosodic word as a whole.

Fourth, the Rhythmic Law can be seen as a manifestation of a more general phenomenon, namely FSR, which cuts across the distinction between light and heavy stems in Sogdian, and which is widely attested among the neighbouring languages. Postulating Stress III to account for the deletion of final vowels in heavy stems leaves other aspects of this phenomenon unexplained. In my account, the Sogdian Rhythmic Law (and subsequently Stress IV) operates on the prosodic structures produced by these changes.

I propose that Sogdian parses its syllables into moraic trochees, and the resulting structures account for various facets of Sogdian grammar, including the loss of final vowels in heavy stems. A moraic trochee is a foot consisting of two moras with prominence on the first mora (Hayes 1995, 71):

Figure 1 The moraic trochee (foot and moras)

Hence either two light syllables or a single heavy syllable can be parsed into a moraic trochee:

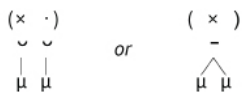

Figure 2 The moraic trochee (foot, syllables, and moras)

Let us assume that Sogdian parses syllables into moraic trochees from right to left. (This analysis may work for left-to-right parsing, but the syllable structure of longer words is not always clear in Sogdian, so I will confine my discussion to the last two to three syllables). Degenerate feet, in this case single light syllables, are not allowed.

Hence we have two parses of the Sogdian words above prior to the operation of the Rhythmic Law:

Figure 3 Parsing of /βaɣi/ and */βa:ɣ/ into moraic trochees

The Rhythmic Law could thus be rephrased simply as: delete a vowel in an unfooted syllable. No shift in word-level accent is necessary. I leave open whether this applies to all unfooted vowels, or just those in final position. In the former case, we may have an explanation of syncope in Sogdian (the extent of which is difficult to determine because Sogdian is written without vowels) in addition to an explanation of the Rhythmic Law. As noted above, 'ambivalent' stems, which count as either light or heavy depending on the inflectional ending, pose a problem for the traditional analysis of the Rhythmic Law. But they are easily accommodated in this account. Let us start with nominal forms. All of the inflectional endings of the singular begin with vowels, and would not make the preceding syllable heavy. But the endings of the plural all begin with /t/, and hence can make a preceding syllable heavy. One example is /rur/ ⟨rwr⟩ 'plant', which was originally light in the singular (e.g. nom.-acc.sg. /rura/ ⟨rwr'⟩) and heavy in the plural (e.g. nom.-acc.pl. /rurt/ ⟨rwrt⟩ , as opposed to the expected light ending */rurta/ * ⟨rwrt'⟩). In this analysis, the heavy form results from the deletion of the final vowel, which is now in an unfooted syllable. (The light forms are sometimes analogically restored).

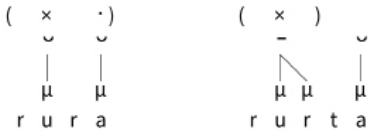

Figure 4 Parsing of /rura/ and */rurta/ into moraic trochees

One might expect all stems ending in a consonant to pattern this way, if a following conjunct consonant really makes the preceding syllable heavy. But that is not the case: light stems ending in 'n' and 'm', like /un/ ⟨wn⟩ 'water' or /rəm/ ⟨rm⟩ 'people', remain light even in the plural (nom.-acc.pl. /unda/ ⟨wnd'⟩ , /rəmta/ ⟨rmt'⟩). I can only aver that

these letters are non-moraic before other consonants, as Sims-Williams did in suggesting they marked the nasalisation of a preceding vowel rather than fully-fledged nasal consonants.

A similar phenomenon occurs in the verbal system. Verbal stems that end in 'r' are light if followed by an inflectional ending beginning with a vowel, and heavy otherwise. The deletion of the final vowel would have taken place in the inherited third person singular **ti* and in the inherited second person plural **ta*, but in both cases (especially in the latter) there has been a tendency to restore the vowel (Gershevitch 1954, §§ 722-45). Here is an example of the stem /βar/ ⟨βr⟩ 'bring'.

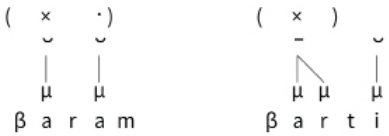

Figure 5 Parsing of /βaram/ and */βarti/ into moraic trochees

This interpretation of the Sogdian Rhythmic Law has parallels in many other languages. In the history of Germanic, high vowels (/i/ and /u/) were lost only in the position targeted by the Sogdian Rhythmic Law, viz. in an unfooted syllable (Dresher, Lahiri 1991; Smith, Ussishkin 2015, 17.3.3.1).

Thus we can see that, in Sogdian, FSR operated on two levels: first, the quantity of all final syllables was reduced across-the-board (except in monosyllables, in observance of word minimality constraints); second, final vowels were targeted for deletion when they were unfooted, that is, at least when they were preceded by a heavy syllable, and perhaps in other cases as well, depending on how exactly syllables were parsed into feet. Sogdian therefore represents an 'intermediate' phase of FSR, when inherited final syllables were already quantitatively reduced, but before they had been eliminated entirely.

Still, even at this intermediate stage, FSR had wrought enormous changes in the inherited system of inflection. Heavy nominal bases ended up with very few inflections: nouns did not distinguish gender, and among cases, only the direct, vocative, and oblique were distinguished. Light bases, by contrast, maintained a two-way gender distinction and a seven-way distinction among cases, if only in the singular.

Besides nominal bases that synchronically ended in a consonant, which admit of a distinction between light and heavy, Sogdian also had nominal bases that synchronically ended in a vowel. These were formed by the addition of the pleonastic suffix **-ka-* or **-kī-* to an

existing base. These 'vocalic stems' maintain the distinction between masculine and feminine only in the direct case of the singular (*-ē* ⟨-y⟩ masc. vs. *-ā* ⟨-'⟩ fem.). In the singular oblique, and in the plural direct and oblique, there is no distinction of gender (Skjærvø 2007, 17).[6]

The use of an extended vocalic declension was grammaticalised in Sogdian. The inherited past passive participle, which was used as such in the earliest documents, is only used as a participle in the later language in the vocalic declension (Gershevitch 1954, § 531: 82). The same applies to the present participle, which is only used as such in the later language in the vocalic declension (132).

4.2 Middle Persian

The final syllables of Old Persian were all lost in Middle Persian. Their loss appears to have been preceded, as one would expect, by a reduction in their quantity. Even later Old Persian inscriptions attest to the loss of distinctions in final syllables. Cantera (2009, 26) notes that, for nouns whose nominative and accusative forms had the same number of syllables, the distinction between these forms was effectively lost by the later fifth century BCE. Korn (2013, 84-5) argues that distinctions of quantity and quality were lost before a final 'm', and *-Vm* had become *-u* by the time Persian words were borrowed into Armenian. Much discussed in connection with the loss of final vowels is the interpretation of the sign ⟨y⟩ at the end of a word in Middle Persian. It had generally been seen as a phonologically insignificant marker of the end of a word. Back then claimed, controversially, that it represented a weak word-final vowel that was lost, in later Old Persian, if the stem was polysyllabic and the preceding syllable contained a long vowel (1978, 41). This is therefore an attempt to apply the Rhythmic Law of Sogdian to Persian.[7] Huyse (2003) revised this account, and concluded instead that the grapheme ⟨y⟩ represented not any final vowel in Old Persian, but specifically the ending */e:h#/ of the genitive singular (from */ahya:/, p. 54), and that this final vowel was lost not in all words at the same time, but at first in (a) polysyllabic words which were (b) neither compounds nor nominal derivatives and where (c) the final vowel */e:h#/ follows a stressed syllable with a long vowel, so long as (d) this long vowel does not represent the end of the word (p. 98). Effectively, as Hintze (2006) notes in her review, this means that monosyllabic stems regularly retain 'y',

6 Sims-Williams (1990, 286) follows Tedesco in deriving the vocalic declension by means of contraction from *-*aki* and then *-*ai*; Gershevitch derived them rather from *-*ak*, as in Middle Persian (with loss of the final vowel).

7 Back 1978, 40-6; cf. MacKenzie 1982; Huyse 2003, 23-5 for a discussion.

while polysyllabic stems retain it only after a long vowel or a short vowel followed by a consonant. There must be a prosodic explanation for this distribution, especially in view of Huyse's observation that in compounds with *bed* 'master', 'y' is written only when the syllable preceding the final member of the compound is prosodically light (p. 80). I will, however, resist trying to state a general prosodic condition for the loss or retention of 'y', given that it was widely extended by analogy to the point where it became, in the late Sassanian period, an orthographic marker of the end of a word.

For the purposes of this paper, it is sufficient to state that the final vowels and vowel + consonant sequences of Old Persian were quantitatively reduced before the Middle Persian stage, resulting in a system wherein every word ended with a short vowel. Whether or not some of those endings were preserved in words with a certain prosodic shape in Middle Persian, the general tendency, evident by the end of the Sassanian period, was the loss of all final vowels.

4.3 Bactrian

The Bactrian data are easy to summarise: "The P[roto-]Ir[anian] [final] vowels are generally lost" (Gholami 2014, 2.7.3). The letter 'o' is typically found at the end of Bactrian words, although it does not have any phonetic significance. Huyse (2003, 61) argues that it originally represented "une voyelle brève réduite" that was later reinterpreted as a word-boundary marker.[8] Gholami (2014, 2.7.3: 61) notes a few Bactrian words that end in vowels other than 'o', without however noting that all of them consist of two light syllables: ιθα 'thus', βαγε 'gods', ταδι 'then', and λιζα 'citadel' (the last from **dizā*). These 'exceptions' clearly indicate that the loss of final vowels was blocked if the resulting form would have been shorter than the minimal prosodic word, i.e. a moraic trochee (two light syllables or one heavy syllable). The vowel is lost even in disyllabic words if the first syllable is heavy, e.g. μαδο /mād/ 'mother' from **mātā* (Gholami 2009, 31). This reminds us of the Sogdian Rhythmic Law: a final syllable is lost if and only if it is preceded by a moraic trochee.

The suffix **-ka* is very common, although as far as I can determine, it occurrence is conditioned lexically rather than by grammatical category (one exception being adjectives from **-ānaka-* > -αγγο, Gholami 2009, 36-7). It is hard for me to say, based on the materials

8 Sims-Williams (*Encyclopedia Iranica* s.v. Bactrian) notes that 'o' also represents an unstressed **a* or **i* in the middle of a word, and that its use at the end of words already ending in a vowel (e.g. ναμωο for Sanskrit *namō*) suggests that it does indeed function as "a word- or morpheme-divider."

I have seen, how this suffix interacted with prosodic principles; in some cases (perhaps most?) a stem-final 'a' was lost by syncope, resulting in forms like ʋαþκο /waːʃk/ 'pupil' < **hāwišta-ka-* and αγγαργο /angaːrg/ 'property' < **ham-kāraka-* (both from Gholami 2009, 36); in other cases (following a conjunct consonant?) it was not, e.g. ανδαγο /andag/ 'borough' < (?) **anta-ka-* (36).

4.4 Khotanese

The following changes are grouped under Kümmel's "Auslautgesetze" (2008, 10). I assume, with Kümmel, that final 'e' and 'o' in Khotanese are short, despite their length not being distinguished in the writing system (cf. also Hitch 2016, 121).

- Inherited final vowels are shortened.
  - */aː#/ becomes /a#/.
- Inherited final consonants are lost.
  - */am#/ becomes /u#/.
  - */ah#/ becomes /ə#/.
  - */ih#/ becomes /ə#/.
  - */aːm#/ becomes /o#/.
  - */aːh#/ becomes /e#/.

These reductions did not apply to monosyllables; cf. the metrically heavy words *o* 'or', *ko* 'if', *co* 'who', etc. (Hitch 2016, 139). Thus the reduction of the quantity of final syllables in Khotanese is exactly parallel to Sogdian, in that it bypassed heavy monosyllables in order to satisfy word minimality constraints.

Long vowels could occur at the end of a word secondarily in Khotanese, i.e. as a result of a contraction of syllables at the end of a word (e.g. *āstai* nom. sg. of the stem *āstaa-* 'bone'). This contraction was usually occasioned by the use of the suffix **-ka-*, which was conditioned, as in the case of Bactrian, lexically rather than by grammatical category.

The reduction of final syllables entailed, as in Sogdian, a loss of distinction between the nominative, accusative, and vocative case forms (Sims-Williams 1990, 284).

Scholars have proposed that Khotanese had a Latin-like stress pattern wherein the penultimate syllable was stressed if it was heavy, and otherwise the antepenultimate syllable was stressed (Kümmel 2008, 4.1.4). This largely (though not entirely) coincides with Maggi's rule that the "stress accent fell on the first heavy syllable from the end of the word but never on the word final syllable" and "[a] word containing only light syllables was stressed on the initial syllable" (Emmerick 2009, 2.2). Hitch noted that "normally the final stem

syllable in O[ld] Kh[otanese] carries primary stress" (2016, 3.3.2.1.5), which was probably meant as a generalisation, not a rule.

Khotanese, somewhat like Sogdian, represents an 'intermediate' phase of the operation of FSR. By the time of the earliest documents, the final syllables had been quantitatively reduced, but never lost. Early Khotanese is therefore the most 'conservative' Middle Iranic language, because final syllables, and hence inflectional endings, are largely retained, albeit reduced in quantity. By the time of the latest documents, however, those syllables were lost entirely (Hitch 2016, 3.2.6; cf. also Emmerick 1987). Bailey (1963, VII-VIII) distinguished between four stages of the language on the basis of this process of loss, illustrated by the inherited genitive plural ending (Proto-Iranian *-ānām*):

- *-ānu* (*/ānam#/ becomes /ānu#/)
- *-āni* (*/ānu#/ becomes /ānə#/)
- *-āṁ* (*/ānə#/ becomes /ān-∅/)
- *-ā* (*/ān#/ becomes /ā-∅/)

Interestingly, he compares the last stage to Apabhramsha (see below), although does not explain it.

4.5 Gandhari

In Gandhari, as in most (if not all) Middle Indic languages, all inherited final consonants other than /m/ were lost, and hence all words ended with a vowel or a vowel followed by /m/ (written as an *anusvāra* in Kharōṣṭhī). The Kharōṣṭhī script in which Gandhari was written does not normally distinguish vowel length. Nevertheless, on the basis of an analysis of the writing of final vowels in post-Aśōkan Gandhari, Fussman (1989, 479) came to the conclusion that at the beginning of the common era Gandhari words ended in one of two vowel sounds:

- ultra-short vowels, resulting from earlier */V/, */V:/, or */Ṽ/ (the last representing a still earlier */Vm/, which, in Fussman's account, yielded a nasalised vowel segment prior to the shortening of final vowels);
- short vowels, resulting from earlier */Ṽ:/ (representing a still earlier */V:m/).

Fussman submitted that qualitative distinctions among the ultra-short vowels were also lost, and that the use of vowel markers such as ⟨-e⟩ , ⟨-o⟩ , ⟨-i⟩ , or ⟨-u⟩ are simply conservative spellings for sounds

that had all become /ə/ in the spoken language.[9] Metrical texts, insofar as they present Gandhari versions of texts that were transmitted in other languages, are very likely to contain such conservative spellings/pronunciations, and hence one important source of information about the length of final vowels is less dispositive than we might like it to be.

A less radical position was taken by Baums (2009, 126-8), who argued that (at least in the British Library fragments 7, 9, 13 and 18) the evidence only supports a merger, in final position, of /e/ and /i/, on the one hand, and of /o/ and /u/, on the other hand: "There is no clear indication that word-final palatal and labial vowels were conflated either with each other or with the neutral vowel" (2009, 128). This would appear to be supported by Brough's observations about the Gandhari *Dharmapāda*: etymological 'ē' and 'ō' are very often written with 'i' and 'u', respectively, but only when they occur in word-final position; in his grammatical sketch he often refers to the "weakness of final syllables" (e.g. 1962, 80, 83).

Loukota offers a different phonological interpretation of Gandhari, according to which there must have been a quantitative contrast between final ⟨-o⟩ (the masculine-neuter direct case ending) and final ⟨-a⟩ (the feminine direct case ending).[10] Evidence for this contrast comes from metrical texts in Gandhari (where a final ⟨-a⟩ often has to be realised as [a:]) as well as the borrowing of Gandhari words into Khotanese and Han Chinese. In Han Chinese, Gandhari words in ⟨-o⟩ are borrowed without the final vowel, whereas Gandhari words in ⟨-a⟩ are borrowed with the final vowel. This contrast could be interpreted in a number of ways, for example as ∅ ~ /a/, /ə/ ~ /a:/, or (following Fussman, albeit with different conditioning factors), ultra-short V ~ short V.

On any of the above interpretations, the status of monosyllables like ⟨ka⟩ is unclear. In other words we cannot know whether the spelling represents /ka:/ or /ka/, or /kə/ or /kă/. In Fussman's and Baums' analyses, as well as one interpretation of Loukota's, the only sound that could occur at the end of a word in Gandhari by the first century or so was a short (or ultra-short) vowel. This probably represents a change from Aśōka's inscriptions at Shabazgarhi and Mansehra. These inscriptions often write *-Vṁ* at the end of a word. On the assumption that 'ṁ' represents a final consonant segment that contributes to syllable weight, rather than a marker of non-moraic

9 "Il est probable qu'en prononciation normale les voyelles ultra-brèves de type a) ne se distinguaient plus les unes des autres et ne jouaient donc plus de rôle phonologique" (Fussmann 1989, 479).

10 I thank Diego Loukota (p.c.) for discussing this with me. This paragraph is my representation of the views he shared with me, which I hope not to have misrepresented. Cf. Loukota 2022 for a detailed discussion of Gandhari loanwords in Khotanese.

nasalisation, and on the assumption that such forms are not borrowings from another dialect, this would mean that Aśōkan Gandhari tolerated heavy syllables at the end of a word, and presumably it would have tolerated long vowels in that position as well. Nevertheless we see already in the Aśōkan inscriptions the use of different endings for the same grammatical category (e.g. *danaṁ*, *dana*, *dane* for the neuter nominative-accusative singular, *jano*, *jana*, *jane* for the masculine nominative singular). Precisely what to make of these alternations is not clear, although it seems significant that they primarily affect those cases that would serve as the subject of a verb (nom. masc. sg. and nom.-acc. neut. sg.; cf. also Caillat 1992, 2.5, 4.4).

The most important grammatical consequence of the reduction of all final syllables was the collapse of the nominative and the accusative into a single form, the direct case. Distinctions between the non-direct cases were retained, because they all added a syllable to the stem. Hence the direct cases were isosyllabic (stem *jana* > direct *jano*), loosely following Cantera (2009) on Middle Persian, whereas the non-direct cases were pleosyllabic (stem *jana* > genitive *janasa*). Gandhari in fact inherited two endings for the locative singular, isosyllabic *-e* (*jane*) and pleosyllabic *-ami*, *-aṁmi*, or *-aspi* (e.g. *janaṁmi*). The inherited ending *-e* was largely used in formulaic phrases (Fussman 1989, 460); a new pleosyllabic ending, *-ae*, came to be used instead.

Since masculine and neuter were only distinguished in the direct cases to begin with, the loss of the distinction between the nominative and accusative would have entailed a loss of the distinction between masculine and neuter stems. There is abundant evidence that historically masculine forms (*-o* and *-e*) were used for erstwhile neuter stems; *-a*, from earlier **-am*, is also often used for the nominative of erstwhile masculine stems, where it might come from either the inherited accusative or the neuter nominative-accusative.

Finally, I will mention that the prevailing interpretation of the Gandhari writing system involves a large number of fricative sounds that are not found in any other Middle Indic language. Single intervocalic consonants were fricativised in Gandhari, which we can tell because the resulting sounds were sometimes written in such a way that distinguished them from their non-fricativised (or voiceless) counterparts:

Table 2 Intervocalic stops in Sanskrit and Gandhari

Sanskrit form	Gandhari form
matam	*mada* [məðə̆]
katham	*kasa* [kəθə̆]
mēdhāvī	*meṣ̲avi* [me:za:vĭ]
māsē	*maṣ̲e* [ma:zĕ]
prabhā	*pravha* [prəʋă] or [prəβă]
svakam	*spaya*, *spaga*, *spaa*, *spaka*, etc. [spəjə̆] (evidently through [spəɣă])

Note that the use of these signs to indicate fricativisation occurs in documents after Aśōka's inscriptions (Konow 1929, lcviii), the significance of which I will return to below (see section 6).

4.6 Apabhramsha

I will leave for section 5 the question of precisely how Apabhramsha fits into this picture in chronological and geographical terms. For now, however, I note that Apabhramsha reduces all final syllables of polysyllabic words (superheavy syllables such as *ām* had already been eliminated in earlier stages of Middle Indic):

- Inherited long vowels are reduced finally. /a:#/, /i:#/, /u:#/, /e:#/ and /o:#/ became /a#/, /i#/, /u#/, /i#/, and /u#/ respectively.
- Inherited vowels followed by a final nasal are reduced. This development is slightly different from the above. Vowels that were followed by a nasal segment are also shortened, with different outcomes (in the case of /a/) depending on whether the vowel is long or short:
  - Old Indic /am#/, /im#/ and /um#/ become /u#/, /i#/ and /u#/ (probably through */Ṽ#/);
  - Old Indic /a:m#/, /i:m#/ and /u:m#/ become /a#/, /i#/ and /u#/ (probably through */Ṽ:#/);

Note that the reduction described above only happens regularly in polysyllabic words. Monosyllables such as *kō* 'who' etc. often retain the length of their final vowel. (Apabhramsha, as a literary language, freely admits forms from Prakrit, another Middle Indic literary language, which obscures to some degree the reduction of final syllables, since Prakrit was not affected by this reduction at all).

There were no contexts in which Old Indic /e:/ and /o:/ were followed by a nasal segment at the end of a word, which is why I did not include them above. However, the regular instrumental singular ending of nominal stems in /-a/ in Apabhramsha is /-ẽ:/ (usually written ⟨-eṁ⟩ and metrically heavy), which seems to derive from the

Old Indic ending /-eːna/. This raises another question: what happens to inherited short vowels in final position in Apabhramsha? Although a thorough investigation would be outside the scope of this paper, we can note a major generalisation that we will return to subsequently (I cite the Prakrit forms as a representative of 'Common Middle Indic'):

- Endings that in Common Middle Indic have the prosodic shape ¯˘(/V:CV#/ or/VC:V#/) are regularly continued in Apabhramsha by endings of the shape ¯ or ˘˘:
  - /aːni#/ (neuter plural nominative-accusative) → /aĩ#/;
  - /eːna#/ (instrumental singular) → /ẽː#/;
  - /eːhĩ#/ (masculine-neuter instrumental plural) → /ehĩ#/ (also locative);
  - /aːhĩ#/ (feminine instrumental plural) → /ahĩ#/;
  - /aːe#/ (feminine instrumental) → /ae#/ (also note the appearance of a new genitive-dative-ablative in /Vhe#/ for feminine stems);
  - /asːa#/ (genitive singular) → /aha#/;
  - /aːnã#/ (genitive plural) → /ahũ#/;
  - /aːmi#/ (first-person plural) → /ami#/;
  - /aːmo(ː)#/ (first-person plural) → /ahũ#/;
  - /anti#/ (third-person plural) → /ahĩ#/.
- Endings that in Common Middle Indic have the prosodic shape ˘˘ (/VCV#/) are regularly continued in Apabhramsha by endings of the same prosodic shape:
  - /asi#/ (second-person singular) → /ahi#/.
  - /aha#/ (second-person plural) → /aha#/.

While the stages in the development of some of these endings are somewhat unclear and a matter of great controversy among linguists – some forms, however, being clearly analogical – we can nevertheless make a generalisation: all terminations were adjusted to match either the template ˘ (if the Common Middle Indic termination was ¯) or the template ˘˘ (if the Common Middle Indic termination was ˘˘ or ¯˘), which I will call 'template A' and 'template B'.

The prosodic form of the word had other implications. The very productive suffix *-*ka*- or *-*kī*- (Jamison 2009) was added to nominal stems – but not in the manner of a derivational suffix, which is how it had been used in Sanskrit and Prakrit, but more in the manner of an augment used to build certain case suffixes. That is, the nominative-accusative singular of all genders could take the suffix (resulting in the endings -*aü* in the masculine and neuter and -*ia* in the feminine) as well as the locative singular of the masculine and neuter (resulting in the ending -*ae*/-*aï*). These are the 'isosyllabic' cases discussed in connection with Gandhari above, and they are precisely the endings that would have taken template A (˘, deriving from a Common

Middle Indic template ¯). The suffix was not used in the remaining ('pleosyllabic') endings.

The effect of the suffix was to make the endings in template A conform to template B. The underlying motivation for this change is the same avoidance of unfooted vowels that we saw in the Sogdian Rhythmic Law: a final ˘˘ constituted a moraic trochee, whereas ˘ did not, and might end up unfooted, depending on the weight of the preceding syllable. Indeed the distribution of the extended endings corroborates this explanation: they appear "mainly after heavy syllables" (Tieken 1998, 1), i.e. in words that would have ended ¯˘ without the extension, resulting in an unfooted vowel. That is not to say that words in Apabhramsha did not or could not end in the pattern ¯˘. It only means that the use of extended endings was much more likely in words that would have ended in that shape than others.

The prosody of the stem and the endings only partly accounts for the distribution of extended endings in Apabhramsha, since – in a way that once again reminds us of Sogdian – the extended endings tend to be used with only certain grammatical categories, namely adjectives and participles. And finally, as Tieken (1998) has emphasised, Apabhramsha exhibits the synchronic availability of different diachronic stages of linguistic development, which effectively means that poets could use extended or unextended forms based on preference and metrical exigencies. In spite of the opportunity for free variation, however, the regularity with which extended endings were used only in template A endings and largely after heavy syllables suggests that prosodic considerations played an important role in the reorganisation of the inflectional system after FSR.

Finally, regarding the phonetic value of intervocalic stop consonants in Apabhramsha, it would appear superficially that Apabhramsha and Prakrit are exactly the same in this respect: namely, aspirate stops are debuccalised (/VC$^{\text{h}}$V/ becomes /VɦV/) and most non-aspirate stops are generally lost (/VtV/, /VkV/, /VdV/, /VgV/ become /VV/). But this account only covers what I will later call, following Bhayani, the second stage of Apabhramsha, which was strongly influenced by Prakrit. In the first stage, intervocalic stops are written with their voiced equivalents, whether aspirated or not, e.g. *kadhidu* for *kathitam*. Although there is no evidence for a fricative pronunciation of these consonants, these forms arguably represent a stage of the language closer to Gandhari (where *kathitam* would have developed to /kaθiðə/) than to Prakrit (where *kathitam* developed into /kaɦiaN/).

5 Analysis

Several questions can be formulated on the basis of the preliminary results above. One is whether FSR itself needs an explanation, and if so, what kind of explanation it requires. Another is whether the occurrence of FSR in Gandhari and Apabhramsha, alone (at least for a time) among the attested Middle Indic languages, requires an explanation, and if so, what kind.

At least within the study of the Indic languages, the reduction of final syllables has traditionally been seen as a natural continuation of widespread phonetic tendencies in the Indo-European languages. Nearly a century ago, Turner called attention to the "phonetic weakness" of "terminational elements" in the Indic languages. He noted that:

> From the time of Aśoka onwards the documents of Middle Indian testify to the complete disappearance of final consonants and the progressive shortening of final vowels, of which in the majority of the modern languages there is now no trace, except where their identity has been partially preserved by contraction with a preceding vowel. (Turner 1927, 229)

He also noted that "in the ancient Indo-European languages the phonetic elements following the initial consonant of the last syllable of words are subject to quite special alterations and weakenings" (Gauthiot, quoted in Turner 1927, 229).[11] If we stipulate, as a law of historical linguistics, that final elements of words are subject to various kinds of weakening and reduction, then no explanation is necessary for the loss of such elements in the Indic languages between the stage represented by Sanskrit and the stage represented by, for example, Hindi.

Earlier philological studies remained vague about the ultimate explanation for a general tendency toward loss of material at the right edge of a word. More recent research, however, has appealed to the general "articulatory and perceptual properties of speech" as a way of explaining why certain sound changes appear to be very widely attested and "natural" (Blevins 2008). In her "Field Guide" to natural and unnatural sound changes, Blevins lists "final vowel shortening" as a change with a primarily perceptual basis (referring to Myers, Hansen 2006).

11 Cf. also Tagare 1948, 49: "There seems to be a tendency in OIA to pronounce the final syllable weakly, as it was probably unaccented" (I thank an anonymous reviewer for this reference).

I note, however, that FSR in the form in which we encounter it in Middle Indic and Middle Iranic is neither identical to a generic "weakness of terminational elements" nor the specific neutralisation of vowel-length contrasts at the end of a word. As a matter of fact, the examples that Turner adduced from Aśōka's inscriptions, as well as earlier stages in the Indic languages, were examples not of the reduction of quantity or quality of final syllables, but of the lenition of intervocalic consonants in inflectional endings. The specific features of FSR discussed above require a more specific explanation, that is, one in which the prosodic structure of a word interacts with constraints on the preservation and loss of linguistic material.

Auer (1993) introduced a graded typological classification of languages based on whether the syllable or the word is more fundamental in their prosodic phonology. A wide range of phonological and morphological features have now been shown to correlate with each other, placing a language on a continuum between 'word languages' and 'syllable languages'. In a contribution to a volume on this topic, Martin Kümmel (2014b) has argued convincingly that, despite their very close relationship, the Indic and Iranic languages have "drifted" toward different ends of the word/syllable continuum. The Iranic languages have exhibited a number of sound changes that are typically associated with word languages, even from the very earliest evidence. For example, the distinction in the quantity of final vowels was neutralised in all of the Old Iranic languages (Old Persian, Avestan, and Young Avestan; Kümmel 2014b, 209). Another diagnostic feature of word phonology in the Iranic languages is complex onsets and codas relative to the Indic languages. By contrast, many of the changes that separate the Indic languages from Proto-Indo-Iranian are associated with syllable languages, including the reduction of complex onsets and codas and the implementation of a variety of *sandhi* rules both within and between words. Taken together, these features allow us to at least formulate the hypothesis that the changes diagnostic of membership in the Iranic branch of Indo-Iranian implicate the word as a significant prosodic unit, while those that indicate membership in the Indic branch implicate the syllable.

From this perspective, it is very easy to see FSR as a word-related development, given that it is at the right edge of a word that syllables are reduced. The changes implicated in FSR can thus be restated:

1. Neutralise syllable quantity at the end of a prosodic word (evident in Old Persian, Avestan, and Young Avestan; also Khotanese and Apabhramsha, and perhaps Gandhari);
2. Neutralise vowel quality or delete a vowel altogether at the end of a prosodic word:
    a. if the vowel is in an unfooted syllable (Sogdian);
    b. generally (Bactrian, Middle Persian, perhaps Gandhari).

In the pre-print version of his paper, Kümmel (2014a) does not discuss FSR *per se* (although he does note, e.g. the loss of all final syllables in Western Iranic), but he does discuss the reduction and loss of unstressed vowels in Middle Iranic in general, which contrasts very clearly with the tendency toward epenthesis in Indic.

Kümmel (2014a) also noted that some of the syllable-linked developments in the Middle Indic languages did not operate in the far northwest. Specifically, the reduction of heterorganic consonant clusters, which operates almost without exception in all of the other attested Middle Indic languages (Ardhamagadhi, Prakrit, Epigraphic Middle Indic, and with the exception of Sanskrit loanwords, Pali), bypassed the languages of the Northwest, including Gandhari as well as modern Dardic languages.

By this point it should be clear that, on the one hand, FSR is 'natural' in the sense that it is based on the articulatory and perceptual properties of speech, and therefore could, in principle, happen anywhere and at any time; on the other hand, however, FSR is part of a 'conspiracy' of sound changes and typological features associated with languages wherein the word is prominent as a unit of prosodic structure, and therefore, in fact, has been confined to languages where such changes have been able to overcome general structure-preserving constraints.

Now we can return to the distribution of FSR in the Middle Indic and Middle Iranic. As noted above, FSR occurs across-the-board in Middle Iranic and only in those Middle Indic languages that are likely to have originated in the Indo-Iranian contact zone, in Gandhara and along the Indus river. Now we can add that FSR is part of a cluster of sound changes – and typological characteristics, which we can see as the ultimate outcomes of structure-modifying sound changes – that are indeed associated with the Iranic language family as a whole, and that the Middle Indic languages that participate in FSR are, by the same token, less likely to participate in the syllable-related sound changes that affect other Middle Indic languages.

Consider, as a parallel case, the loss of final syllables in French. On the one hand, some of the antecedents of this change – the loss of final consonants in Vulgar Latin, and the subsequent reduction in quantity of all final vowels – are shared by all of the Romance languages. Yet not all Romance languages lost their final syllables outright (Vaissière 1996). In fact, it is only the Gallo-Romance branch in which this change was regular and widespread. In other branches, such as Italic, the loss of final syllables is manifestly related to proximity to the Gallo-Romance branch. This very strongly suggests that the decisive sound change, although 'natural' from one perspective, occurred in one branch of the Romance languages, and then spread, by contact, to neighbouring branches.

Thus I propose that contact with Iranic languages is one of the main reasons why FSR, among other word-related phenomena, is found in the Indic languages of the Northwest. I would go further and claim that, despite the long history of contact between Indic and Iranic speakers in that region, attested by many layers of loanwords, evidence suggests that FSR in Gandhari, at least, coincides with waves of Iranian migration in the region that began in the second/first century BCE.

I would even suggest that the extent of FSR in the languages spoken in those waves of Iranian migration – above all the Saka language – can be reconstructed based on its effects in Indic languages such as Gandhari and Apabhramsha. We have no contemporary documents in the Saka language, although words and names in that language were written in the Kharōṣṭhī and Brāhmī scripts with characteristic modifications (e.g. 'ys' for /z/). In that language, final syllables were likely all light, as they were in the oldest stage of the most closely related attested languages, Khotanese and Tumshuqese. Whether the inflectional endings were adjusted to a prosodic template is hard to say: we have observed that such an adjustment is evident in Gandhari (isosyllabic direct cases and pleosyllabic non-direct cases), and much more regularly in Apabhramsha (with two dominant prosodic templates for the ending of a word); perhaps similar phenomena were involved in Saka phonology. The absence of syncope in Gandhari suggests, too, that the Saka language did not exhibit syncope to the same extent as, for example, Bactrian.

We can now turn to other linguistic features that I have associated with FSR. I noted above that the inherited Indo-Iranian suffix *-ka-* is often used to 'repair' the effects of FSR by building new forms which are sometimes grammaticalised. By contrast, consider the explanation of Jamison (2009, 314) for the spread of the suffix in the Indic languages. She invokes the

> relentless, inexorable progress of sound change, which was steadily eroding the ends of Indo-Aryan words, and morphological change, which was streamlining the old complex series of nominal stem formants in favor of vowel-stems, especially *-a-* and *-ā-* stems. These processes set the stage for the 'real' part of the *-ka-* explosion. If *ka*-forms could always have served as lower register doublets to more dignified higher register forms, as those latter forms became threatened by linguistic erosion, speakers would have been inclined to make themselves clearer and to preserve the physical body of a word by, as it were, promoting the more colloquial *-ka*-forms into standard discourse, and also generating more of them – not to mention finding this a convenient way of avoiding consonant stems and other nasty bits of morphological business.

This is an eloquent account of some of the reasons why semantically weak suffixes can spread throughout a language.[12] But I am not sure that sound change in general is "inexorable" and "steady" (despite this serving as a useful assumption in some recent work in historical linguistics); I am convinced that all sound change is sociolinguistic in origin, and hence its progress depends very much on rhythms of interactions that philologists may not have access to. If I am correct, then we need a finer-grained account of the "*-ka*-explosion". There is of course the underlying sociolinguistic phenomenon that Jamison's article concerns in the first place, namely the use of *-ka*-forms in lower registers of Sanskrit and related languages, and their 'percolation', through lexicalisation, to higher registers. In this respect a close parallel is offered by Khotanese and (apparently) Bactrian, where some lexemes are obligatorily formed with the *-ka-* suffix. Prakrit is a bit different, in that *-ka-* can be used, optionally, with any stem at all, although there are of course preferences among different authors, genres, and periods.

The situation we see in Sogdian – where the presence or absence of *-ka-* is conditioned by grammatical category rather than by lexeme – is rare in most of the Middle Indic languages. I can think of two counterexamples. One was noted by Jamison herself: in the Gandhari of the Niya documents, which was probably in even closer contact with Iranic languages than the Gandhari of Gandhara, the past participle in *-ta-* was used for the conjugated past tense, whereas past participles functioning adnominally took the suffix *-taka-* (Jamison 2009, 317). In Apabhramsha, as we have seen, the suffix *-ka-* was virtually integrated into the inflectional endings of isosyllabic cases. And this, too, was conditioned not just by prosodic factors (being more likely after a heavy syllable) but also by grammatical category: "the long and the extended endings… are indeed found almost exclusively with adjectives and past [participles]" (Tieken 1998, 3). Some categories, such as future participles in *-tavya-*, are "always extended" (15).

Finally we can consider the development of fricative consonants. Here, too, scholars have seen the loss of intervocalic stop consonants in the Middle Indic languages as part of a general (perhaps even "inexorable" and "steady") tendency within this language family. And here, even more in the case of FSR, the change is explicable based on general principles of articulatory economy (i.e. speakers will be as lazy as their addressees allow them to be). As noted above, scholars

12 Although Jamison may be right about *-ka-* allowing speakers to "avoid" consonant stems, my impression is that inherited consonant stems were 'thematicised' in the Middle Indic languages with the simple addition of *-a-* (e.g. *śarad-* to Prakrit *saraa-*) or *-ā-* (e.g. *diś-* to Prakrit *disā-*) rather than with *-ka-*.

e-ISSN 2785-5953

sometimes envision this process in several stages: voicing of voiceless stops and fricativisation of voiced stops; fricativisation of the newly-produced voiced stops and total loss of the newly-produced fricatives; repeat (e.g. /VtV/ → /VdV/ → /VðV/ → /VV/). But it must be said that we have no evidence for a fricative pronunciation of any intervocalic stops outside of Gandhari. And it is perfectly possible for intervocalic consonants to be weakened or dropped without an intermediate phase of fricativisation. (Note that intervocalic /t/ is lenited to [ɾ] or [ʔ] in weak positions in English words like *little*, but never fricativised.) In texts written in the Brāhmī script, fricatives could not even be written without introducing new orthographic conventions that distinguished ⟨tt⟩ [t] ~ ⟨t⟩ [d] ~ ⟨d⟩ [ð], as was done for Khotanese (cf. von Hinüber 1981). And even if they were pronounced, the lack of a phonemic distinction between, say, [d] and [ð] would have encouraged a conservative orthographic practice that wrote both sounds with the same sign. All of this raises the question of why fricatives not only did develop in Gandhari but were actually written with newly-developed characters of the Kharōṣṭhī script, generally formed by the addition of a *cauda* sign (Glass 2000, 136). These characters, I repeat, were not used in Aśōka's inscriptions, but only appear in Gandhari documents of the Saka-Kuṣāṇa Age. The most straightforward explanation is that Gandhari adopted fricative sounds from neighbouring Iranic languages.

To support this view we can observe that some of the same letters that are used to represent a fricativised outcome of an inherited stop are also used to represent fricative sounds in Iranic loan words and proper names: 'vh' represents inherited */bʱ/, but also Iranic /f/ (e.g. *vharna*); 'v' and 'vh' both seem to represent Iranic /β/ (e.g. *vaka* or *vhaka*). In other cases, It has also been proposed that some conjunct consonants that look like simple continuations of inherited conjuncts, such as 'ks', are actually used, in some cases at least, to represent Iranian fricative clusters (e.g. *kṣuna* /χʃunə-/; see below on this word). In fact, the fricativisation of stops in Gandhari may have been conditioned by the very same changes in word-level prosodic phonology implicated in FSR. We know that processes of lenition are sensitive to foot structure, thanks especially to the work of Margaret Withgott (1982).

Before moving on to some of the implications of this analysis, I will mention one more possible explanation for the differences observed between the Indic languages. The classification of the Indic languages remains controversial (cf. Ivani, Paudyal, Peterson 2021). Several of the features that appear to distinctively characterise the languages of the Northwest, such as Gandhari and the modern Dardic languages, have been attributed by scholars to the membership of these languages in subgroupings where special developments occur, or where developments that commonly took place elsewhere did not

occur. The idea of an inner/outer divide between the Indic languages is particularly challenging for the analysis proposed here, because on that theory, we would expect at least some features characteristic of the Indic languages of the Northwest to be due not to contact but to their membership in a subgrouping of 'outer' languages that includes, for example, Sinhala – a language as far removed from contact with Iranic speakers as can be imagined.

Claus Peter Zoller, a proponent of the inner/outer thesis, on the one hand, attributes a cluster of phonological features found in the languages of the Northwest to their membership in the outer subgrouping: the weakness of aspiration in Gandhari, and its total loss in some modern Dardic languages as well as Sinhala (2023, 317) and the development -MP- → -P- (345).[13] On the other hand, he acknowledges that certain features of Niya Prakrit (i.e. the Gandhari used in the Niya documents) are suspiciously similar to features of Khotanese. Rather than see these features as evidence of Iranic influence, however, he sees them the other way around, as evidence of the influence of Niya Prakrit on Khotanese (2023, 368).

Zoller clearly believes that the features that set some of the Middle Indic languages apart from common Indic developments are due to membership in the outer group, and resists contact-based explanations wherever possible. I find his arguments unconvincing for several reasons. One is that he does not distinguish between sound changes and the absence of sound changes when discussing diagnostic features of the inner/outer subgrouping. It of course makes sense that languages as widely separated as Gandhari and Sinhala will, in some cases, not participate in sound changes that affect a 'central' group of languages. For example, Gandhari and Sinhala did not exhibit fortition of initial /j/ to /dž/ (Kümmel 2014a). In some cases we can attribute such conservatisms to a weakened influence of syllable-based phonology, which was evidently stronger in the central group. But some explanation is required for sound changes that affect languages widely separated in space. It may be that there are general features of the parent language of the outer languages that made it more likely for them to undergo certain sound changes – for example, a prosodic phonology that was more word-based than the syllable-based phonology of the inner languages – but these would have to be specified and weighed against alternative explanations. Another

13 A few other features he takes to be limited to the Northwest without implicating other outer languages, e.g. the continued preservation of intervocalic stops (Sinhala 2023, 345).

reason for my scepticism is Zoller's arbitrary, imprecise, and sometimes incorrect interpretation of the Middle Indic evidence.[14]

6 The Development of Gandhari

As new Gandhari documents are discovered and studied, we are likely to gain a much clearer picture of how the language developed. But the following represents what I take to be the commonly-accepted account.

Gandhari is first attested in Aśōka inscriptions at Mansehra and Shahbazgarhi in the middle of the third century BCE. For close to two centuries there are very few surviving inscriptions. But then, "around the latter half of the first century BCE, Buddhist inscriptions suddenly become very common in Gandhāra and the surrounding areas" (Salomon 2018, 29). The earlier Gandhari birch-bark scrolls, all containing Buddhist texts, date from around this period as well (although some may be somewhat older). As Salomon noted, the period of Gandhari's use as a "Buddhist literary language" in the Greater Gandhara region coincides with the period of "Indo-Scythian" rule in the northwest (Salomon 2018, 28-31; 2002, 128). Gandhari also came to be used as an administrative language in the Central Asian kingdom of Kroraina, probably because of political, economic, and religious connections between the Tarim Basin and the Greater Gandhara region during the Kuṣāṇa Empire. As a literary language, however, Gandhari was always in competition with, and influenced by, both Sanskrit and other Middle Indic languages. In South Asia it was displaced by Sanskrit by the end of the third century of the common era (Strauch 2012).

I would make two adjustments to this story, one of emphasis and one of fact. Salomon noted that Gandhari "stands apart from all of the other languages" in the Middle Indic family in regard to its phonology (2002, 119), but by this he meant its conservatism rather than its innovations: its preservation of the three-way distinction between sibilants ('ś', 'ṣ' and 's'), and its preservation of consonant clusters with 'r'. But it is important to note that the Gandhari of the Indo-Scythian period is also distinguished from other Middle Indic languages – and, for that matter, from earlier forms of Gandhari – by its innovations,

14 The invocation of Deśya Prakrit as an 'outer' language is one example (*dēśya-* simply refers to Prakrit lexemes that are not obviously derived from Sanskrit equivalents, and they are attested in [Mahārāṣṭrī] Prakrit, a language that Zoller otherwise considers an 'inner' language); cf. 2023, 364. He also attributes the 'inner' features of Gandhari (as opposed to Niya Prakrit) to the influence of Pali, which is unlikely (2023, 368), and interprets the absence of *anusvāra* in Gandhari writing to reflect a sound change -MP- → -P-, rather than what it almost certainly is, the failure to record the nasality of the syllable in writing (345).

and in particular, the quantitative reduction of all final syllables, and consequent upon this, the restructuring of the gender and case system, features I attribute above to contact with Iranic languages.

Among these distinctive features of Gandhari in the Saka-Kuṣāṇa age is the "inconsistency and lack of standardisation in [its] orthography and morphology" (Salomon 2002, 131). Salomon suggests that "[Gandhari] fell out of use" before the developments that would have established it as a standardised literary language took place (131). We can note that Gandhari is not unique among Middle Indic languages in this time: neither Ardhamagadhi, nor inscriptional Middle Indic, nor Pali were, as far as we know, subject to the kind of grammatical description and regulation that Sanskrit was. But Gandhari is uniquely chaotic in its grammar and orthography. Why? I suggest that it had been 'unstandardised' by intensive contact with speakers of Middle Iranic languages.

What do I mean by this? The example of Sanskrit and Pali might lead us to assume literary languages should be relatively uniform over a large space and a long time. But Gandhari's reinvention as a literary language in the first century BCE was premised on its use by a wide variety of people. Among these new users of the language were the political and military elite of the Northwest, many of whom held Iranian names and titles and presumably spoke Iranic languages as well. Gandhari was probably elevated to the status of a literary language by people who were not native Gandhari speakers, which resulted in a prominent and presumably prestigious variety of the language having a phonology that was basically Iranic. (Compare the use of Gandhari in Kroraina, where we presume that it was an administrative language among people whose native languages were forms of Tocharian.) This 'Iranicised' Gandhari presumably coexisted for some time with a 'non-Iranicised' variety; indeed this may be the root of the distinction between spoken and written Gandhari noted by Fussman (1989). But both were subject to a strong and persistent influence from other Middle Indic languages and Sanskrit. All of these influences probably made Gandhari quite heterogenous at exactly the time it was being cultivated as a literary language in the Greater Gandhara region.

I do not mean to say that Iranic influences were absent prior to the first century BCE. Konow (1929, cxiii) noted that the nominative masculine singular ending *-e*, previously considered to betray the influence of eastern Middle Indic languages that continued **-aḥ* as *-ē*, shows the same development of an inherited **-ah* as the neighbouring Iranic languages (Khotanese *-ä* and *-i*). It is quite possible that parallel developments such as these were reinforced by contact. Similarly, Aśōka's inscriptions in the Northwest use Iranic loanwords (see below), probably attesting to the influence of Achaemenid bureaucracy. But Gandhari was a relatively conservative Middle Indic language at

the time of Aśōka. Two centuries later, when it was adopted as a literary language among the Buddhists of the Northwest, it had undergone changes that made it the most innovative of the Middle Indic languages, and those changes were probably due to contact.

7 The Origins of Apabhramsha

In discussing Apabhramsha, we are moving several centuries ahead in time, at least as far as our evidence goes. Although the word *apabhraṁśaḥ* had been in use since at least the time of Patañjali (second century BCE) to describe forms that "fell away from" normative Sanskrit usage, the earliest source to use it as the name of a literary language is evidently Bhāmaha, author of the *Ornament of Literature* (*Kāvyālaṅkāra*), in the sixth or seventh century CE. All that Bhāmaha tells us is that there were compositions in Apabhramsha in the *mātrā* meter.[15]

The next author to mention Apabhramsha after Bhāmaha is Daṇḍin, who identifies it as the literary language of "the Ābhīras and others" in his Mirror of Literature (*Kāvyādarśa*, ca. 700 CE).[16] The association between Apabhramsha and ethnic groups of the Northwest – the Ābhīras and the Gūrjaras – is also found in a Prakrit novel of 779 CE, Uddyōtana's *Kuvalayamālā*, in which a Gūrjara traveler recites a *dōhā* in Apabhramsha.[17]

Bhayani (1998a) distinguished between two "strata" of early Apabhramsha. On the one hand, there was a corpus of literature in the *mātrā* and *raḍḍā* meters, which preserves some archaic phonological features. This corpus is totally lost, but for a few quotations in later grammars and metrical handbooks, but we know the names of several important authors: Gōvinda, Chaïlla, Śuddhaśīla. These appear to have been lyric verses, often with a pastoral character. On the other hand, there is the *sandhibandha*, a longer narrative composition with 'sections' (*sandhi*s) composed in alternating passages, called *khaḍavaka*s, of a carrying meter and a single-verse 'coda' (*ghattā*). The *sandhibandha* appears to have been a formal innovation of a poet named Caturmukha, who lived sometime before the ninth century CE, and it represents the vast majority of surviving Apabhramsha poetry. The archaic features found in the fragments of the earlier *mātrā* literature are not found in *sandhibandha*s, and Bhayani speculated that the change in literary form was linked to a

15 *Ornament of Literature* 1.16*cd*: *saṁskṛtaṁ prākṛtaṁ cānyad apabhraṁśa iti tridhā*; 1.30*ab*: *gāthāślōkamātrādi*.

16 *Mirror of Literature* 1.36*ab*: *ābhīrādigiraḥ kāvyēṣv apabhraṁśa iti smṛtāḥ*.

17 *Kuvalayamālā* § 115: 59, l. 5.

change in language: the latter poems, he said (1998a, 40), "developed under the impact of literary Māhārāṣṭrī" (i.e. Prakrit).

The two key features that distinguish the earlier from the later stratum of Apabhramsha are:

- the preservation of 'r' in consonant clusters;
- the lenition, but not complete elision, of intervocalic stops.

These features are taught by Hēmacandra in one of the earliest grammatical accounts of Apabhramsha, namely in *sūtra*s 8.4.398 and 8.4.396 in his *Siddhahēmacandraśabdānuśāsana* (ca. 1140 CE).[18] Bhayani (1998a, 36-9) notes that the examples given for most of these 'archaic' features are in the *mātrā* meter. He also notes that these features are found in quotations of Apabhramsha prior to Hēmacandra as well. These include a number of *mātrā* verses in the aforementioned *Kuvalayamālā*, and a *dōhā* given as an example in Virahāṅka's metrical handbook, the *Compendium of Mora- and Syllable-Counting Meters* (*Vṛttajātisamuccaya*, perhaps eight century). *R*-preserving forms are also given for Apabhramsha by Rudraṭa in his *Ornament of Literature* (*Kāvyālaṅkāra*, ninth century) and by Namisādhu in his commentary thereon (1068 CE).

Here is one example from the early author Gōvinda, as cited by Hēmacandra (*Siddhahēmacandravyākaraṇa* 8.4.422), in the *mātrā* meter:

ekkamekkaü jaï vi jōēdi
hari suṭṭhu savvāyareṇa
tō vi drēhi jahĩ kahĩ vi rāhī
kō sakkaï saṁvarevi
daḍḍhaṇayaṇa ṇēhēṁ paluṭṭā

Although Hari looks on each
one of them with respect, of course,
he looks at Rādhā whenever he can:
when love draws the cursed eye
somewhere, who can stop it? (Author's transl.)

Here we can see *drēhi*, possibly *drakṣyati* or some other form of the verb 'to see', with a preserved (or intrusive) 'r'. And we can also see, in *jōēdi*, an intervocalic stop consonant that has been lenited (*dyōtayati* to *jōēdi*) but not elided (compare *jōēi* in Svayambhū's version of the same verse in the *Svayambhūchandas*, 4.10.2).

18 8.4.396: *anādau svarād saṁyuktānāṁ kakhatathapaphāṁ gaghadadhababhāḥ* (with examples including *kadhidu* for *kathitam*); 8.4.398: *vādhō rō luk* (with the example *jaï bhaggā pārakkaḍā tō sahi majjhu priēṇa*).

The distinction between earlier and later Apabhramsha was made only in the 1990s by Bhayani. All of the premodern authors we have mentioned so far – Rudraṭa and Namisādhu, Virahāṅka, Svayambhū, and Hēmacandra – operate with a single linguistic category, namely Apabhramsha. But the so-called 'Eastern grammarians' recognised different varieties of Apabhramsha from a relatively early period, and one of them has been speculatively connected to this earlier form. The Eastern grammarians, identified as such by George Grierson (1924), were a group of authors based in Bengal who developed a distinctive approach to Prakrit grammar based on a finer-grained classification of regional varieties.

Kramadīśvara, whose date remains unknown, appears to follow Hēmacandra in teaching the aforementioned archaic characteristics as optional features of standard Apabhramsha (5.1-2), and includes a number of additional substitutions with intrusive 'r' (e.g. *vrāsa-* for *vyāsa-*, 5.5). However, he has a single *sūtra* that says that the 'r' is regularly retained in conjunct consonants in *vrācaṭādau*, which designates some specific varieties of Apabhramsha (5.66). He gives *sarpi* (instead of *sappi*) and *jrũ* and *drũ* (instead of *jō* and *sō*?) as examples. He gives Nāgaraka and Upanāgaraka as further varieties of Apabhramsha, the latter of which is described as a mixture of Apabhramsha and Prakrit (5.67).

Puruṣōttama, another eastern grammarian who was an exact contemporary of Hēmacandra, similarly divides Apabhramsha in general into Nāgaraka, Vrācaḍa, and Upanāgaraka varieties, except here the division is exhaustive: Nāgaraka corresponds to the 'standard' Apabhramsha taught by Hēmacandra. The Vrācaḍa variety is somewhat different from what Kramadīśvara teaches as Vrācaṭa – it converts all sibilants to 'ś', for example – but it also has the retention of both consonantal and vocalic 'r' (18.3). Mārkaṇḍēya's treatment in his *Sum-Total of Prakrit* (*Prākṛtasarvasva*, 1558-69 CE) is essentially the same, except he makes the interesting comment that Vrācaḍa "comes from Sindh" (18.1, commentary). Rāmaśarman says the same (*Wish-Granting Tree of Prakrit* [*Prākṛtakalpataru*], 3.2.1-2).

As Bubeník (1998, 28-9) noted, the preservation of 'r' in conjunct consonants is indeed a feature of the Indic languages of the Northwest, including Sindh. Together with Apabhramsha's associations with the Ābhīras, whom the Mahābhārata places in Sindh, this circumstance lends plausibility to the theory that Apabhramsha originated in the Northwest of the subcontinent, and should reflect linguistic developments specific to that region. And as noted above, FSR is a characteristic of Apabhramsha and Gandhari alone among the Middle Indic languages. This account would make the 'earlier' stratum of Apabhramsha, at least, similar to Gandhari in that some of its characteristic developments might plausibly be attributed to influence from Iranic languages. Although this conclusion appears

to follow rather straightforwardly from the data that Bhayani had gathered, he himself remained committed to the earlier view that "the Apabhraṁśa was but a colloquialised form of literary Prakrit" (Bhayani 1998b, 32), and emphasised how little we actually know about the origins and early history of the language (Bhayani 1998b, 33; 1989, 2).

8 Lexical Evidence

As an appendix to the argument above, I would like to corroborate my hypothesis about language contact by revisiting the much-discussed question of Iranic loanwords in Indic languages. Mayrhofer controversially, but for our purposes conveniently, divided his *Etymologisches Wörterbuch des Altindoarischen* into two parts, one for the older (1992; 1996) and one for the younger language (2001). By 'older language' he understands "those lexemes that are first attested in Vedic literature, or in any case in the ancient grammarians such as Pāṇini and Patañjali"; the 'younger language' includes "lexemes whose first attestation does not appear prior to the Epics and Lawbooks".[19] The 'dividing line' between these two phases is effectively what I have been calling the Saka-Kuṣāṇa Age. Hence, as a general rule, the Iranic words cited as evidence in the first two volumes of Mayrhofer's *Wörterbuch* serve to establish the lemma as inherited from Proto-Indo-Iranian. By contrast, the Iranic words cited as evidence in the third volume serve to establish the lemma as borrowed from an Iranic language. My survey of the Iranic words cited in the third volume revealed, unsurprisingly, a number of semantic clusters: words related to writing; words related to governance and administration; and military terms. Before discussing each of these clusters below, I also wanted to mention names for plants and other realia, which can often not be connected directly to an Iranic word but appear to be *Kulturwörter* of wide extension (e.g. *karkētana-*, *kuñcikā-*, *dāḍima-*, *mātuluṅga-*, *maśaka-*, *rājāvarta-*, *samūra-*, *stavaraka-*, *hiṅgu-*, *hispittha-*). Bailey noted that many words that are peculiar to Buddhist Sanskrit, and hence discussed in Edgerton's *Buddhist Hybrid Sanskrit Dictionary*, are loanwords, "which have been, as it would seem, introduced by the Iranian-speaking 'Indo-Scythians' of northwestern India in the period from the second century

19 Mayrhofer (1992, X): "Gemeint sind in der esteren Gruppe [i.e. die Lemmata der ‚älteren Sprache', AO] jene Lexeme, die in der vedischen Literatur – oder allenfalls bei alten Grammatikern wie Pāṇini oder Patañjali – erstmals belegt sind; solche Wörter, deren Erstbeleg nicht vor den Epen oder den Rechtsbüchern erscheint, werden in der Abteilung ‚jüngere Sprache' behandelt". Cf. also Mayrhofer 1983, 150.

B.C. to the fourth century A.D." (1955, 14); some of these words will be discussed below.

A particularly interesting feature of the contact situation between Iranic and Indic in the Indo-Iranian contact zone is that the languages were already quite closely related, and in some cases it would not have been immediately obvious whether a word was Iranic or Indic. One example is the word *kṣaṇa-*/*kṣuna-* in Gandhari. Both forms appear to refer to a particular time. The former is identical to Sanskrit *kṣaṇa-*, and the latter is identical to Khotanese *kṣuṇa-*. The Khotanese word, like the Gandhari word, is probably borrowed from Bactrian χϸονο, referring to a regnal year, which Tremblay (2005, 436) in turn takes to be a borrowing from Greek χρόνος.[20] It is difficult to know whether Gandhari speakers had a clear sense of the distinction.

8.1 Words Related to Writing

In the realm of writing, almost all of the core Indic vocabulary comes from Iranic loanwords, as Falk (2010a) summarises.

- *lipi-* 'writing' (Chatterji 1960, 129; Mayrhofer 1956-80, 3.103; 2001, 443-4; Falk 2010a, 212): from Old Persian *dipi*, which itself is a borrowing (probably from a word for writing originating in Sumerian *dub*, and found in Elamite *tuppi* and Akkadian *ṭuppu*). In Middle Persian and Sogdian, this word was continued by *nipīk*/*nibīg*, which yielded modern Persian *nivē* (Henning 1957). The variation between Gandhari *dipi-* and Sanskrit *lipi-*/*libi-* suggests, as Henning noted, that the word was borrowed into Indian languages from an eastern Iranian dialect that had /ðipi/. We now know that the Bactrian form was λιβο. Hence we do not even have to invoke the analogical influence of the Sanskrit verbal root *lip* 'smear' to account for the forms starting with 'l'.
- *pustaka-* 'book' (Mayrhofer 1956-80, 2.319; 2001, 331; Falk 2010a, 212): borrowed from a Middle Iranic word **pōstaka-* (itself from *pavasta-*, attested in Old Persian) meaning 'hide'. The Middle Iranian word, apparently meaning 'manuscript' or 'document', travelled widely: Sogdian (*pwstk*), Khotanese (*pūstia*), Tocharian (*postak*), Parthian (*pwstg*), and Bactrian (πωσταγο). The word is attested in Gandhari as *postaga-*, probably as a direct loanword from Iranian, and borrowed into Prakrit as *potthaya-*. Sanskrit *pustaka-*, which appears for the first time in 'classical' literature, might be a re-Sanskritisation of the Prakrit form.

20 Bailey (1979, 69) thought that it was inherited from Proto-Indo-Iranian.

- *mudrā-* 'seal' (Chatterji 1960, 128; Mayrhofer 1956-80, 2.654; 2001, 409): found throughout the Iranian languages in this meaning (Bactrian μολρογο and variants 'sealed document', Persian *muhr* 'signet ring', Khotanese *mūra-* 'coin', etc.). The word appears to come from the Old Persian word for Egypt (*mudrāya*), but Mayrhofer notes that Egypt actually borrowed the sealed letter from Western Asia. In any case, with Gandhari *mudra-*, Sanskrit *mudrā-*, and Pali/Prakrit *muddā-* we are evidently dealing with a loanword from an Iranian language.
- *divira-* 'scribe': from Persian **dipīra-* (see above; in Achaemenid documents it is only attested in Elamite *tup-pi-ra*), whence also Modern Persian *dibīr*. Used in Gandhari in the form *tipira*. The parallel Indian and Iranian terms *divīrapati-* and *dabīrbad* were used to designate a 'chief secretary' (Bailey 1949, 127-8; Falk 2010a, 213; von Hinüber 1989, 46).
- *nipista-*/*nipesita-* '(made to be) written': found in the Gandhari version of Aśōka's fourth Rock Edict at Shāhbāzgarhī. From the widely used Persian word *nipišta-* 'written, inscribed' (Falk 2010a, 209).

As Chatterji (1960) noted, many of these words appear to belong to what he designated as the first period of loanwords, viz. the time in which Gandhara was an Achaemenid territory.

8.2 Words Related to Governance

"In the second and third centuries AD, in the times of the Arsacid and Sasanian dynasties, a multitude of Iranian terms came to India, mostly denoting officials" (Falk 2010a, 212). Many of loanwords are limited, in their attestation, to either Gandhari or Sanskrit inscriptions from the Saka-Kuṣāṇa Age. One well-known example is *erzuna-* 'prince', used once in a Gandhari inscription (Takht-i-Bahi, CKI 53), equivalent to Khotanese *alysānai* (Falk 2010b, 76). Another is *ṣāhi-* 'king' and *ṣāhānuṣāhi-* 'king of kings', found in the royal titulature of the Kuṣāṇas, and clearly a continuation of Old Persian *xšāyaθiya-* and *xšāyaθiya- xšāyaθiyānām*. But the title adopted by earlier Saka rulers, *rājātirāja-* (subsequently appearing as *rājādhirāja-*, which became a standard feature of royal titulature in India), appears to be a calque of the same phrase. Two such loanwords, and a set of titles ending in *-pati-*, call for some further comment.

8.2.1 *kṣatrapa-* 'governor'

The Iranian title **xšaθra-pā-* or **xšaθra-pāvan-* 'protecting the realm' was used by regional governors (satraps) in the Achaemenid empire. The forms, presumably Median, are starred because an original **-θr-* is presupposed by the Greek (σατράπης) and Indian borrowings, whereas the Old Persian form, attested in Darius's Behistun inscription, is *xšaçapāvan-*. Its first attestations in India are all in coins and inscriptions of rulers who took this title, mostly of Iranian background, starting in the first century BCE. The title *mahākṣatrapa-* came into use around the same time. It originally designated a subordinate official, as we can see in the plate of Vasa-Abdagases (Falk 1996, 395) of Azes year 9 (ca. 48 BCE), where Vasa-Abdagases is called a *mahakṣatrava-* and Azes is called *maharaja-*. The title later came to be used for independent rulers, around Mathura, Ujjain, and Bharuch (Salomon 1974).[21]

The Gandhari pronunciation of this word would have been almost identical to the Iranian title, and its transparent etymology permitted a straightforward Sanskritisation to *kṣatrapa-* (Chatterji 1960, 129). The Sanskrit word had a relatively long life in coins and inscriptions, being in use as an official title until the end of the Kārdamaka dynasty in the later fourth century (Falk 2010b, 74). However, as Salomon (1974, 15) points out, it is never used in Sanskrit, Prakrit, or Pali literature.

8.2.2 *bhaṭāraka-* 'lord'

Falk (2010b, 75) has suggested a compelling alternative to the prevailing etymology of the word spelled as either *bhaṭṭāraka-* or *bhaṭāraka-* in Sanskrit. Rather than see it as a Middle Indic pronunciation of the Old Indic word *bhartṛ-* (Mayrhofer 1956-10, 464; 2001, 362; Sircar 1966, 52), he sees it as a continuation of an Iranian title *fratara-*, influenced by a folk etymology from *bhartṛ-*. The latter would regularly give *bhaṭṭa-*, which is well attested; the *-āraka-* has remained unexplained.

The Persian word *fratara-* or *fraθara-*, etymologically meaning 'prior', is used as a title in Achaemenid documents in Aramaic (Skjærvø 1997, 102), referring to "the administrative head of a district or province in Egypt" (Wiesehöfer 2012) or "under-satrap" (Wiesehöfer 1991, 306 *apud* Falk 2010b, 78). It appears on the legends of a series of coins

21 Bailey (1949, 127) suggested that an Iranian **xšaθra-pāvan* also underlies the title *cojhbo* (now transcribed *cozbo*) in the Niya documents; it is now taken to be from **čazdahwant* (Tremblay 2005, 429).

of the kings of Persis who broke free of Seleucid rule in the early second century BCE. Skjærvø argued that, in the phrase found on the legends (prtrk' zy 'lhy') "*fratarakā* is not exactly the title of the dynasty, but an epithet stating priority of the king among others of divine descent" (1997, 102). When Persis was incorporated into the Parthian empire later in the second century BCE, the king, Mithridates I, allowed the local kings of Persis a degree of autonomy.[22]

As far as I know, the word is not used prior to the first century BCE in Sanskrit, Pali, or Ardhamagadhi texts. Falk (2010b, 75) noted that the earliest use of the title in India might be on a Greek legend of a coin issued by Higaraka, who ruled around the middle of the first century BCE. The legend reads BAPTAP///, which Falk interpreted as *bartar*[*akos*], corresponding to an unattested Gandhari **vhartaraka-*, in turn adapted from *frataraka* with metathesis of the 'r', likely influenced by popular etymology from *bhartṛ-*. The form **vhartaraka-* might have subsequently developed into *bhaṭaraka-*. The length of the vowel in Brāhmī writing (*bhaṭāraka-*) may reflect a stress accent in the Iranian source word.

In the first century CE, *bhaṭaraka-* is found in Gandhari inscriptions as a title, much like *kṣatrapa-*. The copper plates of Helagupta (CKI 564), dated to Azes 121 (i.e. 73/74 CE), seem to refer to Yodavharṇa as a *bhaṭaraka* (so Falk 2014, 21; Salomon 2020 takes it as a proper name). The word *bhaṭaraka* occurs on a sandstone reliquary (CKI 536) dated to Azes 147 (i.e. 99/100 CE), and Baums (2012, 238) takes it there as a title. It remained in use as a title and honorific into Kuṣāṇa times. In the Spinwam inscription (CKI 244), dated to year 39, presumably of the Kuṣāṇa era (hence ca. 166 or 266 CE; cf. Falk 2009, 29), the Kuṣāṇa king is called *bhaṭarakasami* (adopting Falk's readings). A Brāhmī inscription dated to the 45th year of Huviṣka (ca. 195 CE or so), now in the Chhattrapati Shivaji Museum in Mumbai, uses the word *bhaṭārikā* (Lüders 1961, § 180 = 205). The inscription records the donation of an image of Śākyamuni in the Rōśikavihāra at Āḷikā by Khvasicā.[23] The image is dedicated to the good health of a number of people, among whom *bhaṭārikā* figures, either referring to a separate person ("his mistress", as per Lüders), or modifying one of the other people as a title ("the mother of Śamaṇikā, the *bhaṭārikā*").

When we come to the Central Asian kingdom of Niya, around the third century CE, *bhaṭaraka-* was used both for lower officials (*cozbo*s and *soṭhaṁga*s) as well as king Aṁgoka (in the Endere inscription,

22 Cf. Wiesehöfer 2012; 2013 and Wiesehöfer 1994, 105-8 for the title.

23 Khvasicā is probably a Saka name, involving the diminuitive suffix *ica*, attested in Khotanese as *īca* (Degener 1989, 128), perhaps after a word meaning 'first' from **fravišta-*, like Khotanese *hvaṣṭa-* (Bailey 1979, 505).

CKI 363, Salomon 1999a) and his successor Mahiri (Boyer, Rapson, Senart 1920, nos. 415, 573). A number of documents make clear that *bhaṭaraga-* is a complementary term to *daza-* 'slave' or 'servant': *taha na dharma bhaṭaragasa tanu dazasa ṛnena giṁnidavya* (Boyer, Rapson, Senart 1920, no. 24, translated by Burrow 1940 as "such is not the law, that the property of the master should be taken for the debt of the slave"); [*da*]*zajaṁna bhaṭaragasa* (Boyer, Rapson, Senart 1920, no. 31; Burrow 1940: "slave people and the […] people of the master").

By the fourth century or so, the title *bhaṭāraka-* (more often *bhaṭṭāraka-*) had been fully absorbed into Indian royal titulature, where it characterises the Guptas (all *paramabhaṭṭāraka-*) and their queens (all *bhaṭṭārikā-*); it is also found in contemporary inscriptions at Chilas (von Hinüber 1989, 53-4). By this time it had perhaps already spread to Indonesia, where it is always a title of a god, and spelled with the more conservative spelling *bhaṭāraka*.

8.2.3 Words Ending in *-pati*

Around the turn of the common era, we find many words ending in *-pati-* as designations for military, political, or bureaucratic positions. On the one hand, this could be explained by reference to earlier Indic lexical patterns (Vedic *vākpáti-*, *vācáspáti-*, *bṛ́haspáti-*, *bráhmaṇaspáti-*) or indeed lexical patterns of the protolanguages (Vedic *viśpáti-* = Avestan *vispaiti-* < Proto-Indo-Iranian **wić-pati-*; Vedic *dámpati-* = Avestan *dəṇg paiti-* = Greek δεσπότης < Proto-Indo-European **dms-poti-*). On the other hand, titles ending in *-pati-* and its cognates were (and remain) widespread throughout the Iranian world, and many such words were borrowed directly into Indian languages at this time. Here are a few examples:

- *gañjapati-* 'treasurer': from an Eastern Iranian form **ganza-pati-*, like its synonym *gañjavara-* (from **ganza-bara-*, cf. Persian *ganjwar*). Used in "hybrid" Sanskrit inscriptions (Damsteegt 1978, 255), Gandhari (Bailey 1949, 127) and Kashmiri Sanskrit.
- *bakanapati-* 'temple attendant': from **baγanə-pati-*, 'master of the gods'. Used in "hybrid" Sanskrit inscriptions from the Kuṣāṇa period (Damsteegt 1978, 255).
- *haysārapati-* 'chiliarch': a military title attested from the Upper Indus Valley (from **hazāra-pati-*; von Hinüber 1986, 149; Falk 2010b, 78).
- *navhapati-* 'clan-master': used as a title by the kings of Oḍi, from **nāfa-pati-*. The first word, though cognate with Sanskrit *nābhi-* 'navel', refers to a clan or family in Iranic languages (Sodgian and Middle Persian *nāf*). The whole compound was loaned into Armenian as *nahapet* (Bailey 1980, 25; Falk 2010b, 75).

- *divīrapati-* 'chief secretary': found in inscriptions of the fifth century from the Upper Indus and in the colophons of the Gilgit manuscripts (von Hinüber 1986, 149), as well as in Kashmiri Sanskrit (the *Rājataraṅgiṇī*). Based on the word *divīra* 'scribe' or 'secretary' discussed above; compare Persian *dabīrbad*.
- *haṁmārapati-* 'chief accountant': in inscriptions of the Upper Indus and in the colophons of the Gilgit manuscripts (von Hinüber 1986, 150).
- *naśćīrapati-* 'hunt-master' (Mayrhofer 2001, 285 compares Parthian ⟨nhšyrpty⟩ /naxcir-pati/ 'Jagdmeister' and Middle and New Persian *naxčīr* 'Jagd'); attested in Kuṣāṇa-era inscriptions.

These words are clearly loans from Iranic languages, given the first element. I pass over a number of other words ending in *-pati-* found in Kuṣāṇa-era inscriptions, since their interpretation and etymology is not secure, but these are likely loans from Iranic languages as well (*kharāsalērapati-*, *manapākapati-*; Falk 2010b, 78).

The word *dānapati-* 'master of giving' i.e. 'patron' is attested in Sanskrit from the *Mahābhārata* onwards. Lüders (1961, 95-6) suggested that this word has an Iranian equivalent attested in Brāhmī inscriptions as *hōramurṇḍaga-* (with much variation) and in Kharōṣṭhī inscriptions as *horamurta*, if this word means 'master of gifts' (cf. Khotanese *hōra-* 'gift'). There are some problems with this interpretation, as Lüders himself admitted. But I am inclined to see this, and a number of Sanskrit words ending in *-pati-*, as calques of Iranian titles.[24] In other words, the use of the element *pati* in compounds referring to the 'chief' person in a particular role or office corresponds with, and is probably influenced by, the use of the etymologically-identical element *bed* (*vel sim.*) in Iranic languages. This tendency would of course have been reinforced by earlier Sanskrit compounds ending in *pati*, such as *sēnāpati-* (first attested in the *Aitarēyabrāhmaṇa*).

One problematic example is the word *sthapáti-* 'architect', which is first attested in the Atharvavēda. It presents certain problems of derivation as a Sanskrit word (namely the use of *stha-* as the initial element of a compound), and hence Mayrhofer (1996, 764) suggests that it may be an "Umformung eines [iran.?] LW", probably on the basis of other Iranian loanwords in *-pati-*. If Mayrhofer is right, this would indicate that the process of borrowing such words from neighbouring Iranic languages had already begun prior to the Saka-Kuṣāṇa Age.

24 I do not include *dārapati-*, which von Hinüber (1986, 149) suggested to read in several inscriptions from the Upper Indus, since he subsequently changed the reading to *dānapati-* (1989, 56). Still, the context of the word in these inscriptions (alongside Iranian terms such as *divīra-*) suggests it comes from the sphere of Iranian influence.

8.3 Military Terms

When scanning the third volume of Mayrhofer's *Etymologisches Wörterbuch des Altindoarischen* I noted a relatively large number of words that pertain to military equipment and roles, which I will merely list here, with the corresponding pages in Mayrhofer (2001); the references to cognates are his; I have tried to find the earliest attestations.

- *aśvavāra-* 'horse-rider' (p. 18) (Old Persian *asabāra*, East Middle Iranic **aspa-βāra-*). On this word, first attested in label inscriptions at Bharhut (second century BCE, Chatterji 1960, 129), cf. especially Morgenstierne 1974, 275-6.
- *khōla-* 'helmet' (p. 148) (Bāṇa's *Kādambarī* and *Harṣacarita*, seventh century) (Old Persian *xaudā-*, Avestan *xaoδa-*, East Middle Iranic **xōla-*).
- *tīrī-* 'arrow' (p. 248) (Halāyudha's *Abhidhānacintāmaṇi*, tenth century) (Persian *tīr*, from Old Persian *tigra-*).
- *druṇā-* 'bow' (p. 274) (Halāyudha) (Middle Persian *drōn*, Khotanese *durna-*).
- *nipaka-* 'pledge' or 'hostage' (p. 291) (*Divyāvadāna*, third century) (Sogdian ⟨npq⟩ , Khotanese *nvi-*); cf. Bailey 1955, 18; the Kashmiri Sanskrit word *nīvī*, used in the sense of 'hostage' in the *Rājataraṅgiṇī*, is probably also related.
- *paryāṇa-* 'saddle' (pp. 307-8) (Varāhamihira, sixth century) (Sogdian *pyrδn*); Bailey 1955, 14.
- *padāti(ka)-* (Mayrhofer 1996, 79): attested already in the *Taittirīyabrāhmaṇa*, which evidently made Mayrhofer reluctant to consider it an Iranic loanword outright despite the close parallel with Middle Persian *payādag* and Persian *piyāda*; Morgenstierne has "no doubt" that the word was borrowed from Persian (1974, 275: 8).
- *padika-* 'footsoldier' (p. 303) in the *Amarakōśa*: derivable from Sanskrit *pada-*, but Mayrhofer mentions the possibility that it might come from Iranian **padik*, continued by Persian *payg* 'footsoldier' (see above).
- *lastaka-* 'bow-grip' (p. 441) in the *Amarakōśa* ("Veilliecht iran. *δasta(k) *'Griff', *'Handstelle'")
- *vārabāṇa-* 'cuirass' (p. 467) in the *Amarakōśa* ("Iran. **varo-pāna* 'Brustschutz'").

8.4 Names

There are of course many Iranian names attested in Indian inscriptions and manuscripts during the Saka-Kuṣāṇa Age, and in certain areas of the Indo-Iranian border regions (e.g. Gilgit) they continued to appear alongside Indian names well into the first millennium. There is no point in listing the names here, but I did want to make two points about the impact of Iranic languages on the Indian onomasticon.

First, in a multilingual and multiethnic culture, it might not have been obvious which elements are 'Indic' and which elements are 'Iranic' (Morgenstierne's 1974, 271). This is especially true of cognate elements, such as the words derived from Proto-Indo-Iranian **priHa-*: Indic *priya-* and Iranic *friya-* must have sounded very similar. Hence it is unsurprising to have names such as Indrafriya (*iṁtavhria* in CKI 60) in Gandhara, where we might have expected Indrapriya. But even elements such as *spāla-*, the Eastern Iranic word for 'army', were essentially treated as Indic words, the exact equivalent of *sēnā-*, in names such as Suśpāla (= Suṣēṇa; Falk 2006, 396) and Dharmaspāla (= Dharmasēna; von Hinüber 1986, 151). The emergence of Sanskrit as a 'cosmopolitan' language in the period immediately following the Saka-Kuṣāṇa Age might explain the disappearance of Iranic elements from the onomasticon, both because the formal study of the Sanskrit language would have made their foreignness more apparent, and because names were increasingly regulated by the norms of both Sanskrit grammar and the various *dharmaśāstra*s. Consider, for example, the names of the Kārdamaka kings of Ujjayinī: the first few are completely Iranic (Zamōtika, Caṣṭana), and then they switch over to Indic names (Jayadāman, Rudradāman), with a few ambiguously Iranic elements (Dāmazāda = Dāmajāta; Tandon 2009).

Second, it appears that many Iranic names were given an *interpretatio indica*, and appear in this form in Sanskrit texts. I found all of these in Mayrhofer's *Etymologisches Wörterbuch*; more could certainly be found.

Parṇadatta, governor of Saurāṣṭra mentioned in Skandagupta's Junāgaṛh inscription, is probably a Sanskritisation of the common Iranic name Farnadāta (Charpentier 1931; Mayrhofer 2001, 306). *Parṇa-* stands in for *farna-* also, evidently, in Ṛtuparṇa = *Ṛtafarnah, a king of Kośala (Mayrhofer 2001, 38). As in the case of Parṇadatta, Indic *datta-* probably corresponds to Iranic *dāta-* in the name of a character in the *Mahābhārata,* Bhagadatta = *Bagadāta (Mayrhofer 2001, 360), since the theophoric element more naturally represents Iranic *baga-* 'god' than Indic *bhaga-*, which might suggest something untoward. Similarly, in the *Mahābhārata*, the king of Sauvīra in the Lower Indus Valley is called Dattāmitra, recognizable as the common Iranic name Dātamiθra (compare Mithradates, Miθradāta; Mayrhofer 2001, 259), the equation being more likely given the importance of

Mithra as an Iranian god. Mayrhofer also noted (281) that the name of Sahadēva's son, Dhruvāśva, might represent an attested Iranic name Druwasp (Λροοασπο in Bactrian and *druuaspa* in Avestan). Finally, Mēghākṣa, the Persian king allied to Malayakētu in the *Mudrārākṣasa*, probably contains the element **xšaya-* 'rule' (Mayrhofer 2001, 414, who connects the first element to Pamir *mēγ* 'name of a mythical horse').

9 Conclusions

I expect that my argument about the linguistic influence of Middle Iranic languages on the Indic languages of the Northwest will be sufficiently clear by now. I will conclude, then, with three features of the analysis offered here that might have broader historical or methodological implications.

First, contact is an important feature in the history of language, and a major focus of linguistic research in the present day. But the study of historical languages, that is, languages to which we have access only through written records, approaches linguistic change by default through the model of accumulated changes over time that differentiate one speech community from another, i.e. the *Stammbaum* model. The study of contact in historical languages, and especially in the historical languages of South Asia, is much less developed than the study of neogrammarian sound change, and it often figures in 'marginal' cases where words or forms cannot be explained through neogrammarian sound change; there is a whole genre of Indological research attempting to identify words in Sanskrit (and rarely other languages) that have come from Dravidian, Munda, or other 'substrate' languages.

There are good reasons for this imbalance, of course. We have neither the plenum of evidence that would tell us, for example, about whether and how the linguistic practices of communities in contact differ from each other. Nor do we have the kind of evidence that would tell us, unambiguously, how certain forms were pronounced, and indeed as we have seen with the Kharōṣṭhī script, there are a number of different phonological interpretations of the orthography. Nevertheless, the linguistic circumstances prevalent in South Asia – where 'linkages' of related languages are likely to be found, and where areal phenomena have already been documented – should invite us to consider other features of the historical languages as possibly resulting from contact-induced change. One parallel case is the change of 's' to 'h' in Greek, Armenian, and Iranian (Parpola 2002, 82).

Second, philology – here narrowly understood as the study of historical languages – can and should avail itself of new concepts and methods. By 'new' I do not mean to suggest that prosodic phonology,

which I have used in this paper, is new, but only that its application to the study of historical languages is still limited. In fact prosodic phonology in particular can be an important tool for detecting contact phenomena, precisely because the 'same' language can be pronounced with different prosodic features by different communities of speakers. Kümmel's work (2014) correlating the features of Indic and Iranic languages with the parameters of syllable and word languages, respectively, is exemplary in this respect.

Third and finally, Apabhramsha figures in this analysis as the 'bridge' whereby sound changes produced by contact in a specific region at a specific time continued to be represented in the literary record of India in the later first millennium. That is, as Gandhari was superceded as a literary language by Sanskrit, and as Middle Iranic languages receded from the linguistic horizons of India, Apabhramsha remained as the single surviving Indic language with an Iranic phonology. Now the *Sprachwirklichkeit* of Apabhramsha is a large and complex issue. Most philologists have seen Apabhramsha as a somewhat crystallised or frozen form of a popular language (Bubeník 1998), and some have seen the 'underlying' popular language(s) as very widely distributed vernaculars that were the antecedent to the modern vernaculars of North India (cf. Ghosal 1956). Whatever spoken language(s) to which Apabhramsha was connected might, in any analysis, have served as the 'vector' by which a number of important sound changes entered the linguistic communities of North India. The loss of final vowels is of course found in all of the modern North Indian vernaculars, and for that reason it might seem attractive to analyse it as the result of tendencies or pressures internal to the Indic languages. But it was not inevitable of course. If we require an explanation for it, the long-term influence of FSR from Iranic languages could be considered.

Bibliography

Primary Texts

Gandhari documents and inscriptions are cited by their CKI or CKD number on `http://gandhari.org` (maintained by Stefan Baums and Andrew Glass).

Bhāmaha (6th c. CE?). "Ornament of Literature (Kāvyālaṅkāra)". Trivedî, K.P. (ed.) (1909), *The Pratâparudrayośobhûshaṇa of Vidyânâtha. With the Commentary, Ratnâpaṇa, of Kumârasvâmin, Son of Mallinâtha, and With a Critical Notice of Manuscripts, Introduction, Critical and Explanatory Notes and an Appendix Containing the Kâvyâlaṅkâra of Bhâmaha*. Bombay: Government Central Press.

Daṇḍin (ca. 700 CE). "Mirror of Literature (Kāvyādarśa)". Thakur, A.; Jha, U. (eds) (1957), *Kavyalakṣaṇa of Daṇḍin (Also Known as Kāvyādarśa) with Commentary Called Ratnaśrī of Ratnaśrījñāna*. Darbhanga: Mithila Institute of Post-Graduate Studies and Research in Sanskrit Learning.

Hēmacandra (12th c. CE). "Siddhahēmacandraśabdānuśāsana". Pischel, R. (Hrsg.) (1877), *Hemacandra's Grammatik der Prâkritsprachen (Siddhahemacandram Adhyâya VIII), mit kritischen und erläuternden Anmerkungen*. Halle: Waisenhaus.

Kramadīśvara (s.d.). "Distilled Essence (Saṁkṣiptasāra)". Banerjee, S.R. (ed.) (1980), *Prākṛtādhyāya (The Prākṛta Book of the Samkṣiptasāra) by Kramadīśvara*. Ahmedabad: Prakrit Text Society. Prakrit Text Series 22.

Mārkaṇḍēya (1560s CE). "Sum-Total of Prakrit (Prākɼtasarvasva)". Acharya, K.C. (ed.) (1968), *Mārkaṇḍeya's Prākṛta-Sarvasva*. Ahmedabad: Prakrit Text Society. Prakrit Text Society Series 11.

Puruṣōttamadēva (12th c. CE). "Prakrit Grammar (Prākɼtānuśāsana)". Nitti-Dolci, L. (éd.) (1938), *Le Prākṛtānuśāsana de Puruṣottama*. Paris: Société Asiatique.

Rāmaśarman (s.d.). "Wish-Granting Tree of Prakrit (Prākɼtakalpataru)". Ghosh, M. (ed.) (1954), *Rāmaśarman's Prākṛta-kalpataru*. Calcutta: The Asiatic Society. Bibliotheca Indica 278.

Uddyōtana (778 CE). "Kuvalayamālā". Upadhye, A.N. (ed.) (1959), *Uddyotanasūri's Kuvalayamālā (A Unique Campū in Prakrit)*. Part I, *Kuvalayamālā, Prākrit-Text & Various Readings*. Bombay: Bhāratīya Vidyā Bhavana. Singhi Jain Series 45.

Secondary Literature

Auer, P. (1993). *Is a Rhythm-Based Typology Possible? A Study of the Role of Prosody in Phonological Typology*. Konstanz: Universität Konstanz. KontRI Working Paper 21. `http://paul.igl.uni-freiburg.de/auer/userfiles/downloads/Phonotypo_Kontri1.pdf`.

Back, M. (1978). *Die sassanidischen Staatsinschriften*. Téhéran: Bibliothèque Pahlavi.

Bailey, H.W. (1946). "Gāndhārī". *Bulletin of the School of Oriental and African Studies*, 11(4), 764-7. `https://doi.org/10.1017/S0041977X00089801`.

Bailey, H.W. (1949). "Indo-Iranica II". *Bulletin of the School of Oriental and African Studies*, 13(1), 121-39. `https://doi.org/10.1017/S0041977X0008188X`.

Bailey, H.W. (1955). "Buddhist Sanskrit". *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 87(1-2), 13-24. https://doi.org/10.1017/S0035869X00106975.

Bailey, H.W. (1963). *Khotanese Texts V*. Cambridge: Cambridge University Press.

Bailey, H.W. (1979). *Dictionary of Khotan Saka*. Cambridge; New York: Cambridge University Press.

Bailey, H.W. (1980). "A Kharoṣṭrī Inscription of Seṇavarma, King of Oḍi". *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 112(1), 21-9. https://doi.org/10.1017/S0035869X00135877.

Bashir, E. (ed.) (2016). "Contact and Convergence: The Northwest". Hock, H.H.; Bashir, E. (eds), *The Languages and Linguistics of South Asia: A Comprehensive Guide*. Berlin: De Gruyter Mouton, 264-99.

Bashir, E. (2006). "Indo-Iranian Frontier Languages". *Encyclopædia Iranica*. https://iranicaonline.org/articles/indo-iranian-frontier-languages-and-the-influence-of-persia.

Baums, S. (2009). *A Gāndhārī Commentary on Early Buddhist Verses: British Library Kharoṣṭhī Fragments 7, 9, 13 and 18* [PhD dissertation]. Seattle: University of Washington.

Baums, S. (2012). "Catalog and Revised Texts and Translations of Gandharan Reliquary Inscriptions". Jongeward, D.; Errington, E.; Salomon, R.; Baums, S. (eds), *Gandharan Buddhist Reliquaries*. Seattle: Early Buddhist Manuscripts Project, 200-51.

Bhayani, H.C. (1998a). "On Early Apabhraṁśa". *Indological Studies: Literary and Performing Arts, Prakrit and Apabhraṁśa Studies*, vol. 2. Ahmedabad: Parshva Publication, 35-43.

Bhayani, H.C. (1998b). "The Rise of Apabhraṁśa as a Literary Language". *Indological Studies: Literary and Performing Arts, Prakrit and Apabhraṁśa Studies*, vol. 2. Ahmedabad: Parshva Publication, 26-34.

Bhayani, H.C. (1989). *Apabhraṃśa Language and Literature: A Short Introduction*. Delhi: Bhogilal Leherchand Institute of Indology.

Blevins, J. (2008). "Natural and Unnatural Sound Patterns: A Pocket Field Guide". Willems, K.; De Cuypere, L. (eds), *Naturalness and Iconicity in Language*, 121-48.

Boyer, A.M.; Rapson, E.J.; Senart, E. (1920). *Kharoṣṭhī Inscriptions Discovered by Sir Aurel Stein in Chinese Turkestan. Part I, Text of Inscriptions Discovered at the Niya Site*. Oxford: Clarendon Press.

Brough, J. (1962). *The Gāndhārī Dharmapada*. London: Oxford University Press.

Bubeník, V. (1998). *A Historical Syntax of Late Middle Indo-Aryan (Apabhraṃśa)*. Amsterdam; Philadelphia: John Benjamins.

Burrow, T. (1940). *A Translation of the Kharoṣṭhi Documents from Chinese Turkestan*. London: The Asiatic Society.

Caillat, C. (1992). "Connections Between Asokan (Shahbazgarhi) and Niya Prakrit?". *Indo-Iranian Journal*, 35(2-3), 109-19.

Cantera, A. (2009). "On the History of the Middle Persian Nominal Inflection". Sundermann, W.; Hintze, A.; de Blois, F. (eds), *Exegisti monumenta: Festschrift in Honour of Nicholas Sims-Williams*. Wiesbaden: Harrassowitz, 17-30.

Charpentier, J. (1931). "Pherendates–Parṇadatta". *Journal of the Royal Asiatic Society*, 63(1), 140-1. https://doi.org/10.1017/S0035869X00066296.

Chatterji, S.K. (1960). "Some Iranian and Turki Loans in Sanskrit". Dil, A.S. (ed.), *Shahidullah Presentation Volume*. Lahore: Linguistic Research Group of Pakistan, 123-40.

Damsteegt, T. (1978). "Epigraphical Hybrid Sanskrit: Its Rise, Characteristics and Relationship to Buddhist Hybrid Sanskrit". *Orientalia Rheno-Traiectina*, vol. 23. Leiden: Brill.

Degener, A. (1989). *Khotanische Suffixe.* Wiesbaden: Steiner.

Dresher, B.E.; Lahiri, A. (1991). "The Germanic Foot: Metrical Coherence in Old English". *Linguistic Inquiry*, 22(2), 251-86.

Emeneau, M.B. (1956). "India as a Linguistic Area". *Language*, 32(1), 3-16.

Emeneau, M.B. (1980). "The Indian Linguistic Area Revisited". Dil, A.S. (ed.), *Language and Linguistic Area: Essays by Murray B. Emeneau*. Stanford: Stanford University Press, 197-249.

Emmerick, R.E. (1987). "The Transition from Old to Late Khotanese". *Transition Periods in Iranian History = Actes du Symposium de Fribourg-en-Brisgau* (22-24 mai 1985). Leuven: Association pour l'avancement des études iraniennes, 33-42.

Emmerick, R.E. (2009). "Khotanese and Tumshuqese". Windfuhr, G. (ed.), *The Iranian Languages*. Milton Park: Routledge, 377-415.

Falk, H. (1996). "Three Inscribed Buddhist Monastic Utensils from Gandhāra". *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 156, 393-412.

Falk, H. (2006). *Aśokan Sites and Artefacts: A Source-Book with Bibliography*. Mainz: Von Zabern.

Falk, H. (2009). "The Pious Donation of Wells in Gandhara". Mevissen, G.; Banerji, A. (eds), *Prajñādhara: Essays on Asian, History, Epigraphy and Culture in Honour of Gouriswar Bhattacharya*. New Delhi: Kaveri Books, 23-6.

Falk, H. (2010a). "Foreign Terms in Sanskrit Pertaining to Writing". de Voogt, A.; Finkel, I. (eds), *The Idea of Writing: Play and Complexity*. Leiden: Brill, 207-17.

Falk, H. (2010b). "Names and Titles from Kuṣāṇa Times to the Hūṇas". Alram, M.; Klimburg-Salter, D.; Inaba, M.; Pfisterer, M. (eds), *Coins, Art and Chronology II: The First Millennium C.E. in the Indo-Iranian Borderlands*. Vienna: Österreichische Akademie der Wissenschaften, 73-90.

Falk, H. (2014). "The First-Century Copper-Plates of Helagupta from Gandhāra hailing Maitreya". *Annual Report of the International Research Institute for Advanced Buddhology at Soka University*, 17, 3-26.

François, A. (2015). "Trees, Waves and Linkages: Models of Language Diversification". Bowern, C.; Evans, B. (eds), *The Routledge Handbook of Historical Linguistics*. Abingdon: Routledge, 161-89. https://doi.org/10.4324/9781315794013.

Fussman, G. (1989). "Gāndhārī écrite, Gāndhārī parlée". Caillat, C. (éd.), *Dialectes dans les littératures indo-aryennes*. Paris: Collège de France, 433-98.

Gershevitch, I. (1954). *A Grammar of Manichean Sogdian*. Oxford: Blackwell.

Gholami, S. (2014). "Historical phonology of Bactrian". *Selected Features of Bactrian Grammar*. Wiesbaden: Harrassowitz, 31-71.

Gholami, S. (2009). *Selected Features of Bactrian Grammar* [PhD Dissertation], Georg-August-Universität Göttingen: Göttingen.

Ghosal, S.N. (1956). "Controversy Over the Significance of Apabhraṃśa and a Compromise Between the Views of Jacobi and Grierson". *Journal of the Asiatic Society. Letters*. 12, 23-37.

Glass, A. (2000). *A Preliminary Study of Kharoṣṭhī Manuscript Paleography* [MA Thesis]. Seattle: University of Washington.

Grierson, G. (1924). *The Prakrit Dhātv-ādeśas According to the Western and the Eastern Schools of the Prakrit Grammarians*. Calcutta: Asiatic Society of Bengal.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Henning, W.B. (1957). "The Inscriptions of Tang-i Azao". *Bulletin of the School of Oriental and African Studies*, 20, 335-42.

Hintze, A. (2006). Review of *Le Y final dans les inscriptions moyen-perses et la 'loi rythmique' proto-moyen-perse*, by Huyse, P. *Kratylos*, 51, 161-7. http://dx.doi.org/10.29091/kratylos/2006/1/13.

von Hinüber, O. (1981). "Die Paiśācī und die Entstehung der Sakischen Orthographie". Bruhn, K.; Wezler, A. (Hrsgg), *Studien zum Jainismus und Buddhismus: Gedenkschrift für Ludwig Alsdorf*. Wiesbaden: Franz Steiner, 121-7.

von Hinüber, O. (1986). "Zu einigen iranischen Namen und Titeln aus Brāhmī-Inschriften am oberen Indus". *Studia Grammatica Iranica: Festschrift für Helmut Humbach*. München: R. Kitzinger, 147-62.

von Hinüber, O. (1989). "Brāhmī Inscriptions on the History and Culture of the Upper Indus Valley". Jettmar, K. (ed.), *Antiquities of Northern Pakistan: Reports and Studies (Band 1): Rock Inscriptions in the Indus Valley*. Mainz: Philipp von Zabern, 41-72.

Hitch, D. (2016). *The Old Khotanese Metanalysis* [PhD dissertation]. Cambridge (MA): Harvard University.

Huyse, P. (2003). *Le y final dans les inscriptiosn moyen-perses et la 'loi rythmique' proto-moyen-perse*. Paris: Association pour l'avancement des études iraniennes.

Ivani, J.K.; Paudyal, N.; Peterson, J. (2021). "Indo-Aryan – A House Divided? Evidence for the East-West Indo-Aryan Divide and its Significance for the Study of Northern South Asia". *Journal of South Asian Languages and Linguistics*, 7(2), 185-202. https://doi.org/10.1515/jsall-2021-2029.

Jamison, S. (2009). "Sociolinguistic Remarks on the Indo-Iranian *-ka-Suffix: A Marker of Colloquial Register". *Indo-Iranian Journal*, 52, 311-29. http://dx.doi.org/10.1163/001972409x12562030836615.

Kielhorn, F. (1905-06). "Junagadh Inscription of Rudradaman". *Epigraphia Indica*, 8, 36-49.

Konow, S. (1929). *Kharoshṭhī Inscriptions, With the Exception of Those of Aśoka*. Calcutta: Government of India.

Korn, A. (2013). "Final Troubles: Armenian Stem Classes and the Word-end in Late Old Persian". Tokhtasev, S.; Luria, P. (eds), *Commentationes Iranicae: Vladimiro f. Aaron Livschits nonagenario donum natalicum*. Petersburg: Nestor-Historia. 74-91.

Korn, A. (2019). "Isoglosses and Subdivisions of Iranian". *Journal of Historical Linguistics*, 9, 239-81. https://doi.org/10.1075/jhl.17010.kor.

Kümmel, M. (2008). *Khotansakisch* [Handout]. https://www.academia.edu/1748443/Introduction_to_Khotan_Saka_in_German_.

Kümmel, M. (2014a). "Syllable- and Word-related Developments in Earlier Indo-Iranian." Pre-print version of Kümmel 2014b. https://www.academia.edu/379001/Syllable_and_word_related_developments_in_earlier_Indo_Iranian.

Kümmel, M. (2014b). "Syllable- and Word-Related Developments in Earlier Indo-Iranian". Reina, J.C.; Szczepaniak, R. (eds), *Syllable and Word Languages*. Berlin: De Gruyter, 204-21.

Kümmel, M. (2018). "Areal Developments in the History of Iranic: West vs. East". https://ia801207.us.archive.org/15/items/dokumen.tips_areal-developments-in-the-history-of-iranic-west-vs-east/

dokumen.tips_areal-developments-in-the-history-of-iranic-west-vs-east.pdf.

Loukota, D. (2022). "'Ne Hāḍe Vajrropamä Vaśärä': Indic Loanwords in the Khotanese Book of Zambasta and the Chronology of the Spread of Buddhism to Khotan". *Buddhist Road: Dynamics in Buddhist Networks in Eastern Central Asia 6th-14th Centuries*, 1(7), 3-32.

Lévi, S. (1903). "Notes on the Indo-Scythians". *Indian Antiquary*, 32, 381-9, 417-26.

Lüders, H. (1961). *Mathurā Inscriptions*. Göttingen: Vandenhoeck & Ruprecht.

MacKenzie, D.N. (1982). Review of *Die sassanidischen Staatsinschriften*, by Back, M. *Indogermanische Forschungen*, 87, 280-98.

Masica, C. (1976). *South Asia: Defining a Linguistic Area*. Chicago: University of Chicago Press.

Mayrhofer, M. (1956-80). *Kurzgefasstes etymologisches Wörterbuch des Altindoarischen*. Heidelberg: Carl Winter.

Mayrhofer, M. (1983). "Überlegungen zu einem neuen etymologischen Wörterbuch des Altindoarischen". Bammesberger, A. (Hrsg.), *Das etymologische Wörterbuch: Fragen der Konzeption und Gestaltung*. Regensburg: Friedrich Pustet, 147-54.

Mayrhofer, M. (1992, 1996, 2001). *Etymologisches Wörterbuch des Altindoarischen*, 3 Bde. Heidelberg: Carl Winter.

Morgenstierne, G. (1974). "Early Iranic Influence upon Indo-Aryan". *Commémoration Cyrus. Hommage Universel*, vol. 1. Téhéran; Liège: Bibliothèque Pahlavi, 271-9.

Morgenstierne, G. (1975). "Ancient Contacts Between N.E. Iranian and Indo-Aryan?". *Mélanges linguistiques offerts à Émile Benveniste*. Paris: Société Linguistique de Paris, 431-4.

Müller, M. (1883). *India: What Can It Teach Us?*. London: Lovell's Library.

Myers, S.; Hansen, B.D. (2006). "The Origin of Vowel Length Neutralization in Final Position: Evidence from Finnish Speakers". *Natural Language and Linguistic Theory*, 25(1), 157-93.

Nespor, M.; Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.

Novák, Ľ. (2013). *Problem of Archaism and Innovation in the Eastern Iranian Languages* [PhD dissertation]. Prague: Charles University.

Parpola, A. (2002). "From the dialects of Old Indo-Aryan to Proto-Indo-Aryan and Proto-Iranian". Sims-Williams, N. (ed.), *Indo-Iranian Languages and Peoples*. Oxford: Oxford University Press, 43-102.

Salomon, R. (1974). "The Kṣatrapas and Mahākṣatrapas of India". *Wiener Zeitschrift für die Kunde Sudasiens*, 18, 5-25.

Salomon, R. (1999a). "A Stone Inscription in Central Asian Gāndhārī from Endere (Xinjiang)". *Bulletin of the Asia Institute*, 13, 1-13.

Salomon, R. (1999b). *Ancient Buddhist Scrolls from Gandhāra*. London: The British Library.

Salomon, R. (2002). "Gāndhārī and the Other Indo-Aryan Languages in the Light of Newly Discovered Kharoṣṭhī Manuscripts". Sims-Williams, N. (ed.), *Indo-Iranian Languages and Peoples*. Oxford: Oxford University Press, 119-34. Proceedings of the British Academy 136. https://doi.org/10.5871/bacad/9780197262856.003.0005.

Salomon, R. (2018). *The Buddhist Literature of Ancient Gandhara*. Somerville (MA): Wisdom.

Salomon, R. (2020). "The Copper Plates of Helagupta: A New Edition and Study". *Indo-Iranian Journal*, 63, 3-69. http://dx.doi.org/10.1163/15728536-06301006.

Sims-Williams, N. (1984). "The Sogdian 'Rhythmic Law'". *Middle Iranian Studies: Proceedings of the International Symposium Organized by the Katholieke Universiteit Leuven from the 17th to the 20th of May 1982*. Leuven: Peters, 203-15.

Sims-Williams, N. (1990). "Chotano-Sogdica II: Aspects of the Development of Nominal Morphology in Khotanese and Sogdian". Gnoli, G.; Panaino, A. (eds), *Proceedings of the First European Conference of Iranian Studies*, vol. 1. Rome: Istituto italiano per il Medio ed Estremo Oriente, 275-96.

Sims-Williams, N. (1996). "The Sogdian Manuscripts in Brāhmī Script as Evidence for Sogdian Phonology". Emmerick, R.E. (ed.), *Turfan, Khotan und Dunhuang: Vortrage der Tagung Annemarie v. Gabain und die Turfanforschung', veranstaltet von der Berlin-Brandenburgischen Akademie der Wissenschaften*. Berlin: Akademie Verlag, 307-15.

Sims-Williams, N. (2011). "Bactrian Language". *Encyclopædia Iranica*, 3(4), 344-9. https://www.iranicaonline.org/articles/bactrian-language.

Sircar, D.C. (1966). *Indian Epigraphical Glossary*. Delhi: Motilal Banarsidass.

Skjærvø, P.O. (1997). "The Joy of the Cup: A Pre-Sasanian Middle Persian Inscription on a Silver Bowl". *Bulletin of the Asia Institute*, 11, 93-104.

Skjærvø, P.O. (2007). "An Introduction to Manichean Sogdian". https://web.archive.org/web/20150106094817/http://www.fas.harvard.edu/~iranian/Sogdian/sogdiancomplete.pdf.

Smith, L.C.; Ussishkin, A. (2015). "The Role of Prosodic Templates in Diachrony". Honeybone, P.; Salmons, J. (eds), *The Oxford Handbook of Historical Phonology*. Oxford: Oxford University Press, 262-86. https://doi.org/10.1093/oxfordhb/9780199232819.013.028.

Strauch, I. (2012). "The Character of the Indian Kharoṣṭhī Script and the 'Sanskrit Revolution': A Writing System Between Identity and Assimilation". de Voogt, A.; Quack, J.F. (eds), *The Idea of Writing: Writing Across Borders*. Leiden: Brill, 131-68.

Tagare, G.V. (1948). *Historical Grammar of Apabhraṁśa*. Poona: Deccan College Post-Graduate and Research Institute.

Tandon, P. (2009). "The Western Kshatrapa Dāmazāda". *The Numismatic Chronicle*, 169, 173-87.

Tieken, H. (1998). "The Distribution of the Variant Endings -u, -aü, and -ā in the Apabhraṃśa Verses in Hemacandra's Prākrit Grammar". *Indo-Iranian Journal*, 41(1), 1-31.

Tikkanen, B. (2008). "Some Areal Phonological Isoglosses in the Transit Zone between South Asia and Central Asia". Ud-din, I. (ed.), *Proceedings of the Third International Hindu-Kush Cultural Conference*. Oxford: Oxford University Press, 250-62.

Tremblay, X. (2005). "Irano-Tocharica and Tochario-Iranica". *Bulletin of the School of Oriental and African Studies*, 68, 421-49. https://doi.org/10.1017/S0041977X05000248.

Turner, R.L. (1927). "The Phonetic Weakness of Terminational Elements in Indo-Aryan". *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 59(2), 227-39. https://doi.org/10.1017/S0035869X00057439.

Vaissière, J. (1996). "From Latin to Modern French: On Diachronic Changes and Synchronic Variations". *AIPUK. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel,* 31, 61-74.

Wiesehöfer, J. (1991). "PRTRK, RB μYL', SGN und MR". Sancisi-Weerdenburg, H.; Kuhrt, A. (eds), *Asia Minor and Egypt: Old Cultures in a New Empire = Proceedings of the Groningen 1988 Achaemenid History Workshop*. Leiden: Nederlands Instituut voor het Nabije Oosten, 305-9.

Wiesehöfer, J. (1994). *Die'dunklen Jahrhunderte' des Persis: Untersuchungen zu Geschichte und Kultur von Fārs in frühhellenisticher Zeit (330-140 v. Chr.)*. München: Beck.

Wiesehöfer, J. [2000] (2012). "Frataraka". *Encyclopædia Iranica*. https://iranicaonline.org/articles/frataraka.

Wiesehöfer, J. (2013). "Fratarakā and Seleucids". Potts, D.T. (ed.), *The Oxford Handbook of Ancient Iran*. Oxford: Oxford University Press, 718-27.

Withgott, M. (1982). *Segmental Evidence for Phonological Constituents* [PhD Dissertation]. Austin: The University of Texas at Austin.

Zoller, C.P. (2018). "'Pagan Christmas: Winter feast of the Kalasha of the Hindu Kush' and the true frontiers of 'Greater Peristan'". *Acta Orientalia*, 79, 163-337.

Zoller, C.P. (2023). *Indo-Aryan and the Linguistic History and Prehistory of North India*. Wiesbaden: Harrassowitz. Neuindischen Studien 20. https://doi.org/10.5617/ao.7672.