



Universiteit
Leiden

The Netherlands

Traces of language contact in Niya Prakrit: Bactrian and other foreign elements

Schoubben, N.

Citation

Schoubben, N. (2024, November 6). *Traces of language contact in Niya Prakrit: Bactrian and other foreign elements*. Retrieved from <https://hdl.handle.net/1887/4108454>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4108454>

Note: To cite this publication please use the final published version (if applicable).

CHAPTER 2 Phonological preliminaries

This second chapter treats selected topics in the (historical) phonology of Niya Prakrit (§2.1) and Bactrian (§2.2). These preliminaries are needed for the later arguments about Bactrian loanwords into Niya Prakrit and the sound substitutions shown by them (treated in Chapters 3–4).

The main difficulty with the phonological interpretation of Niya Prakrit is the large sign inventory of the Kharoṣṭhī script. Sometimes, one phoneme can be indicated by two or even three akṣaras, but at the same time, many akṣaras have more than one phonological value, depending on the position in the word and the etymological source of the represented sound.

The issue with Bactrian is that there were more phonemes than there were graphemes in the Graeco-Bactrian alphabet. As a result, some letters had to be used for more than one sound, thus requiring one to determine when which letter represents which phoneme. Mostly, arguments from historical phonology and comparisons with the orthography of a Bactrian text in Manichaean script (M 1224) are sufficient to determine this. Yet, sometimes, the phonological interpretation is complicated by the late date and fragmentary nature of the Manichaean evidence.

For a summary of the phonological values which I assume for Niya Prakrit, see Table 1 (p. 37); for Bactrian, see Table 2 (p. 62).

2.1 Kharoṣṭhī and Niya phonology

2.1.1 Kharoṣṭhī and its characteristics

The history of the Gāndhārī language is inextricably connected to the history of the Kharoṣṭhī script. Together with Brāhmī, Kharoṣṭhī is one of two abugida scripts (= alphasyllabaries) developed in South Asia in the last few centuries before the common era.³⁰ Brāhmī was from its beginnings conceived as a pan-Indian script. Kharoṣṭhī, by contrast, never lost its

³⁰ Despite many proposals, the etymology of the name “Kharoṣṭhī” has never been clarified (Falk 1993: 84–90; Salomon 1998b: 50f.; Glass 2000: 10f.). Kharoṣṭhī had its own alphabetical order, called *arapacana*, because it started with the sequence 𑀧 *a*, 𑀦 *ra*, 𑀢 *pa*, etc. (Thomas 1950; Brough 1977 = 1996: 450–460; Salomon 1990; 1993; Baums 2009: 194–197). In modern scholarship, the *varṇamālā* system known from Brāhmī is used instead. Modified akṣaras are placed behind the unmodified form, e.g. 𑀧𑀢 *ga* /ɣ/ after 𑀧 *ga* /g/.

association with the northwestern regions, from where the script also spread to large parts of Central Asia.³¹ Being attested from the 3rd century BC until at least the 4th century AD,³² Kharoṣṭhī naturally developed over time, a process which included the creation of new signs in order to more faithfully represent certain phonological (and phonetic) distinctions.³³

Kharoṣṭhī is written from right to left. This is one out of many indications of a close connection with the Aramaic script, of which Kharoṣṭhī most likely represents a reworking to make the script suit the needs of an Indo-Aryan language (Falk 1993: 77–81; 84–105; Salomon 1998b: 51–54; Glass 2000: 11–20). As the chancery language of the Achaemenid empire, Aramaic was used over a vast territory stretching from present-day Egypt to Afghanistan and NW India (Gzella 2015: 157–211; 2023: 139–183). So far, no archive of administrative documents in Achaemenid Official Aramaic has been discovered in Gandhāra. The discovery of such an archive in neighbouring Bactria (Naveh & Shaked 2012 = ADAB) makes it, however, plausible that they must have existed in Gandhāra too. Further support for this hypothesis comes from the Aramaic translations of some Aśokan edicts and the continuation of Aramaic writing habits and phraseology in Gāndhārī (e.g. Sims-Williams 1996b: 81; Yakubovich 2006: 338; Baums 2014; 2022).³⁴

Three defining features of the Kharoṣṭhī script sometimes hamper our understanding of the phonology of Niya Prakrit.

First, Kharoṣṭhī orthography has never been completely standardised (see e.g. Fussman 1989).³⁵ Inconsistencies include, for instance, writing an

³¹ For an overview of Kharoṣṭhī materials discovered outside of the Gandhāran heartland, see now Salomon (2024). In this article (p. 537), Salomon also convincingly rejects Falk’s recent proposal that an ancient knife hilt excavated in Austria (Wels, ancient Ovilava) would be written in a Central Asian variety of Kharoṣṭhī (Falk, Miglbauer & Pfahl 2023: 202–208).

³² The Kuča Prakrit materials have been argued to be more recent, but their dating is difficult to establish. See e.g. Ching (2013: 86f.), who cautiously dates them to the 5th or 6th century AD.

³³ A comprehensive palaeography of Kharoṣṭhī, taking the various newly discovered materials into account, remains to be written. See Glass (2000) for a preliminary investigation. Although outdated in some respects, see also KD (295–322) for the specific type of Kharoṣṭhī used in the Niya documents. For conjunct consonants unique to the Niya corpus, see now Glass (2024).

³⁴ The Aramaic heritage in Gāndhārī deserves more comprehensive research in the future.

³⁵ It still seems that the Niya documents are slightly more standardised than many of the literary documents in South Asian Gāndhārī (cf. also §7.3).

anusvāra (᳚) where it is etymologically unjustified, e.g. in opt.2sg. *preṣeyamṣi* ‘you should send’ for ***preṣeyasi* (CKD 399), or leaving it out where it is historically expected, e.g. in *abhyadara* ‘inside’ instead of ***abhyamḍara* < *abhyantara-* (CKD 291, cf. also LKD: §47).³⁶

Second, geminate consonants are generally left unmarked in Kharoṣṭhī. Consequently, one has to base oneself on arguments from historical phonology and comparisons with other MIA languages to decide when, for instance, ᳚ *cha* writes a single consonant /c^h/ or a geminate /cc^h/.

Third, vowel length is only inconsistently indicated. In its earliest forms, the Kharoṣṭhī script could not distinguish short from long vowels at all. By the time of the Niya documents, a length mark had been adopted which could be used to mark a vowel as long, e.g. ᳚ *ā* next to ᳚ *a* or ᳚ *kā* next to ᳚ *ka*. Yet, the scribes did not use this length mark with any consistency (see e.g. KD: 298f.; Glass 2000: 36f.).

An unfortunate consequence of the last two features is that we do not know for sure whether Gāndhārī participated in the MIA sound change known as the *Zweimorengesetz* or “law of two morae” (von Hinüber 2001²: §§108–110). This law states that “overlong” syllables consisting of a long vowel followed by more than one consonant were generally prohibited in MIA (but note Turner 1967; 1973 = 1975: 405–415; 430–435). Accordingly, either the vowel or the consonant cluster could be shortened, as in Pāli *issara-* ‘lord’ < *īśvara-* or *dīgha-* ‘long’ < *dīrgha-*. The currently available data do not allow us to reasonably confirm the extent to which the same procedures were active in Gāndhārī (see Baums 2009: 121–124). To account for this uncertainty, I will, in phonological transcriptions of Gāndhārī words, mark originally long vowels followed by more than one consonant with the anceps signs *ā̃*, *ī̃*, *ū̃*, *ē̃*, and *ō̃*.

Short /a/ was, as in OIA, probably phonetically realised as a schwa (Allen 1953: 57–61; Baums 2009: 112 fn. 26), but because of its phonemic relationship with long /ā/, I will still use /a/ and not **/ə/ in phonological transcriptions. Word-final vowels seem to have been pronounced only weakly, if at all (LKD: §13; Fussman 1989: 45).

**

The following sections will briefly introduce some of the phonological changes characteristic of Niya Prakrit and their relation to developments in

³⁶ Unetymological anusvāras are also common in Khotanese.

Gāndhārī more generally.³⁷ The discussion consists of two parts. §§2.1.2–2.1.5 each present one feature of Niya (historical) phonology, i.e. the lenition of intervocalic consonants; confusion of aspirated and unaspirated stops; confusion of voiced and unvoiced stops; and the treatment of stops after nasals. §§2.1.6–2.1.7 respectively concern the phonological value of some selected akṣaras and the transliteration of the Kharoṣṭhī sign 𑀘 as *jha*.

2.1.2 The lenition of intervocalic consonants

In line with the phonological developments in contemporary MIA languages, most of the OIA intervocalic aspirated stops lost their occlusion in Niya Prakrit, i.e. they became *-h-* (LKD: §27).³⁸ Some examples are *suha* ‘pleasure’ < OIA *sukha-*, *lahu* ‘light, small’ < OIA *laghu-*, and *parihaṣa* ‘claim’ < OIA *paribhāṣā-*. At the same time, it is possible to find historical spellings with the aspirates preserved, e.g. *lekha* ‘letter’ < OIA *lekha-*.³⁹

The treatment of OIA intervocalic *-th-* and *-dh-* requires more discussion. In Niya Prakrit and some other forms of Central Asian Gāndhārī, the default outcome of *-th-* and *-dh-* appears to be *-h-*, as exemplified by frequent words such as *taha* ‘so’ < OIA *tathā* and *gohoma* (*gohomi*, also contracted *goma*) ‘wheat’ < OIA *godhūma-* (LKD: §27; 88 s.v. *goma*; Loukota 2023: 10). Historical spellings of the type *tatha* exist as well. In other dialects of Gāndhārī, OIA *-th-* and *-dh-* underwent spirantisation to /z/, e.g. *tasa* / *taṣa* ‘so’ < OIA *tathā* or *krosa* / *kroṣa* ‘anger’ < OIA *krodha-* (cf. e.g. Salomon 2000: 85; 2008: 114; 339f.).⁴⁰ At least for *-dh-*, there are occasional instances of the

³⁷ For more details on Niya Prakrit phonology, see LKD (§§1–50). For the (historical) phonology of other types of Gāndhārī, see Baums (2009: 110–200).

³⁸ Palatals are unaffected by this process. For the lenition of intervocalic retroflexes, see below. On similar sound changes in other MIA languages, see e.g. Bubeník (1996: 56f.); von Hinüber (2001²: §§187–194); De Clercq (2009: 41).

³⁹ Contrast *leharaga* ‘courier’ < **leharaga* < OIA *lekha-hāraka-*, whose haplogy must postdate the lenition of *-kh-* to *-h-*. For confusions between aspirated and unaspirated stops, see §2.1.3.

⁴⁰ Probably, *-dh-* (and *-th-*) first became a fricative **-ḍ-*, before developing further to /z/ (Brough 1962: 95f.; Pinault 1987a: 143; Baums 2009: 144f.). The dir.pl. *sasumate* ‘considered good’ < OIA *sādhumata-* in SHĀH I 2 (the parallel in MĀN I 3 has *sa[dhu]mata*) has been considered an early instance of the change to /z/ (cf. e.g. Caillat 1992: 110; Salomon 1999a: 126 “anticipates the later development”). One could also hypothesise that the akṣara 𑀘 *su* in *sasumate* was intended as an approximation for **-ḍu-* because intervocalic 𑀘 *da* probably still indicated a

same reflex in Niya Prakrit, i.e. *masu* ‘wine’ < OIA *madhu-* (§3.7.1 s.v.); *asimatra* / *aṣimatra* / *ajhimatra* ‘above measure’ < OIA *adhimātram*; and *anavarajhi* ‘faultless’ < OIA *anaparādh(y)a-* (LKD: 110 s.v. *masu*; TKD: 74; also §2.1.7). Since *-dh-* > *-h-* seems regular in Niya Prakrit, these examples may be loans from a different dialect of Gāndhārī.

Intervocalic non-aspirated stops, voiced and unvoiced, were affected by a process of lenition which generally resulted in voiced fricatives, even though historical spellings are not infrequent (see LKD: §§16–23).

OIA intervocalic *-k-* and *-g-* became a fricative /ɣ/, written in Kharoṣṭhī with the sign 𑀧 *ga* (older transcription *ḡa*). Two such examples are *anega* ‘many’ < OIA *aneka-* and *aroga* ‘health’ < OIA *aroga-* (LKD: §16).

The lenited outcome of OIA intervocalic *-c-* and *-j-* could be written with 𑀧 *śa* or 𑀧 *ja* (older transcription *ḡa*), thus suggesting a phonological value /ʒ/.⁴¹ This lenition can be exemplified by words such as *praśura* ‘abundant’ < OIA *pracura-* or *yajitaga* ‘requested’ < OIA **yācitaka-* (LKD: §17).

The sound resulting from intervocalic weakening of the retroflex stops *-ṭ-* and *-ḍ-* is spelled with 𑀧 *ḍa*, phonologically most likely /r̥/ (see §2.1.6).⁴² This sound change is seen in *kukuḍa* ‘cock’ < OIA *kukkuṭa-*, *vaḍavi* ‘mare’ < OIA *vaḍavī-*, and other examples (LKD: §18). The aspirated *-ṭh-* of OIA *vihethayati* ‘(s)he injures’ underwent the same change, having become Niya *viheḍeti*. The only inherited word in Niya Prakrit containing an intervocalic *-ḍh-* is *oṣaḍha* ‘medicinal herb’ < OIA *auṣadha-* (CKD 225). In view of the lenition process undergone by the other intervocalic retroflexes, it is conceivable that *-ḍh-* stands here for /r̥^h/ (so also Baums 2009: 144 on other forms of Gāndhārī).⁴³

plosive /d/ in Aśokan Gāndhārī (cf. Fussman 1989: 464), and there was thus no straightforward way to write /ḍ/.

⁴¹ In a few sources in South Asian Gāndhārī (e.g. CKI 48; 174; 176; 328), the phoneme /ʒ/ is written with the special sign 𑀧 *śa* (Glass 2000: 100).

⁴² Because of this lenition to /r̥/ and by lack of an example in initial position, 𑀧 *ṭa* is used in Niya Prakrit for a geminate /t̪t̪/. Burrow (LKD: §18) noted some exceptions to this rule, but they can all be explained differently. The reading of *coṭaga* ‘clothing’ < *coṭaka-* (CKD 316) and *viheṭa* ‘injury’ < *vihetha-* (CKD 621, 2x) can be corrected to *coḍaga* and *viheḍa*, respectively. For *guṭa* ‘hidden’ < *gūḍha-* (CKD 17), I propose to change the reading to *guja* ‘idem’ < *guhya-* (as in CKD 376). The element *ṭera* in *saṃghaṭera* ‘community elder’ (CKD 419) is probably a mistake for ***thera* < *sthavira* (cf. LKD: 93 s.v.).

⁴³ The same lenition happened in NIA languages, e.g. Hindi *paūrḥā* ‘stomach’ < **prauḍhaka-* (CDIAL: no. 9021). Further support for the existence of /r̥^h/ in (later

For OIA intervocalic *-t-* and *-d-*, the akṣaras ᵛ *ta* and ᵝ *da* were used indiscriminately, e.g. *lihitaga* besides *lihidaga* ‘written’ < OIA **likhitaka-* (LKD: §19). Because the other intervocalic consonants had been weakened to voiced fricatives, I assume, with Baums (2009: 141), that ᵛ *ta* and ᵝ *da* wrote /δ/ in this type of context.⁴⁴ In some other forms of Kharoṣṭhī, intervocalic /δ/ could be written with the special signs ᵛ̣ *ṭa* and ᵝ̣ *ḍa* (older transcriptions *ṭa* and *ḍa*; cf. Glass 2000: 77; 79f.).

The unaspirated labial stops *-p-* and *-b-* were lenited to /v/, as shown by lexemes such as *avi* ‘also, too’ < OIA *api* and pres.3pl. *pivaṃti* ‘they drink’ < OIA *pibanti*. Historical spellings were common nevertheless (LKD: §20).⁴⁵

Lenition also affected intervocalic sibilants. The outcome of OIA *-ś-* can, in between vowels, be written with either ᱦ *śa* or ᱪ *ja*, e.g. *daśa* / *daja* ‘ten’ < OIA *daśa*. As the sign ᱪ *ja* has the phonological value /ǰ/ (cf. supra), one can conclude that intervocalic *-ś-* became a voiced /ǰ/ (LKD: §21).

Because of the voicing of intervocalic *-ś-* (and *-s-*), one suspects retroflex *-ṣ-* to have become a voiced /z/ as well. Niya Kharoṣṭhī did not have a separate sign for /z/, and the outcome of OIA intervocalic *-ṣ-* is simply spelled with ᱦ *ṣa*, as in *śeṣa* ‘remainder’ < OIA *śeṣa-* (LKD: §23). Some other types of Kharoṣṭhī, including the one used in the Khotan Prakrit document CKD 661, made use of a special sign ᱦ̣ *ṣa* to write the phoneme /z/ unambiguously (cf. e.g. Baums 2009: 150).⁴⁶

Between vowels, OIA *-s-* became /z/ (cf. LKD: §22), a phoneme which could be spelled with ᱦ̣ *ṣa*, ᱪ̣ *ṣa*, or ᱪ̣ *jha* (see in more detail §2.1.7).⁴⁷ A word appearing in all three spellings is *dasa* / *daśa* / *dajha* ‘slave’ < OIA *dāsa-*.

forms of) Gāndhārī comes from the recent discovery of a Kharoṣṭhī sign ᱦ̣ *ḍha* in the Bajaur fragments (Schlosser 2022: 55). The sign *ḍha* found in foreign names in Niya Prakrit, e.g. *ḍhagiya*, is palaeographically distinctive; it probably indicates a different, yet still undetermined sound (Burrow 1935b: 668).

⁴⁴ In a few frequent words, intervocalic dentals are sometimes even weakened to \emptyset , e.g. in *caūra* ‘four’ < OIA *catur-*.

⁴⁵ Occasionally, *avi* ‘also, too’ < OIA *api* results further into *ami* (e.g. CKD 189). In this case, *-m-* conceivably stands for a nasalised labial approximant [ũ], which, like Baums (2009: 138f.), I consider to be an allophone of /v/ and /m/.

⁴⁶ In CKD 661, Boyer et al.’s (1920–29: 249) readings *maṣa* (2x) ‘coins’, *niravaśiṣo* ‘without remainder’, and *ajiṣanayi* ‘at the request of’ are to be corrected to *maṣa*, *niravaśiṣo*, and *ajiṣanayi*, respectively.

⁴⁷ The few cases when ᱪ̣ *ṣa* occurs word-initially are in words of an enclitic nature, such as *ṣaca*. See LKD (§22) and Glass (2000: 107–109).

2.1.3 Confusion of aspirated and unaspirated consonants

Except for the intervocalic lenitions described above, OIA aspirated consonants were normally preserved unchanged in Niya Prakrit. Like in South Asian Gāndhārī, there is nevertheless an occasional tendency for the aspiration to be lost, both word-initially and in other positions (LKD: §§24f.). Examples include *gaṣa* besides *ghaṣa* ‘fodder’ < OIA *ghāsa-*; *grida* besides *ghrida* ‘ghee’ < OIA *ghṛta-*; *śavata* besides *śavatha* ‘oath’ < OIA *śavatha-*; and *kumba* instead of **kumbha* ‘jar’ < OIA *kumbha-*. Although etymologically correct forms still make up the majority of cases, instances like these may suggest that aspiration was no longer phonologically distinctive. Indeed, scribes sometimes also wrote aspirated consonants where unaspirated ones are expected etymologically, e.g. *dhura* besides *dura* ‘away, far’ < OIA *dūra-* or *upaśamgh-* besides *upaśamg-* ‘to suspect, to worry’ < OIA *upa + √śamk* (LKD: §26). These confusions between aspirated and unaspirated consonants have been plausibly ascribed to Iranian influence (cf. KI: ci; Salomon 1999a: 128; 2008: 340; 2016: 375f.; §5.1.2).

2.1.4 Confusion of voiced and unvoiced stops

Confusions of voiced and unvoiced stops can also occur, including in Anlaut (Burrow 1935b: 667f.; LKD: §§14f.). Unvoiced stops can sometimes be written instead of a voiced one, e.g. *kilane* besides *gilani* / *gilanaga* < OIA *glāna(-ka-)*; *can-* besides *jan-* ‘to know’ < OIA *√jñā-*; and *tura* besides *dura* ‘away, far’ < OIA *dūra-* (LKD: §14). Occurrences are sporadic, and the etymologically expected spelling is typically more frequent. Yet, there is at least one well-attested word which consistently undergoes initial devoicing, i.e. *palýi* ‘tax’ < OIA *bali-* (ca. 90 attestations; cf. Burrow 1935b: 675; LKD: §14; Lüders 1936b: 647). At times, voiced stops can also be written instead of an etymologically expected unvoiced stop, e.g. ins.sg. *dena* besides *tena* ‘therefore’ < OIA *tena* (LKD: §15). Confusion of voiced and unvoiced consonants is typical of Niya Prakrit, but there are rare examples of the same phenomenon in other forms of Gāndhārī (cf. §6.3).

2.1.5 The treatment of consonants after nasals

In some variants of Gāndhārī, including Niya Prakrit, etymologically unvoiced consonants became voiced after a nasal, while originally voiced (aspirated) consonants were assimilated to a preceding nasal (LKD: §§45f.; Baums 2009: 159; 2015: 76). In the variety of Central Asian Gāndhārī used in

the *Khotan Dharmapada*, these sound changes are regularly observed in writing, whereas in the Niya documents historical spellings frequently occur (Brough 1962: 98f.).⁴⁸ Niya examples of the first change are *upaśaṃg(h)*- besides *upaśaṃk*- ‘to suspect, to worry’ < OIA *upa* + $\sqrt{śaṅk}$ and *mahaṃda* besides *mahaṃta* ‘big’ < OIA *mahat*-.⁴⁹ Instances of the second are *chiṃn*- ‘to cut’ < OIA \sqrt{chid} (cf. Pāli *chindatī*) and *baṃn*- ‘to bind’ < OIA \sqrt{bandh} .

Because only the etymologically voiced (aspirated) consonants were assimilated to a preceding nasal, this change needs to be dated before the voicing of originally unvoiced consonants after nasals. To what extent the latter process had been phonologised is hard to say. In the phonological representations of Niya words, I tentatively assume it had.

2.1.6 The phonological value of selected akṣaras

What follows is a list of akṣaras (i) which are specific to Niya Kharoṣṭhī or at least used there more frequently than elsewhere (𑖦 *kṣā*; 𑖧 *lyā*; 𑖨 *va*; 𑖩 *śpa*; 𑖪 *śpa*); (ii) whose phonological value needs more extensive argumentation (𑖫 *tha*; 𑖬 *ṭha*; 𑖭 *ḍa*); or (iii) which are used slightly differently than in other forms of Gāndhārī (𑖮 *ka*; 𑖯 *na*; 𑖰 *vha*). For each akṣara, I shall indicate the most common source(s) of the sound it represents; summarise the arguments supporting the proposed phonological (or sometimes phonetic) value; and point out if its transliteration has changed over time.⁵⁰

𑖮 *ka*

The sign 𑖮 *ka* is formed by adding a vertical stroke on top of 𑖫 *ka* (Glass 2000: 52f.; Schoubben 2021: 48f.). *ka* most often corresponds to OIA *-sk-*, e.g. in *namaḱaro* ‘homage’ < OIA *namaskāra*-. In the *arapacana* poem in Bajaur fragment 5 (CKM 268) to be published by Gudrun Melzer (Munich), the same akṣara can reflect OIA *-skh-* or *-ṣk-*, viz. in *kaḷiṭa* / *kaḷamaṇa* ‘stumbling’ < \sqrt{skhal} and *duḱara* ‘difficult’ < *duṣkara*- (Schoubben 2021: 49f.). In Niya

⁴⁸ Some scholars assume these sound changes to be restricted to Central Asian Gāndhārī (e.g. Barchi & Peschl 2022: 426f.). I will return to this issue in §5.4.6.

⁴⁹ I am unaware of a Niya Prakrit example whereby a voiceless aspirate becomes voiced after a nasal, but there are examples of this type in the *Khotan Dharmapada*, e.g. abs. *saghaï* < OIA *saṃkhyāya* ‘having considered’ (DHP^K 68).

⁵⁰ Compound akṣaras whose transliteration has also changed over time but which are not discussed here are 𑖱 *tsa* (older transliteration *nka*); 𑖲 *tga* (*nḡa*); 𑖳 *śpa* (*n̄sa*); 𑖴 *rma* (*mā*); 𑖵 *kla* (*kra*); 𑖶 *pla* (*pra*); and 𑖷 *śla* (*śra*).

Prakrit, *ka* is further used to substitute Bactr. -*pk-* /*šk*/, as exemplified by *maḱa* ‘bean’ ← **μαḱko* /*māškə*/ and other words (cf. §3.2.2 s.v. and §4.3.18). I therefore assume that in Niya Prakrit *ka* could stand for either /*sk*/ or /*šk*/.⁵¹ Yet, one should not exclude Brough’s idea (1962: 75) that it stood instead for “a sound developed from an earlier *ska* [or *ška*]”.

𑀓 *kṣa*

In Niya Kharoṣṭhī, 𑀓 *kṣa* occurs much more often than 𑀔 *kṣa*, the sign commonly used in other forms of the script; the only palaeographic difference is the horizontal line added on top of 𑀓 *kṣa*.⁵² These signs are used to represent the outcome of OIA -*kṣ-*, e.g. *kṣetra* (*kṣetra*) ‘field’ < OIA *kṣetra-*, or as a substitution of Ir. *-*xš-*, e.g. *kṣuna* (*kṣuna*) ‘date, reign, regnal year’ ← Bactr. *χpovo* /*xšunə*/ (§4.3.18). Given that *kṣa* and *kṣa* freely interchange in the Niya documents, both may have been pronounced in the same way (KD: 302; Glass 2000: 116). Yet, it is possible that there originally was a distinction between the two: *kṣa* may have been a way to write the Iranian cluster *-*xš-* unambiguously, while *kṣa* stood for the outcome of OIA -*kṣ-* in inherited words (Hitch 1984: 198–202). Based on Gāndhārī loanwords into Chinese and evidence from Dardic historical phonology, some scholars assume that OIA -*kṣ-* yielded Gandh. [tʃ] (e.g. Bailey 1946a: 774; Brough 1962: 72f.). This interpretation is not universally accepted, however (cf. Salomon 1998b: 75; 2008: 123; Cheung 2013: 624; Schoubben 2022a: 150 fn. 12).

𑀔 *ṭha*

In inherited words, the akṣara 𑀔 *ṭha* typically represents the outcome of OIA -*ṣṭ(h)-*, e.g. *naṭha* ‘destroyed’ < OIA *naṣṭa-* or *pratiṭhaveti* ‘(s)he establishes’ < OIA *pratiṣṭhāpayati*.⁵³ In contrast to the corresponding Brāhmī sign, *ṭha* probably does not indicate a voiceless aspirated retroflex /*ṭʰ*/, but the cluster /*ṣṭʰ(h)*/. This is suggested by foreign renderings of Gāndhārī words containing *ṭha*, such as Gr. *Τιαστάνης*; TB *Caṣtane*; Uyghur *čšt’ny* ← Gandh. *Caṭhana* (PN) or TA *ṣoṣṭaṅk** (nom.pl. *ṣoṣṭāṅkāñ*) ← Niya *ṣoṭhaṅga* ‘finance officer’. Corroborating evidence for the value /*ṣṭʰ(h)*/ comes from occasional spelling variations as between *suveṭha* and *suveṣṭa* (a title) and from the preservation

⁵¹ In contrast to Bajaur fragment 5 (CKM 268), there is no Niya example of 𑀔 *ka* standing for /*skʰ*/.

⁵² Burrow still transliterated these signs as *čḥa* and *chā*, respectively.

⁵³ OIA -*ḥsth-* is another source, cf. *paraṭhida* ‘standing in front’ < *paraḥ-sthita-*.

of the relevant cluster in modern Dardic languages, as in Khowar *oṣṭ* < OIA *aṣṭā* ‘eight’. For more details, see Bailey (1949a: 123–125; 1950a: 398); Brough (1962: 75–77); Baums (2009: 164).

१ *ḥha*

१ *ḥha* is more or less consistently distinguished from ७ *ṭha* /ṣṭ^(h)/, which implies that their pronunciations must have been sufficiently different to have called for the creation of separate graphemes. Yet, the phonological value of *ḥha* is contended (see e.g. Konow 1943: 68–72; Brough 1962: 75–77; Baums 2009: 164–167). For reasons of space, I cannot delve very deeply into the debate here, so I shall limit myself to the Niya evidence. In Niya Prakrit, only two words containing this akṣara have a secure etymology, i.e. *aḥhovaga* ‘fit for a purpose, suitable’ < OIA **arthopāka*- and *vaḥhayaga* ‘attendant’ < OIA *upasthāyaka*-.⁵⁴ In *aḥhovaga*, *-ḥh-* corresponds to OIA *-rth-*. Its phonological value seems to be /ṭṭ^h/, as unaspirated *-rt-* yields Niya *-ṭ-* /ṭṭ/ (e.g. *kaṭavo* ‘to be done’ < *kartavyah*). In *vaḥhayaga*, *-ḥh-* is the reflex of OIA *-sth-*.⁵⁵ The Pāli cognate of this word is *upaṭṭhāka*- < *upasthāka*-, which suggests that the *-ḥh-* in Niya *vaḥhayaga* could likewise stand for /ṭṭ^h/ . This view receives further support from the variant spelling *vaṭayaga*,⁵⁶ whose *-ṭ-* /ṭṭ/ could result from deaspiration of /ṭṭ^h/ (cf. §2.1.3).

५ *ḍa*

In older literature (e.g. LKD), ५ was transcribed as *ḍa*. Because this akṣara is distinguished from ५ *ḍa* by an extra line at the bottom right, *ḍa* is these days the preferred transliteration. Except for occurrences in loanwords, ५ *ḍa* ordinarily represents the outcome of a lenited intervocalic *-ṭ-* or *-ḍ-*, as in *kukuḍa* ‘cock’ < OIA *kukkuṭa*- or *vaḍavi* ‘mare’ < OIA *vaḍavī*-. ५ *ḍa* probably stood for a rhotic sound, here phonologically transcribed as /r/ (cf. LKD: §18

⁵⁴ The meaning and etymology of *kuḥhakṣira* are obscure (see §3.7.2 s.v.); the same is true of the adj. *raḥhi* in CKD 574; and the PN *Kruḥhaka* in CKD 513 is etymologically unclear as well. In CKD 422 (perhaps also 852), there is a variant spelling *uḥha* for the more common *uṭa* ‘camel’, traditionally assumed to be from OIA *uṣṭra*- but perhaps rather a loanword (see fn. 940).

⁵⁵ The outcome of OIA *-sth-* can alternatively be written with † *tha*, as in *thavitaga* ‘deposit’ < **sthāpitaka*-.

⁵⁶ Attestations: dir.sg. *vaḥhayaga* in CKD 419; 576 (*rayavaḥhayaga* ‘royal attendant’); 579 (2x); 581; 782 (2x); gen.pl. *vaḥhayagana* in CKD 357; 637; 796.

and §2.1.2). One can deduce this pronunciation from Gāndhārī loanwords into, for instance, Tocharian, given that speakers of this language substituted *-ḡ-* with *-r-*, e.g. in TB *kor* ‘myriad’ ← Gandh. *koḍi* /kōḍī/ < OIA *koḍī-* (DoT: 215). Another indication of the value /r/ is the spelling variation between loc.sg. *śarataṃmi* and *śaḍataṃmi* ‘in the autumn’ in CKD 574.⁵⁷ Phonetically, 𑖅 *ḍa* may have been a retroflex flap [ɽ] as in NIA languages, including Dardic ones (see von Hinüber 2001²: §201; Baums 2009: 141; cf. also §4.3.26).

𑖅 *na*

In Niya Kharoṣṭhī, there was a strong preference for writing dental 𑖅 *na* everywhere, also in places where retroflex 𑖅 *ṇa* would be expected etymologically (KD: 305).⁵⁸ In other words, the two types of nasal had probably merged in Niya Prakrit (LKD: §34), as they had in some other forms of Gāndhārī (cf. Brough 1962: 97f.; Baums 2009: 134–136; 147f.). Because in Niya Prakrit the outcome of this merger is by default written as 𑖅 *na*, I phonologically represent it as /n/.⁵⁹

𑖅 *lyā*

The sign above occurs mainly in borrowings, but it can occasionally be found in inherited vocabulary too. By Boyer et al. (KD), it had been transcribed as *lpa*, but Burrow (1935b: 670) and Lüders (1936b: 637ff.) changed the reading of this akṣara to *lyā*. They did so because *lyā* and *lya* (written 𑖅) interchange in the case of the PN *Kalyānadharma* / *Kalyanadharma* and because *lyī* and *li* alternate in examples such as *lyihida* ‘written’ next to *lihida*, and *Lipe* (a foreign PN) next to *Lýipe(ya)*.⁶⁰ These interchanges suggest that *lyā* indicates a palatal /l̥/. See further §6.3.6.

⁵⁷ Baums (2009: 141; cf. also Emeneau 1931) further mentions the prefix variation *paḍi-* and *pari-* in South Asian Gāndhārī; for a Niya example of this variation, cf. *parichitama* in CKD 249, which Burrow (1940: 46) interprets as a variant spelling of *paḍichitama* ‘we received’.

⁵⁸ The akṣara 𑖅 *ṇa* is hardly used in Niya Kharoṣṭhī.

⁵⁹ The exact phonetic value of /n/ is difficult to determine, as it could have been a dental [n], an alveolar [n̪], or a retroflex [ɳ].

⁶⁰ *Kalyānadharma* is attested in CKD 536; 605; *Kalyanadharma* in CKD 123; 477; 560; 597; 601; 611 (2x); 618; 619; 762; 763; *lyihida* in CKD 575; and *Lipe* in CKD 754.

𑀮 va

The akṣara 𑀮 *va* (by Burrow transliterated as *va*) is distinguished from 7 *va* by the addition of a vertical line at the bottom right, the so-called *cauda* (Glass 2000: 98). Outside the Niya documents, 𑀮 *va* has hitherto only been discovered on some Kuṣāṇa-period coins. In Niya Prakrit, 𑀮 *va* commonly occurs in foreign personal names, such as *Varpa*. It can additionally be used as a hiatus filler after back vowels like *u*, e.g. *hetuvena* ‘by reason’ < *hetu-*. Therefore, 𑀮 *va* most likely represents a semivowel /w/, sufficiently distinct from /v/ to have warranted the creation of an additional grapheme. Nonetheless, scribes often confused 𑀮 *va* and 7 *va*; they did so in foreign words, such as *vasu* next to *vasu* (a title), but also in inherited words, such as *viṃṇaveti* beside *viṃṇaveti* ‘(s)he orders’ < OIA *vijñāpayati*. See also LKD (§29).

𑀯 vha

Boyer et al. (KD) and Burrow (LKD; TKD) transliterated the Kharoṣṭhī sign 𑀯 as *p̄ha*. Nowadays, Lüders’ (1909: 657f.) transliteration *vha* is preferred for the following two reasons: (i) “the graphic form [of *vha*] appears to consist of a *va* with a diacritic ... attached to the right of the vertical” (Brough 1962: 66; cf. also Glass 2000: 118); (ii) *vha* can represent the outcome of an OIA *-bh-* that underwent intervocalic lenition, as in Khotan Pkt. *avhiṇanu* ‘recognition mark’ < OIA *abhijñāna-* (CKD 661). As argued in detail by Baums (2009: 145–147), the initially intended phonological value of *vha* was probably /v^h/ (perhaps ~ IPA [v^h]). In late forms of Gāndhārī, including Niya Prakrit, a subsequent lenition to *-h-* had happened, as in *prahuḍa* ‘gift’ < OIA *prābhṛta-* (cf. LKD: §27). Consequently, there are no inherited words in Niya Prakrit containing *vha*: this akṣara only occurs as a substitution of Ir. **-f-*, e.g. *Tiravhara* (PN) ← Bactr. **τ(ε)ιροφαρρο* /t̪irəfarrə/ < **tira-farnah-* ‘glory of the god Tir’. The same usage of *vha* occurs in other forms of Gāndhārī, and we may presume that Ir. **-f-* was approximated there as /v^h/. Since Niya Prakrit no longer had a phoneme /v^h/ in inherited words (cf. supra), I cautiously assume that in this late form of Gāndhārī *vha* had become a way to write a loan phoneme /f/.

𑀓 śpa and 𑀔 s̥pa

The Kharoṣṭhī signs 𑀓 śpa and 𑀔 s̥pa have so far not been discovered outside of the Niya documents. The element transcribed as -ś- “consists of a loop at the base of the character and crossing the stem slightly” (Glass 2000: 126). 𑀓 śpa only occurs in the foreign personal names Śpālu (CKD 781) and Śpālyaya / Śpālyiya, the latter having a variant with 𑀕 sva, i.e. Śvalyaya (CKD 709). 𑀔 s̥pa can also be found in inherited vocabulary, e.g. in *paros̥para* ‘one another’ < *paraspara*-. Similar to how 𑀓 śpa can alternate with 𑀕 sva, 𑀔 s̥pa interchanges with 𑀖 sva, e.g. *s̥parna* ‘gold’ next to *svarna* (CKD 113). It seems therefore likely that the element -ś- indicates a fricative close to /v/, possibly /f/. See also LKD (§49).

2.1.7 The transliteration of Kharoṣṭhī 𑀣 as *jha*

The transliteration of the Kharoṣṭhī script adopted in this book is the same as that used in the online Gāndhārī dictionary and text catalogue—with the exception of one essential point, i.e. the transliteration of the akṣara 𑀣. Whereas Baums and Glass (2002ab–) transliterate 𑀣 as *za*, I use the older transliteration *jha*.

The most extensive argumentation in favour of the transliteration *za* that I know of is by Brough (1962: 60–62); he advanced the following three arguments (see also Glass 2000: 67; 110–112):

- (i) 𑀣 is used to render a foreign *z* in loanwords from Iranian and Greek.
- (ii) When other MIA languages have *-(j)jh-*, the corresponding sound in Gāndhārī is typically written as *Y ja* or *Ȳ ja*, and not as 𑀣.⁶¹ One such example is *uvaṅjae* (CKI 46), also gen.sg. *upajayasa* (CKI 161), < *upādhyāya*- ‘teacher’ (cf. Pāli *upajjhāya*-).
- (iii) It is potentially confusing to use the transliteration *jha* when 𑀣 is also used for the outcome of intervocalic OIA *-dh-* and interchangeable with 𑀇 *sa* and 𑀈 *śa* /*z*/, as e.g. in *asimatra* / *aśimatra* / *ajhimatra* (or *azimatra*) ‘above measure’ < *adhimātram*.

This section hopes to show that none of these arguments is compelling. Although it is undeniable that 𑀣 can also stand for /*z*/, it seems that spelling variations, chronological distributions, and arguments from palaeography

⁶¹ Originally, *Ȳ ja* may have been used to write a geminate /*jj*/ (cf. Baums 2009: 197–200), but *Ȳ ja* and *Y ja* have become interchangeable in the extant texts.

bolster the conclusion that the originally intended value of ʃ was /j^(h)/. For this reason, I prefer to transliterate ʃ as *jha* and not as *za*.

2.1.7.1 The substitution of foreign z

Brough's first argument needs the assumption that ʃ has always been a loan phoneme that exactly replicates a foreign z. Yet, if there is no exactly corresponding phoneme in the language, it is cross-linguistically more common to adapt foreign sounds to the phonological system of the recipient language, meaning that the foreign sound is approximated, not exactly replicated (cf. also §4.3).⁶² This process of phonological adaptation is also evidenced in Gāndhārī, as ʃ is not the only way to approximate z in loanwords (see further §4.3.19).

Most importantly, there are examples of z being substituted with ʃ *ja* in both Niya Prakrit and South Asian Gāndhārī: a Niya example is *jenavida* 'chief of the armour' ← Bactr. ζηνοβιδο /zēnəvidə/ < **zaina-pati-* (§3.2.2 s.v.);⁶³ an instance from South Asian Gāndhārī is the PN *Damijada* (CKI 42, gen.sg. *Damijadasa*) ← **dāmisād(ə)* < OIr. **dāmi-zāta-* 'born through (the will of) the creator'.⁶⁴ Otherwise, Iranian names containing *-zād(ə) < OIr. *-zāta- are usually rendered with ʃ (cf. e.g. Salomon 1997: 233).⁶⁵

⁶² /z/ was only added later to the phonemic inventory of Gāndhārī due to intervocalic lenition of OIA *-th-*, *-dh-*, and *-s-* (cf. §2.1.2). To the best of my knowledge, the earliest evidence for ʃ used to substitute a foreign z are a few coins of the middle of the 2nd century BC, on which the name of the Indo-Greek king Zoilus I (Ζωϊλος) is rendered in Kharoṣṭhī as *Jhoīla* (gen.sg. *Jhoīlasa* in CKC 37; 38; 39).

⁶³ Contrast *jheniga* 'in charge of, under the care of' ← Bactr. *ζηνιγο (attested ζινιγο) /zēnīgə/ < **zainiya-kā-*.

⁶⁴ For the idea that *-jada* is to be identified as Ir. *-zāta-, see already Konow (KI: 14), who points to variants of the same name written in Brāhmī script as *Dāmajada* and *Dāmayasada*. The reconstruction of the compound as **dāmi-zāta-* can be found in Harmatta (1989: 301), but only the etymology of the second part of *Damijada* is relevant for the argument about ʃ.

⁶⁵ One such example is the PN *Mahajhada* (*Mahazada*) (CKI 102; 327; 332), in my view a rendering of Bactr. *μασοζαδο /māhəzādə/ < **māha-zāta-* 'born through (the will of) the Moon (god)'. Another instance is the PN *Arajha(ṃ)da* (CKI 108; 564), which I consider a rendering of Bactr. *αυροζαδο /ahrəzādə/ 'born through (the will of) Ahura' < **ahura-zāta-* (differently Harmatta 1989: 301; Falk 2015: 10).

Also in Sanskrit, *j* can be employed as an approximation of a foreign *z*.⁶⁶ The word *jagara*- ‘armour’, for instance, is commonly assumed to have been borrowed from a cognate of Khwar. *zγryk*; Oss. I. *zγær*, D. *æsqær*; and Psht. *zγæra* ‘idem’ (Abaev 1958–89 IV: 308f.; EWAia III: 205; EDP: 101). A second instance is *jena*- / *jayana*- ‘saddle, armour (for a horse)’, which has been identified as a loanword from Mir. **zēn(ə)* < **zainu*- (EWAia III: 212). The example of *gañja*- ‘treasury’ ← OP **ganza*- reveals that the same rendering was also possible after a nasal (see also §3.5.1 s.v. *gaṃṇā*).

The fact that a foreign *z* could be substituted with *j* in Gāndhārī and other Indo-Aryan languages shows that the use of ʃ as a substitute for a foreign *z* does not necessarily imply that ʃ represents a phoneme /z/ rather than /j^h/. As Burrow (1937: 112) puts it, “the Indians themselves in those days as well as now used their own *j* or *jh*, or modifications, to express the sound [i.e. *z*] which was foreign to them”.⁶⁷ Consequently, Brough’s first argument in favour of transliterating ʃ as *za* is not compelling.

2.1.7.2 The Gāndhārī equivalent(s) of MIA *-(j)jh-*

As a second argument in support of ʃ being *za*, Brough observed that the Gāndhārī equivalent of *-(j)jh-* in other MIA languages is spelled with ʃ *ja* or Ȳ *ĵa*, not with ʃ. Brough’s observation is, however, no longer fully accurate: at least one example of ʃ corresponding to MIA *jh-* has meanwhile come to light. The Gāndhārī cognate of MIA *jhāna*- ‘meditation’ < OIA *dhyāna*- is usually spelled with ʃ *ja* or Ȳ *ĵa*, and in Senior fragment 12 (CKM 244) also with ʃ *ca* (cf. Silverlock 2015: 166). Yet, in the Senavarma inscription (CKI 249), there is one instance of the same word spelled with ʃ, i.e. *jhāṇa* (*zaṇa*).⁶⁸ How should we interpret this variant form?

⁶⁶ In what follows, I exemplify this substitution with borrowings from Iranian, but the same rendering is found with Greek loanwords. Note e.g. Skt. *jūka*- ‘the constellation Libra’ ← ζυγόν / ζυγός ‘idem’ (EWAia III: 211). This astrological term occurs in Varāhamihira’s *Bṛhajjātaka*, a text datable to the 6th century AD, i.e. when Gr. ζ no longer stood for /zd/ or /dz/, but for /z/ (VG³: 58).

⁶⁷ Speakers of Indian languages copying Avestan manuscripts also frequently confused *-j-*, *-z-*, and *-ž-*.

⁶⁸ In its other instance in CKI 249, the word is written as *ĵaṇa*. There is another sequence *jhana* (*zana*) in CKI 841, but the context makes it unlikely that the word for ‘meditation’ is intended there (Scherrer-Schaub, Salomon & Baums 2012: 165 with fn. 25).

Preferring the reading *zaṇa*, Glass (2000: 110) has argued that a sound change from *dhy-* > *jh-* > *z-* operated in this word. As support, he cited examples of intervocalic *-dhy-* becoming *-z-* in modern Dardic languages.⁶⁹ Glass' argument has, however, been cogently refuted by Baums (2009: 171f.), who noted that all palatal sounds became *-z-* in Dardic, meaning that the Gāndhārī outcome of OIA *-dhy-* could still have been palatal.

Baums (2009: 171f.) approached the problem posed by *jhaṇa* (*zaṇa*) from a different angle. Seeing that the Gāndhārī outcome of OIA *-dhy-* is typically written with Υja or $\bar{\Upsilon} ja$, Baums searched for other evidence that could tell us the value of Υja and $\bar{\Upsilon} ja$. He mentions in this respect the Gāndhārī outcome of OIA *dahyamāna-* 'burning' (cf. Pāli *dayhamāna-*; *ḍayhamāna-*). This word is currently attested in three different spellings:

- (i) *dajamaṇa* (NIRD^{L2} recto 79);
- (ii) *ḍajamaṇa* (DHP^K 75; 159 ins.sg. *ḍajamaṇeṇa*);
- (iii) *daśamaṇa* (NIRD^{L2} recto 76; DHP^{SP} 37a; 37b ins.sg. *daśamaṇeṇa*).

The latter spelling illustrates the default change of OIA *-hy-* to Gandh. /ʒʒ/, here written with $\sqcap śa$, but elsewhere also with Υja . On this basis, Baums concluded that the Υja and $\bar{\Upsilon} ja$ in *dajamaṇa* / *ḍajamaṇa* also stand for /ʒʒ/.

The same conclusion was earlier reached by Brough (1962: 62), who, however, added that "this does not exclude the possibility that in initial position and where representing an older double stop, the sign might still have represented a palatal plosive or affricate." Indeed, it seems suspicious that, apart from this one example cited by Baums, Υja and $\bar{\Upsilon} ja$ from OIA *-dhy-* never interchange with $\sqcap śa$ or Υja .⁷⁰ Note, for instance, Niya *parib(h)uj-* 'to understand' < *paribudhy-*, which in its more than 70 attestations is never written with $\sqcap śa$ or Υja , suggesting that the Υja in *parib(h)uj-* stands for a geminate /jj/, deaspirated from earlier /jj^h/ (cf. infra). Because initial stops are generally resistant to lenition in Gāndhārī (as

⁶⁹ Cf. e.g. Pashai *waraz-* 'to protect, keep, feed' < OIA **uparādhyati* 'to favour' (CDIAL: no. 2217).

⁷⁰ Moreover, the outcomes of *dahyamāna-* listed above all come from the same verse, attested with minor variants in the *Khotan Dharmapada*, the *Split Dharmapada*, and the *Nirdeśa* commentary contained in British Library fragment 9. The outcome of *-hy-* is always written with $\sqcap śa$ in the *Khotan Dharmapada* (Brough 1962: 105), and *ḍajamaṇa* would be the only example where $\bar{\Upsilon} ja$ is not used for the outcome of *-dhy-* (cf. Falk 2015a: 57 "treats *hya* as if it was *dhya*").

in MIA in general), it also stands to reason that in initial position $Y ja$ and $\bar{Y} ja$ from OIA *-dhy-* represent /j/ and not a fricative /ʒ/.

The value /j/ is further confirmed by the fact that Gandh. *jaṇa* / *jāṇa* < *dhyāna-* was borrowed into Khotanese as *jāna-* and into Chinese as *chán* 禪 ‘Chán/Zen Buddhism’, in (Early) Middle Chinese *dzyen*: the initial *j-* in Khot. *jāna-* and the initial *dzy-* in (E)MC *dzyen* stand for a voiced affricate (approximately IPA [d͡z]) and thus corroborate that Gandh. *jaṇa* / *jāṇa* did not begin with a fricative.

The Gāndhārī development of OIA *-dhy-* to /j/ implies an intermediary stage when aspiration was still preserved, i.e. */jʰ/ (see von Hinüber 1982: 248 = 2009–19: 308).⁷¹ Hence, one can simply assume that *jaṇa* is a historical spelling that preserves the original aspiration.

Another possibility is that *jaṇa* is a loanword. Because of the high frequency of Buddhist formulae in the Senavarma inscription (cf. von Hinüber 2003), *jaṇa* could, instead of being inherited, also have been borrowed from an eastern MIA form **jhāṇa-*. Also in this case, 𑀘 would stand for /jʰ/, thus warranting the transliteration of this akṣara as *ja*.

A second word where 𑀘 seems to correspond to MIA *-(j)jh-* is *gujha* (*guza*) (SĀ^{S1} recto 3). Because this word occurs in a list of body parts, Glass (2007: 124; 159–161) interpreted it as a Gāndhārī cognate of OIA *guhya-* in its derived meaning ‘pudenda, anus’. This etymology is accepted by Baums (2009: 171) and taken as evidence that 𑀘 could also stand for /ʒʒ/, the default outcome of OIA *-hy-* in Gāndhārī (cf. supra). In my view, there is no need to do so. Since it occurs in a canonical *sūtra* text based on a now lost original in an eastern MIA language, *gujha* / *guza* could instead be a loan from eastern MIA **gujjha-* < *guhya-* (CDIAL: no. 4222).⁷² In this case, *gujha* would be the preferred transliteration, and one does not have to take recourse to complicated phonological arguments: *gujha* could simply write /gujjʰa/.

In brief, there is nowadays at least some evidence that Kharoṣṭhī 𑀘 can correspond to *(j)jh* in other MIA languages. As a result, Brough’s second argument in favour of the transliteration *za* can no longer be upheld.

⁷¹ It is probably due to this deaspiration that 𑀘 is absent from the Aśokan Gāndhārī inscriptions. The fact that 𑀘 *ja* was scarcely used in inherited vocabulary may also explain why this akṣara became the standard way to render a foreign *z* in loanwords from Greek and Iranian.

⁷² For the phonological development, cf. e.g. Ved. dat.sg. *máhya(m)* ‘I’ > Pāli *mayha(m)* > Pkt. *majjha(m)* (von Hinüber 2001²: §245; §366).

2.1.7.3 Akṣaras alternating with 𑀘 and changing phonological values

This brings us finally to Brough's third argument, i.e. the potential confusion that can result from transliterating 𑀘 as *pha* when this grapheme is also used for the outcome of OIA intervocalic *-dh-* and more generally interchanges with 𑀚 *sa* and 𑀛 *sa /z/*. Even though I agree that this type of alternation would be less confusing when transliterating 𑀘 as *za*, two observations nonetheless weaken the force of this argument significantly.

First, it should be noted that 𑀘 can also alternate with the palatal obstruents 𑀜 *ca*, 𑀝 *cha*, and 𑀞 *ja*. So, if one transliterates 𑀘 as *za*, this type of interchange is no less confusing than the one with 𑀚 *sa* and 𑀛 *sa* when adopting the transliteration *pha*.

The alternation between 𑀘 and palatals occurs in both Central Asian and South Asian Gāndhārī. In Niya Prakrit, for instance, *phorita* is attested as a variant of *c(h)or-* 'to steal' < OIA $\sqrt{c}ur$, while *phuṭhi*, a type of corn, occurs as a variant spelling of *cuṭhi(ye)* and *juṭhi*.⁷³ In both cases, the currently adopted transcriptions *zorita* and *zuṭhi* fail to properly visualise this type of interchange. For the same reason, it seems preferable to transcribe the South Asian Gāndhārī PN gen.sg. *Budhazi[a]sa* < OIA *Buddhajīva-* as *Budhajhi[a]sa* instead (CKI 557).

Still more data come from the *arapacana* poem transmitted as fragment 5 of the Bajaur collection of Gāndhārī manuscripts (CKM 268). This poem, cited from a preliminary transcription by Gudrun Melzer (Munich), is a collection of Buddhist verses arranged according to the *arapacana* sequence of the Kharoṣṭhī script. The two hemistichs of verse 39 start with the words *phariṭa* (*zariṭa*) 'aged' and *phara* (*zara*) 'old age', whose etymologies from respectively *jarita-* and *jarā-* strongly suggest that they ought to be transcribed with an initial *ph-*.

Second, it seems that the interchange between 𑀘 and 𑀚 *sa* and 𑀛 *sa /z/* is a characteristic of Central Asian as opposed to South Asian Gāndhārī, suggesting that it may be a late (or perhaps dialectal) development. I could find only one instance in a South Asian source of 𑀘 being used where one would expect 𑀚 *sa* or 𑀛 *sa*, i.e. in the Kuṣāṇa inscription CKI 149, where the

⁷³ *phorita* is attested in CKD 177; 678; 694; *phuṭhi* in CKD 677; *cuṭhi(ye)* in CKD 422; 782; and *juṭhi* in CKD 291; 571; 703 (2x).

outcome of the OIA loc.sg. *māse* ‘in the month’ is spelled as *majhe* (*maze*) and not as *mase* or *maṣe* as in other inscriptions.⁷⁴

The alternation between ṛ *sa*, 𑀓 *ṣa*, and 𑀣 allows one to conclude that /z/ was a possible phonological value of 𑀣. It does not follow that 𑀣 has always been /z/ or that /z/ was the only value possible. The evidence assembled above supports there to have been an alternative phonological value /j(j)^h/.

The key question is which phonological value is the originally intended one. In my view, the confusion between ṛ *sa*, 𑀓 *ṣa*, and 𑀣 being restricted to late sources counts as an argument that /j(j)^h/ is the originally intended value. Further support for this view is to be found in palaeography. If 𑀣 had always been a sign for /z/, one wonders why a new grapheme 𑀓 *ṣa* was developed to unambiguously write /z/ as the outcome of a lenited intervocalic *-s-*, *-th-*, and *-dh-* (so also Burrow 1935a: 780; LKD: §22).⁷⁵ What is more, the shape of the akṣara 𑀣 potentially reveals that its originally intended value was *jha* and not *za*, as it may not be a coincidence that 𑀣 not only looks like a vertically mirrored image of Kharoṣṭhī 𑀧 *ṇa*, another palatal sound (cf. Baums 2009: 171), but also as a mirror of the Early Brāhmī sign 𑀧 *jha* (Stefan Baums p.c.).⁷⁶ I therefore find it preferable to transcribe Kharoṣṭhī 𑀣 as *jha*.

The subsequent use of Kharoṣṭhī 𑀣 to write /z/ may be conceived of as the result of internal phonological developments in Gāndhārī. The lenition of OIA *-th-*, *-dh-*, and *-s-* had resulted in the addition of /z/ to the phonemic inventory of Gāndhārī—at least in intervocalic position—and the sign 𑀓 *ṣa* was developed to write this new phoneme /z/ in inherited words (cf. supra). 𑀣 *jha*, on the other hand, was still the default way to render a foreign *z* in loanwords. Since, however, intervocalic /z/ was now part of Gāndhārī’s phonemic inventory, foreign *z* could at once be rendered more faithfully as /z/ rather than /j(j)^h/. This new way to render a foreign *z* could, in turn, have caused a change in the phonological value of 𑀣 *jha*, i.e. from /j(j)^h/ to /z/, therewith enabling 𑀣 *jha* to start interchanging with ṛ *sa* and 𑀓 *ṣa* in inherited words (and ṛ *sa* and 𑀓 *ṣa* to be used as a substitution for a foreign /z/, see §4.3.19).

⁷⁴ For the spelling *mase*, see CKI 61; 155; 225; 226; 244; 245; 309; 312; 320; 368; 455; 497; 828. For *maṣe*, see CKI 152; 328; 509, and the Niya documents.

⁷⁵ On the historical development of 𑀓 *ṣa*, see recently Schlosser (2022: 85f.).

⁷⁶ Baums (2009: 172; 196) additionally points out that the position of Kharoṣṭhī 𑀣 in the *arapaçana* sequence would fit with its originally being an aspirated voiced palatal.

Most likely, this change in the phonological value of 𑀧 *jha* was originally restricted to intervocalic position. Yet, in some Niya Prakrit names, 𑀧 *jha* can be seen alternating with 𑀧 *sa* (not 𑀧 *sa*) also initially and before /m/. Note, for instance, *Jhagamoya* (CKD 630) and *Jhagimoya* (CKD 632), otherwise spelled as *Sagamoya*, and names ending in *-rajhma* / *-rasma* ← Ir. **razman-* 'battle'. Such alternations evidence that the value of 𑀧 *jha* had changed to /z/ also in these positions. The fact that initial 𑀧 *jha* can still interchange with 𑀧 *ja* and other palatal obstruents (see above on *jhuthi*, a type of corn and *jhorita* 'stolen') leads one to believe that this process may not have been fully completed. Therefore, I cautiously assume that 𑀧 *jha* still had two phonological values in Niya Prakrit, i.e. /j^(h)/ and /z/.

TABLE 1 Niya Prakrit phonology

This table outlines the most likely phonological value(s) of the akṣaras used in Niya Kharoṣṭhī. For the assumed value(s) of *kā*; *kṣa*; *kṣā*; *jha*; *ṭha*; *ḥha*; *ḍa*; *na*; *lā*; *ya*; *vha*; *ṣā*; *śā*, see §§2.1.6–2.1.7. I do not mention /gg^h/ and /bb/ as possible values of respectively *gha* and *ba* because I did not find a relevant example in Niya Prakrit. These values do occur in other types of Gāndhārī.⁷⁷ As in OIA (Nikolaev 2021), geminate /rr/ was probably absent from the language.

LETTER	VALUE	EXAMPLE(S)
𑀓 a	/a/ /ā/	<i>aja</i> / <i>ajja</i> / ‘now, today’ < <i>adya</i> <i>gamana</i> (<i>gamam</i>) / <i>gamana</i> / ‘going’ < <i>gamana-</i> <i>agach-</i> / <i>āyacch</i> ^h / ‘to come’ < <i>āgaccha-</i> <i>kala</i> (<i>kāla</i>) / <i>kāla</i> / ‘time’ < <i>kāla-</i>
𑀓 i	/i/ /ī/	<i>ich-</i> / <i>icch</i> ^h / ‘to wish’ < <i>iccha-</i> <i>lihidaḡa</i> / <i>lihiḍaya</i> / ‘written’ < * <i>likhitaka-</i> <i>nidaḡa</i> / <i>nīḍaya</i> / ‘taken’ < * <i>nītaka-</i>
𑀓 u	/u/ /ū/	<i>udaga</i> / <i>uḍaya</i> / ‘water’ < <i>udaka-</i> <i>puno</i> / <i>punō</i> / ‘again’ < <i>punar</i> <i>dura</i> / <i>dūra</i> / ‘away, far’ < <i>dūra-</i>
𑀓 r	/ri/ /ru/	<i>kṛta</i> (<i>kṛita</i>) / <i>kṛiḍa</i> / ‘made’ < <i>kṛta-</i> <i>pr̥ch-</i> (<i>pruch-</i>) / <i>prucc</i> ^h / ‘to ask’ < <i>pr̥ccha-</i> (cf. LKD: §5)
𑀓 e	/ē/ /ē/	<i>eka</i> / <i>ēkka</i> / ‘one’ < <i>eka-</i> (cf. Pkt. <i>ekka-</i>) <i>pitupetri</i> / <i>piḍupētri</i> / ‘ancestral’ < <i>pitṛpaitṛka-</i> (cf. Pāli <i>pitupettika-</i>) <i>eṣa</i> / <i>ēṣa</i> / ‘this’ < <i>eṣa-</i> <i>lekha</i> / <i>lēk</i> ^h <i>a</i> / ‘writing’ < <i>lekha-</i>
𑀓 o	/ō/ /ō/	<i>yoga</i> / <i>yōḡga</i> / ‘proper’ < <i>yogya-</i> (cf. Pāli <i>yogga-</i>) <i>omaga</i> / <i>ōmaya</i> / ‘short’ < * <i>avamaka-</i> (cf. Pāli <i>omaka-</i>) <i>kośala</i> / <i>kōśala</i> / ‘well-being’ < <i>kauśala-</i>
𑀓 m	/ṃ/ /ñ/ /ṇ/ /n/ /m/	<i>upaśamk-</i> (<i>upaśamḡ(h)-</i>) / <i>upaśanḡ</i> / ‘to worry’ < <i>upa</i> + <i>√śaṅk</i> <i>paṃca</i> / <i>pañja</i> / ‘five’ < <i>pañca</i> (cf. DHP ^K <i>paja</i>) <i>piṃḍa</i> / <i>piṇḍa</i> / ‘total’ < <i>piṇḍa-</i> <i>mahaṃta</i> (<i>mahaṃda</i>) / <i>mahanda</i> / ‘big’ < <i>mahat-</i> <i>gamam</i> / <i>gaman</i> / ‘going’ < <i>gamana-</i> (cf. LKD: §13) <i>samṃujita</i> / <i>sambūziḍa</i> / ‘honoured’ < <i>samṃujita-</i>

⁷⁷ Cf. *ughuṭhago* ‘proclaimed’ < OIA **udghuṣṭaka-* (Pāli *ugghuṭṭha-*) in Av^{L4} verso 306 and *babaka* ‘made from grass’ < *bālbaja-* (Pāli *babbaja-*) in DHP^K 169.

᳚ ka	/k/ /kk/	kama /kāma/ 'desire' < <i>kāma-</i> ukas- / uk kaz/ 'to go away' < <i>ut + √kas</i> (cf. Pāli <i>ukkas-</i>) eka /ē kk a/ 'one' < <i>eka-</i> (cf. Pkt. <i>ekka-</i>)
᳚ ka	/sk/ /šk/	nama ka ro /namaskārō/ 'homage' < <i>namaskāra-</i> ma ka /mā ṣ ka/ 'bean' ← Bactr. *μαῖσκο /māš̄kə/ (§3.2.2 s.v.)
᳚ kṣa ᳚ kṣa	/kṣ/ /kṣa/	kṣetra (kṣetra) /kṣētra/ 'field' < <i>kṣetra-</i> sakṣi (sakṣi) /sākṣi/ 'witness' < <i>sākṣin-</i>
᳚ kha	/k ^h / /kk ^h /	khara /k ^h ara/ 'ass' < <i>khara-</i> nikhal- /nikk ^h āl/ 'to remove, provide' < <i>niškālaya-</i> dukha /dukk ^h a/ 'sorrow' < <i>duḥkha-</i> (cf. Pāli <i>dukkha-</i>)
᳚ ga	/g/ /gg/	goṭha /gōṣṭ ^(h) a/ 'farm, household' < <i>goṣṭha-</i> yoga /yō gg a/ 'proper' < <i>yogya-</i> (cf. Pāli <i>yogga-</i>)
᳚ ḡa	/gn/	viḡa /vigna/ 'obstacle' < <i>vighna-</i>
᳚ ga	/γ/	anega /anēya/ 'many' < <i>aneka-</i> aroga /arōya/ 'health' < <i>aroga-</i>
᳚ ḡa	/γn/	Suḡuta /sū γ nūḍā/, foreign PN ~ <i>Sugnuta</i>
᳚ gha	/g ^h /	ghrida /g ^h riḍa/ 'ghee' < <i>ghṛta-</i> śighra /śiḡ ^h ra/ 'quickly' < <i>śiḡhram</i>
᳚ ca	/c/ /cc/	cora /cōra/ 'thief' < <i>caura-</i> kica /kicca/ 'task' < <i>kṛtya-</i> (cf. Pāli <i>kicca-</i>)
᳚ cā	/śc/	pa cā /paścā/ 'afterwards' < <i>paścāt</i>
᳚ cha	/c ^h / /cc ^h /	chagalaga /c ^h aḡalaya/ 'she-goat' < * <i>chagalaka-</i> pruch- /prucc ^h / 'to ask' < <i>pr̥ccha-</i> (cf. Pāli <i>pucch-</i>)
᳚ ja	/j/ /jj/	jeṭha /jēṣṭ ^(h) a/ 'eldest' < <i>jyeṣṭha-</i> raja /rājja/ 'kingdom' < <i>rājya-</i> (cf. Pāli <i>rajja-</i>)
᳚ ja	/ž/ /žž/	pradeja (<i>pradeśa</i>) /praḍēža/ 'district' < <i>pradeśa-</i> yajitaga /yāžiḍaya/ 'requested' < * <i>yācitaka-</i> guja /gužža/ 'hidden' < <i>guhya-</i> (cf. Pāli <i>guyha-</i> , Pkt. <i>gujha-</i>)
᳚ jha	/j ^(h) / /z/	jhorita (<i>c(h)or-</i>) /j ^(h) ōriḍa/ 'stolen' < <i>corita-</i> divajha (<i>divasa</i> ; <i>divaṣa</i>) /divaza/ 'day' < <i>divasa-</i>
᳚ ṅa	/ñ/ /ññ/	ṅana /ṅāna/ 'knowledge' < <i>jñāna-</i> (cf. Pāli <i>ñāna-</i>) a(m)ṅa /aṅṅa/ 'other' < <i>anya-</i> (cf. Pāli <i>añña-</i>)
᳚ ṭa	/ṭṭ/	paṭa /paṭṭa/ 'silk roll' < <i>paṭṭa-</i>
᳚ ṭha	/ṣṭ ^(h) /	naṭha /naṣṭ ^(h) a/ 'destroyed' < <i>naṣṭa-</i>
᳚ ṭha	/ṭṭ ^h /	vaṭhayaga (<i>vaṭayaga</i>) /vaṭṭ ^h āyaya/ 'attendant' < <i>upasthāyaka-</i> aṭhovaga /aṭṭ ^h ōvāya/ 'fit for a purpose, suitable' < * <i>arthopāka-</i>

ᳵ ḍa	/ḍ/	<i>daṃḍa</i> / <i>daṃḍa</i> / 'stick, punishment' < <i>daṃḍa-</i>
ᳶ ḍa	/ṛ/	<i>kukuḍa</i> / <i>kukkuṛa</i> / 'cock' < <i>kukkuṭa-</i> <i>vaḍavi</i> / <i>vaṛavi</i> / 'mare' < <i>vaḍavi-</i> <i>giḍa</i> / <i>giṛa</i> / 'taken' < <i>gṛhīta-</i>
᳷ ḍha	/ḍ ^(h) /	<i>muṣḍhaṣi</i> / <i>muṣḍ^(h)aṣi</i> /, woman exchanged within exogamous marriage practices ← Bactr. *μοζδορι / <i>muždōšī</i> / * <i>mižda-strī-</i> 'reward woman' or * <i>miždušī-</i> 'gnädige (Frau)' (§3.2.3 s.v.)
	/ṛ ^h /	<i>oṣaḍha</i> / <i>oṣaṛ^ha</i> / 'medicinal herb' < <i>auṣadha-</i>
	/ḍḍ ^h /	<i>aḍha</i> / <i>aḍḍ^ha</i> / 'half' < <i>ardha-</i> (cf. Pāli <i>aḍḍha-</i>)
᳸ ta	/t/	<i>taha</i> / <i>tahā</i> / 'thus' < <i>tathā</i>
	/tt/	<i>utara</i> (<i>utvara</i>) / <i>uttara</i> / 'higher' < <i>uttara-</i>
	/ḍ/	<i>lihitaḡa</i> / <i>lihiḍaya</i> / 'written' < * <i>likhitaka-</i>
᳹ tha	/t ^h /	<i>thavitaḡa</i> / <i>t^hāviḍaya</i> / 'deposit' < * <i>sthāpitaka-</i>
	/tt ^h /	<i>iṃthu</i> / <i>itt^hu</i> / 'thus' < <i>ittham</i> (Schlosser 2022: 186; 295)
ᳺ da	/d/	<i>doṣa</i> / <i>dōṣa</i> / 'fault' < <i>doṣa-</i>
	/dd/	<i>udiṣa</i> / <i>uddiśśa</i> / 'concerning' < <i>uddiśya</i> (cf. Pāli <i>uddissa</i>)
	/ḍ/	<i>lihiḍaḡa</i> / <i>lihiḍaya</i> / 'written' < * <i>likhitaka-</i>
᳻ dha	/d ^h /	<i>dhanu</i> 'bow' / <i>d^hanu</i> / < <i>dhanu(s)-</i>
	/dd ^h /	<i>nadha</i> / <i>nadd^ha</i> / 'bundle' < <i>naddha-</i>
᳼ na	/n/	<i>nama</i> / <i>nāma</i> / 'name' < <i>nāman-</i> <i>manuśa</i> / <i>manuśśa</i> / 'man' < <i>manuśya-</i> (cf. Pāli <i>manussa-</i>)
	/nn/	<i>a(m)na</i> / <i>anna</i> / 'food, cereal' < <i>anna-</i>
᳽ pa	/p/	<i>pita</i> / <i>piḍā</i> / 'father' < nom.sg. <i>pitā</i> ← * <i>pitṛ-</i>
	/pp/	<i>pipali</i> / <i>pippali</i> / 'pepper' < <i>pippalī-</i>
᳾ pha	/p ^h /	<i>phalophala</i> / <i>p^halōp^hala</i> / 'all kinds of fruit' < <i>phalāphala-</i>
	/pp ^h /	<i>pupha</i> / <i>pupp^ha</i> / 'flower' < <i>puṣpa-</i> (cf. Pāli <i>puppha-</i>)
᳿ ba	/b/	<i>budha</i> / <i>budd^ha</i> / 'Buddha' < <i>buddha-</i>
ᳺ bha	/b ^h /	<i>bhata</i> / <i>b^hatta</i> / 'food' < <i>bhakta-</i>
	/bb ^h /	<i>abhomata</i> / <i>abb^hōmaḍa</i> / 'disrespected' < <i>abhyavamata-</i> (cf. Pāli <i>abbhomata-</i>)
᳻ ma	/m/	<i>muli</i> / <i>mūli</i> / 'price' < <i>mūlya-</i>
	/ram/	<i>ram-</i> / <i>ram</i> / 'to enjoy' < \sqrt{ram}
	/mm/	<i>samuha</i> / <i>sammuha</i> / 'face to face' < <i>sammukham</i>
᳼ ya	/y/	<i>yadi</i> / <i>yaḍi</i> / 'if' < <i>yadi</i>
	/priya/	<i>priya</i> / <i>priya</i> / 'dear' < <i>priya-</i>
	/yy/	<i>payati</i> / <i>payyātti</i> / 'capacity' < <i>paryāpti-</i>

𑀓𑀲 <i>ra</i>	/r/	<i>raya</i> /rāyā/ 'king' < nom.sg. <i>rājā</i> ← <i>rājan-</i> <i>karana</i> /kārana/ 'reason' < <i>kāraṇa-</i> <i>parikraya</i> /parikraya/ 'rent' < <i>parikraya-</i>
𑀓𑀸 <i>la</i>	/l/	<i>lih-</i> /lih/ 'to write' < √ <i>likhati</i>
	/ll/	<i>kila</i> /kīla/ 'wedge' < <i>kīla-</i> <i>vulasi</i> /vullāzi/ 'admiration' < <i>ullāsa-</i>
𑀓𑀸𑀲 <i>lyā</i>	/l̥/	<i>lyihida</i> /lihiḍa/ 'written' < <i>likhita-</i> <i>Lyipe</i> /Īip(p)ē/, foreign PN
𑀓𑀶 <i>va</i>	/v/	<i>vivada</i> /vivāḍa/ 'dispute' < <i>vivāda-</i> <i>avi</i> /avi/ 'also, too' < <i>api</i>
	/vv/	<i>sava</i> /savva/ 'all' < <i>sarva-</i> (cf. Pāli <i>savva-</i>)
𑀓𑀺 <i>ya</i>	/w/	<i>yaṃti</i> (<i>vaṃti</i>) /wāṇdi/ 'in, at, to' < <i>upānte</i> <i>hetuṇena</i> /hēḍuwēna/ 'by reason' < <i>hetu-</i>
𑀓𑀺𑀲 <i>vha</i>	/f/	<i>Tiravhara</i> /tirafara/, PN ← Bactr. *τ(ε)ιροφαρρο /tirəfarrə/ < *tīra-farnah- 'glory of the god Tir'
𑀓𑀻 <i>śa</i>	/ś/	<i>śata</i> /śaḍa/ 'hundred' < <i>śatam</i>
	/śś/	<i>avaśa</i> /avaśśa/ 'certainly' < <i>avaśyam</i> (cf. Pāli <i>avassaṃ</i>)
	/ž/	<i>daśa</i> (<i>daja</i>) /daža/ 'ten' < <i>daśa</i>
𑀓𑀻𑀲 <i>ṣa</i>	/ṣ/	<i>ṣamanera</i> /ṣāmanēra/ 'novice' < <i>śrāmaṇera-</i>
	/ṣṣ/	<i>vaṣa</i> /vaṣṣa/ 'year' < <i>varṣa-</i> (cf. Pāli <i>vassa-</i>)
	/z/	<i>śeṣa</i> /śēza/ 'remainder' < <i>śeṣa-</i>
𑀓𑀻𑀲𑀲 <i>ṣā</i>	/ṣn/	<i>kriṣāga</i> /kriṣnaya/ 'blackish' < <i>kṛṣṇaka-</i>
𑀓𑀻𑀲𑀸 <i>ṣpā</i>	/ṣf/	<i>ṣpālyaya</i> /ṣfālāya/, foreign PN ~ <i>ṣvalāyaya</i>
𑀓𑀻𑀲𑀻 <i>sa</i>	/s/	<i>saṃvatsara</i> /samvatsara/ 'year' < <i>saṃvatsara-</i>
	/ss/	<i>osuka</i> /ōssukka/ 'eagerness' < <i>autsukya-</i>
	/z/	<i>dasa</i> (<i>daṣa</i> ; <i>dajha</i>) /dāza/ 'slave' < <i>dāsa-</i>
𑀓𑀻𑀲𑀺 <i>śa</i>	/z/	<i>śaca</i> /zacā/ 'with' < <i>sacā</i>
		<i>viśaj-</i> /vizajj/ 'to send' < <i>visarjaya-</i>
𑀓𑀻𑀲𑀻𑀲 <i>śā</i>	/sn/	<i>śatra</i> /snātra/ 'bath' < <i>snātra-</i> ins.sg. <i>pritiśehena</i> /prīḍisnehena/ 'with love and affection' < <i>prīti-</i> + <i>sneha-</i>
𑀓𑀻𑀲𑀻𑀲𑀸 <i>śpā</i>	/sf/	<i>śpāna</i> /sfarna/ 'gold' < <i>suvarṇa-</i> <i>parospara</i> /parōsfara/ 'one another' < <i>paraspara-</i>
𑀓𑀻𑀲𑀻𑀲𑀻 <i>ha</i>	/h/	<i>hasta</i> /hasta/ 'hand' < <i>hasta-</i> <i>suha</i> /suha/ 'happiness' < <i>sukha-</i>
	/hn/	opt.3sg. <i>gr̥hēyāti</i> /grihnēyāḍi/ '(s)he may take' < √ <i>grah</i>

2.2 Graeco-Bactrian script and Bactrian phonology

2.2.1 The Graeco-Bactrian script and its characteristics

Except for one fragment in an adapted form of the Manichaean script (M 1224), Bactrian is written in the so-called Graeco-Bactrian script, a local derivative of the Greek alphabet. The inventory of letters is more or less the same as in Greek, with two notable deviations: the digraph ξ /ks/ only occurs in its respective numerical value 60,⁷⁸ and one grapheme has been added to the script inventory to write a phoneme unknown to Greek phonology, i.e. β /ś/ (on which see Falk 2022: 155–167). This letter is conventionally put at the end of the alphabet.

Palaeographically, one has to distinguish between two variants of the Graeco-Bactrian script. On the one hand, a “monumental” or “rectangular” variant was used for the inscriptions of the Kuṣāṇa period at Surkh Kotal, Rabatak, and other sites. On the other hand, a “cursive” variant was developed to write administrative documents, letters, literary texts, etc. Due to its distinctive ductus, this cursive script had to be gradually deciphered by modern scholars (cf. Sims-Williams 2006; SCBD: 53–59).

The identification of loanwords from Bactrian into Niya Prakrit requires at least a preliminary idea of the phonological value of the Graeco-Bactrian graphemes (summarised in Table 2, p. 62). For some letters, this is sufficiently straightforward. But for others the situation is more complex because there were more phonemes in Bactrian than there were graphemes in the Graeco-Bactrian alphabet. Three topics in particular will need to be discussed at some length: (i) what does the vowel system of Bactrian look like, and how is /h/ graphically rendered (§2.2.2); (ii) when do <β>, <γ>, and <δ> stand for plosives and when for fricatives (§2.2.3); and (iii) which phonemes do <ζ>, <σ>, and <ϐ> indicate (§2.2.4)?

In what follows, the evidence used to address these questions will mainly be drawn from Bactrian historical phonology and the orthography of the Manichaean fragment.⁷⁹ Some of the linguistic features of the latter are

⁷⁸ ψ /ps/ does not seem to be attested, also not in its expected numerical value 700.

⁷⁹ This is not the place to give a comprehensive overview of Bactrian historical phonology, so I will only discuss points that are directly relevant to my argumentation. SFBG (31ff.) outlines the main changes but differs in some crucial

typical of Late Bactrian and thus indicative of a late composition of this text.⁸⁰ As a result, one could be inclined to regard also those linguistic features of Manichaean Bactrian not deducible from texts in Graeco-Bactrian orthography, in particular some of the palatalisations treated in §2.2.4, as being of a late nature and hence of little to no value to our understanding of Kuṣāṇa Bactrian phonology. In this study, I adhere to a pragmatic policy: linguistic features observable in the Manichaean fragment are posited for Kuṣāṇa Bactrian (i) when internal evidence (e.g. from the relative chronology of sound changes) proves that these features are indeed not recent innovations, or (ii) when they fit with the loanword data discussed in Chapters 3–4, and counter-evidence against their being old is lacking.

2.2.2 The Bactrian vowel system and the notation of /h/

The phonological system of Bactrian included at least ten vowels: the vowels /a/, /e/, /i/, and /u/ in both a short and a long version; a long /ō/;⁸¹ and a central vowel /ə/ (cf. Sims-Williams 1989b: 233f.).

The notation of long vowels varies. /a/ and /ā/ are both written with <α>, so one must decide on etymological grounds which one is intended. In the word ανδαρζο ‘instruction’, for instance, both alphas most likely stand for short /a/ because the word etymologically derives from *han-darza-, whereas the etymology of αζαδο ‘noble’ from *āzāta- suggests that its two alphas stand for long /ā/. The vowels /ē/ and /ō/ have separate symbols, i.e. <η> and <ω>. They occur in words such as ηβ- ‘to dispute’ <*āpaya-, οηλ- ‘to lead’ <*wādaya-, ωζο ‘power’ <*aujah-, and ρωσο ‘day’ <*rauḥah-. In Kuṣāṇa Bactrian, /ī/ and /ū/ could also be distinguished from their short counterparts /i/ and /u/: the digraph <ει> writes /ī/, e.g. in λαδειγο /lādīgə/ ‘righteous’ <*dāt-īya-ka-, and <ου> could, when not standing for /uh/ (cf.

details from the views adopted here. Francesca Michetti is preparing a comprehensive historical phonology of the language. For the Manichaean evidence, I build further on the comments made by Sims-Williams (2011a), whose conclusions are on many, though not all, points the same as mine. In §§4.3.8–4.3.21, I will return to some of the issues involved once the additional evidence of Bactrian loanwords into Niya Prakrit can be considered.

⁸⁰ The exact dating remains uncertain. Sims-Williams (2011a: 168f.) tentatively settles on “the late seventh or early eighth century C.E.” but notes that a still later date is not excluded.

⁸¹ As discussed below, conclusive evidence for a phonemic short /o/ is missing.

infra), spell /ū/, e.g. in (υ)αρουγο / (h)arūgə/ ‘all’ < *harwa-ka-.⁸² In later texts, <ι> and <ο> are the normal way to spell /ī/ and /ū/, as in λιδο /līdə/ ‘seen’ < *dīta- or σοδο /sūdə/ ‘plenty’ < *sūtV-. But for some words, the digraphs <ει> and <ου> can still be used in post-Kuṣāṇa times, sometimes even to write an apparently short vowel, as in σπουριγο /spurīgə/ < *us-pr̥na- + -ιγο < *-iya-ka- (jb 12).

The Graeco-Bactrian script did not have a distinct sign for /ə/. Yet, spelling variations and arguments from historical phonology can often be used to ascertain the presence of a schwa in a given lexeme (see Sims-Williams 1989b: 234). One of the most reliable indications is when a particular word is sometimes spelled with one vowel and sometimes with another, e.g. ναβιχτο next to νοβιχτο /nəvixtə/ (and archaic νιβιχτο /nivixtə/) ‘written’ < *ni-pixšta-. In Anlaut, /ə/ can occur as a prothetic vowel before consonant clusters or as the result of vowel weakening. Examples of the former are αχρονο ‘calendar/regnal year’ < earlier χρονο and αρκαρ- (also ερκαρ-) ‘to chase’ < *škara-; an example of the latter is αζοαστο (also ζοαστο) next to the more archaic οζοαστο ‘brought out, released, issued’ < *uz-wāsta-.

The majority of Bactrian words end in -o (BD II: 40), by the time of the later Bactrian documents probably no more than a graphic device to indicate word end. In Kuṣāṇa Bactrian, word-final -o may still have been pronounced as /ə/ (cf. e.g. Morgenstierne 1970: 126; Schwartz 1974: 409f.; Sims-Williams 1989b: 234), and this is how I will transcribe it.⁸³

GBactr. <o> can additionally be used at the border between a prefix and a root and between two members of a compound, e.g. πιδοβανδο ‘response, connection’ < *pati-banda- next to πιδβηγο ‘revered’ < *pati-baya-ka- (BD II: 254) and ανδαροζαγγο besides ανδαρζαγγο ‘of another kind’ < *antara-zana-ka- (BD II: 192). In many such cases, it is difficult to decide if <o> is only used as a graphic device or if it indicates a schwa (phonemic or not).⁸⁴

⁸² These spelling habits probably find their origin in the Hellenistic Greek pronunciation of <ει> and <ου> (VG³: 69f.; 77f.; Kantor 2023: 516–544; 704–717). For the etymologies of λαδειγο and (υ)αρουγο, see Sims-Williams & Cribb (1996: 90; 92).

⁸³ Jügel (2015: 168 fn. 429) plausibly suggests that “Im Baktrischen könnte -o, das nahezu jedes Wort wie ein Marker des Wortendes abschließt (...), auf den Nominativ oder Akkusativ Singular zurückzuführen sein (*am > *u > ?ə)”. See further Kreidl (2024: 208f.) and Michetti (fthc.: §7).

⁸⁴ When discussing Bactrian loanwords into Niya Prakrit, I will, for reasons of convenience, transcribe <o> in examples of this type with /ə/ when there is a

As already briefly indicated above, GBactr. <o> is additionally used to write /u/ (and /ū/), e.g. in ορλαγνο /urlaynə/ ‘name of a divinity’ < *vr̥θra-gna- or κοδο /kudə/ ‘dog’ < *kutV-. In addition, <o> can represent the semivowel /w/, as shown by instances like οαρο /wǎrə/ ‘there’ < *awaθra or φρογαοο /frəγāwə/ ‘profit’ < *fra-gāwa-.⁸⁵

A more intricate question is whether <o> is also used for /o/. In view of the existence of a short vowel ε /e/, there may have been a phoneme /o/ as well, but it is not so easy to find a straightforward example of it. Some scholars assume that when <o> indicates the outcome of *u*-umlaut of **a*, it stands for /o/ (so e.g. SFBG: 65; Bichlmeier & Benvenuto 2022ab). Given that *i*-umlaut of **a* results in /i/, e.g. -βιδο /vidə/ < *pati-, I consider it more plausible that *u*-umlaut of **a* yields /u/. According to Sims-Williams (1989b: 234), the first o in βο(o)- ‘to be(come)’ < *bawa- could also be an example of /o/, but it is equally possible that βο(o)- stands for /vuw/ rather than /vow/.⁸⁶ I therefore believe that we can for now do without the postulation of a Bactrian phoneme /o/.

Not being used for /u/, GBactr. <v> instead serves to write the phoneme /h/ (see already Herzfeld 1930: 9f.). Examples of this spelling are υρζ- ‘to release’ < *hr̥z- (cf. MBactr. *hyrz-*) and λαυ- ‘to give’ < *daθa- (cf. MBactr. *lh-*).⁸⁷ Most likely, this spelling habit stems from the default aspiration of initial *v-* in Ancient Greek, i.e. *v̥-* /hu/ > Attic and Koiné /hü/; this hypothesis is confirmed by the early Bactrian usage of <v> to write initial /hi/ in the word υνδο ‘India’ (Rab. 4; 7; 19; NPlate 3) and its derivative υνδοοαο ‘in Indian’ (Rab. 10).⁸⁸

corresponding vowel in Niya Prakrit, but assume it to be a dead vowel when there is no Niya correspondent. This practice sometimes results in minor inconsistencies, for which the reader’s indulgence is requested.

⁸⁵ After a vowel, /w/ can also be spelled with double <oo> in texts after 465 AD (SCBD: 62).

⁸⁶ In SCBD (62 fn. 198), Sims-Williams and de Blois also transcribe subj.1sg. βοοαο ‘I will be(come)’ as “[vuwān]”.

⁸⁷ Because of the development of OIr. **θ* to Bactr. /h/ in initial and intervocalic positions, instances of <θ> in writing (e.g. ιθαο ‘thus’ < *iθā) are probably historical spellings (cf. BD II: 218; SFBG: 38).

⁸⁸ Sims-Williams (2008c: 59f.) appears to attribute the value /hi/ to the “later” Greek sound change of /ü/ to /i/, but this change is usually assumed not to have occurred before the 10th or 11th century AD (Holton et al. 2019: 11–13; cf. also Kantor 2023: 544–558). Therefore, I deem it more likely that Gr. /ü/ was substituted with Bactr. /i/ as the next closest available sound, in the same way

Starting from the variant spellings $\mu\upsilon\rho\omicron$ / $\mu\upsilon\rho\omicron$ / $\mu\omicron\rho\omicron$ ‘Mihr’ < **miθra-* in Kuṣāṇa-period sources, Sims-Williams (2008c: 60) has come up with a hypothesis on how the value of <υ> changed from /hi/ to /h/. According to him, <υ> could first only be used for word-initial /hi/ since /mihirə/ < **miθra-* with its word-internal /hi/ is written as $\mu\upsilon\rho\omicron$ in Rab. 10 and Kuṣāṇa-period coins.⁸⁹ Starting with coins datable to Kaniška’s reign, the Bactrian outcome of **miθra-* can alternatively be written as $\mu\upsilon\rho\omicron$, the standard spelling in later times. Assuming that $\mu\upsilon\rho\omicron$ likewise stands for /mihirə/, Sims-Williams considers this form to represent the next stage in the development when <υ> could also be used for word-internal /hi/. The final step would have been when $\mu\upsilon\rho\omicron$ came to be regarded as equivalent to $\mu\omicron\rho\omicron$ (also found on coins of Kaniška). The latter is transcribed by Sims-Williams as /mihərə/, on the assumption that <ο> marks a svarabhakti vowel /ə/ and /h/ is left unwritten.

Although Sims-Williams’ hypothesis seems plausible overall, one may still wonder if the <ορ> in $\mu\omicron\rho\omicron$ really writes /hərə/ and not simply /hr/. In the case of $\mu\omicron\rho\omicron$, the variant $\mu\upsilon\rho\omicron$ /mihirə/ provides clear evidence that the cluster /hr/ < **θr* was (eventually) broken up by a svarabhakti vowel, but <ορ> is also used to write the outcome of **-θr-* in words such as $\beta\alpha\rho\omicron$ ‘kingdom, city’ < **xšaθra-* and $\pi\omicron\rho\omicron$ ‘son’ < **puθra-* (cf. Lazard, Grenet & de Lamberterie 1984: 222f.).⁹⁰ If, as per Sims-Williams, the cluster /hr/ < **-θr-* was broken up by a schwa also in these words, I find it somewhat surprising that in later Bactrian they are always spelled as $\beta\alpha\rho\omicron$ /šahrə/ and $\pi\omicron\rho\omicron$ /puhrə/, never as ** $\beta\alpha\upsilon\rho\omicron$ /šahərə/ or ** $\pi\omicron\upsilon\rho\omicron$ /puhərə/. I therefore have a slight preference to interpret the digraph <ορ> as an archaic spelling of /hr/, even though it is admittedly hard to understand why <ορ> would have been adopted to write /hr/.⁹¹ This potentially contentious point does

that speakers of Coptic approximated Gr. /ü/ as /i/ (Gignac 1976: 266f.). Another instance of the same substitution may be Bactr. $\omicron\iota\beta\omicron\eta\gamma\gamma\omicron$ (and variant) ‘made from linen or cotton’, an adjective made from * $\omicron\iota\beta\omicron$ ‘linen, cotton’ ← Gr. $\beta\acute{\upsilon}\sigma\sigma\omicron\varsigma$ ‘idem’ (Sims-Williams 2002: 228; BD II: 248).

⁸⁹ Cf. also the PN $\mu\omicron\rho\omicron\zeta\alpha\delta\alpha$ /mihirəzādā/ < **miθra-zāta-* in Ayrtaṃ 6, but note that Michetti (fthc.: Appendix) prefers to read $\mu\omicron\rho\omicron\zeta\alpha\delta\alpha$ with only one iota.

⁹⁰ A further example is obl.sg. $\pi\omicron\rho\omicron$ ‘father’ < gen.sg. **piθrah* (NPlate 1, Sims-Williams 2015: 256; 259).

⁹¹ Julian Kreidl (p.c.) made the interesting (but speculative) suggestion that <ορ> could result from a graphic simplification of earlier *<θρ> given that <θ> and <ο> are palaeographically similar, and **θr* is the most common source of Bactr. /hr/. A word where <ορ> could be used to write a cluster /hr/ that is not derived from

not affect the arguments about Bactrian loanwords into Niya Prakrit in any significant way, and I will spell /hr/ as <υρ> when reconstructing Bactrian words.

2.2.3 The phonological value of β, γ, and δ

In Greek, the value of <β>, <γ>, and <δ> changed from /b/, /g/, and /d/ to /v/, /ɣ/, and /ð/ at some undetermined point in time after the classical period, i.e. after the 4th century BC, possibly even much later (VG³: 29–32; Kantor 2023: 94–131). This raises the question which value these graphemes had in Bactrian. Were they used for plosives or fricatives, or were they used for both? If the latter, how can we distinguish when, for instance, <β> stands for /b/ and when for /v/?

2.2.3.1 -δ- /d/

Let us begin with the least complicated letter, i.e. <δ>. In most of its occurrences in inherited words, Bactrian <δ> derives from OIr. *t. Firstly, *t lenited to <δ> intervocally, e.g. λαδο ‘law’ < *dāta-, πιδο ‘father’ < *pitā-, nom.sg. of *pitar-, and καδο ‘house’ < *kata-ka-.⁹² This lenition is to be dated before the occasional procope of initial vowels, as demonstrated by, amongst others, δανομανο ‘so-and-so’ < *aitāt-nmāna-. Secondly, *t became <δ> when it stood in contact with a (secondarily) voiced consonant. Words belonging to this category are μαρδο ‘man, slave’ < *martiya-, ανδαρο ‘in between’ < *antara-, and αλωγδο ‘fully privileged’ < *hada-uxta-. After a nasal or a voiced sibilant, Bactr. <δ> can alternatively be from OIr. *d, e.g. βανδαγο ‘servant, slave’ < *banda-ka- and αζδο ‘knowledge, information’ < *azdā (cf. also de Blois 2013: 269).⁹³

*θr is the divine name αρομοζδο < *ahura-mazda- (Rab. 10); I would transcribe this name as /ahrəmuzdə/, assuming syncope of the *u in *ahura-, as also happened in the month name αρηζνο /ahrēznə/ < *ahura-yazniya- ‘(month) of the worship of the ahura(s)’ (BD II: 198).

⁹² When in secondary contact with another *t, the outcome is τ /t/, e.g. στο ‘and’ < *uta + *uti or βαταρο ‘happier’ < *čyāta-tara-. τ /t/ can also represent the outcome of older *-št-, e.g. αταο ‘eight’ < *aštā / *aštāu (SFBG: 47).

⁹³ In other positions, *d developed to Bactr. λ /l/. From the 7th century AD onwards, λ from *d is in some clusters reversed to δ, e.g. λρ and ρλ (SCBD: 60). In late texts, there accordingly are more environments in which Bactr. <δ> can come from *d.

Unfortunately, it cannot be established with certainty if GBactr. <δ> stood for /d/, /δ/, or both. The Manichaean fragment only uses a grapheme <d>, never <δ>. This could be taken as evidence that there was no phoneme /δ/ in Bactrian, seeing that the Manichaean fragment does distinguish /b/ from <β> /v/ and <g> /g/ from <γ> /γ/ (see below). It is, however, also possible that Manichaean Bactrian simply did not differentiate in writing between /d/ and /δ/, as Manichaean Middle Persian and Parthian do not do so either (cf. Sims-Williams 2011a: 170; SFBG: 34).

By lack of evidence to the contrary,⁹⁴ I follow Morgenstierne (1970: 126), Sims-Williams (1989b: 233), and Gholami (SFBG: 17) in assuming that GBactr. <δ> by default writes /d/ and not /δ/ (cf. Table 2, p. 62). This assumption will not affect the arguments about Bactrian loanwords in Niya Prakrit made in Chapters 3–4 (cf. §4.3.13).

2.2.3.2 -γ- /g/ and /γ/

GBactr. <γ> can, to begin with, represent the outcome of OIr. *k. The latter could undergo intervocalic lenition in Bactrian, as in ωγο ‘one’ <*aiwa-ka- or αγπο ‘ineffective’ <*a-kārya- (SFBG: 35),⁹⁵ and it could be assimilated to an adjacent voiced consonant, as in οισπαργο ‘everything’ <*wisparkā- or αγγιτο ‘received’ <*ham-kašta- (SFBG: 46).

⁹⁴ It is possible, as Sims-Williams (BD II: 39) points out, that in Late Bactrian <Δ> and <δδ> “were adopted as unambiguous ways of indicating [d] at a period when δ in most positions had come to be pronounced as a fricative [d̪]”. Yet, conclusive evidence in support of this hypothesis is lacking (cf. SFBG: 34), and the implied fricativisation of /d/ to /δ/ would, in any case, be a late change that does not affect the interpretation of Kuṣāṇa Bactrian phonology. Regarding loanwords, a potentially relevant item is the word *wrδn* ‘chariot’ in a multilingual Manichaean calendar from Turfan (U 130). Sims-Williams (2019b: 258f.) persuasively argues that *wrδn* represents Bactr. *οαρδανο < OIr. *wartana-. However, the fact that *wrδn* is written with <rδ> does not necessitate that Bactr. *οαρδανο was pronounced by native speakers as /warδanə/ and not as /wardanə/. One could also say that /warδanə/ was “Bactrian with a Sogdian accent”, seeing that /d/ is only found after nasalised vowels in Sogdian (Sims-Williams 1989a: 178; Yoshida 2009: 286).

⁹⁵ For the reconstruction *a-kārya-, see Sims-Williams (2015: 260). At a morpheme boundary or in the case of the hypocoristic suffix *-k(k)a-, intervocalic *-k- is sometimes preserved, e.g. in νακανδο /nəkandə/ ‘dug’ <*nikanta- or ζαδακο /zādak(k)ə/ ‘child’ <*zāta-k(k)a-.

As the Manichaean Bactrian script consistently differentiates <g> from <γ>, this can help us to decide whether GBactr. <γ> from *k stands for /g/ or /γ/. In the Manichaean fragment, the voiced outcome of OIr. *k is always spelled with <g>, as exemplified by, amongst others, *‘sprγmyyg* ‘flower’ < **sprgma-ka-*, *βrg* ‘fruit’ < **bara-ka-*, and *βyžg* ‘sin’ ~ *βιζαγο* < **bazdya-kǎ-*. These examples prove that *k probably had nowhere developed further than the plosive /g/ when the Manichaean folio was written (Sims-Williams 2011a: 170). Consequently, the same must be true of older stages of the language.

A second source for GBactr. <γ> is OIr. *g, be it word-initially, intervocally, or as part of a consonant cluster (SFBG: 36; 44). Words belonging to this category are, for instance, *γαμβο* ‘damage’ < **gamba-*, *λωγω* ‘falsehood’ < **drauga-*, *οιγνο* ‘famine’ < **waignā-*, and *αβυαγγο* ‘deduction, levy’ < **apa-θanga-*.

When resulting from *g, GBactr. <γ> corresponds to MBactr. <γ>, so we can posit a sound change from OIr. *g to Bactr. /γ/ (Sims-Williams 2011a: 170).⁹⁶ This change was likely inhibited by a preceding nasal, given that OIr. *d was also continued unchanged after *n (cf. *supra* and Sims-Williams 1989b: 233).⁹⁷

Relative chronology suggests that the fricativisation of *g occurred before the start of Bactrian literacy in the 1st century AD. Considering that Bactr. /g/ from *k apparently did not develop to **/γ/, the change of *g to Bactr. /γ/ must predate the voicing of *k, and since this voicing had already been completed in our earliest texts, *g > /γ/ needs to have happened before

⁹⁶ There are two examples of this correspondence in initial position, i.e. *γ’w* ~ *γαι(ι)* ‘cow’ < **gǎw-* and *γlyh* ~ **γαλιω* (earlier *γαλιγο*) ‘theft’ < **gadya-kǎ-*, and seven instances in non-initial position, i.e. *’γd* ~ *αγαδο* ‘came’ < **ā-gata-*; *’wγwt* ~ *ωγοτο* ‘secrecy (?)’ < **awa-gušta-*; *‘sprγmyyg* ‘flower’ < **sprgma-ka-*; *γγwng* ‘as’ ~ **σαγωγγο* or **σαγογγο* (later *σογγο*) < **čiyāt-gauna-ka-*; *fr’w* ~ *φρογασο* ‘profit, wealth’ < **fra-gāwa-*; *myγdyg* ~ **μυγδιο* ‘fruit’ < **migda-ka-*; and *rwγwn* ~ **ρωγνο* ‘butter’ < **raugna-*. I leave out *r’γγg* ‘evil’ (cf. GBactr. *ραγο* ‘idem’) since its etymology is unknown (BD II: 259; Sims-Williams 2009a: 264).

⁹⁷ Every instance of MBactr. <ng> in words whose etymology is known continues **-nVk-*; cf. the suffix *-’ng* < **āna-ka-*; *γγwng* ‘as’ < **čiyāt-gauna-ka-*; *frwng* ‘foreseeing’ < **fra-waina-ka-*; *hβyśznng* ‘of many kinds’ < *hβyś* ‘much’ + **zana-ka-*; and the suffix *-yng* < **aina-ka-*.

Kuṣāṇa times as well. It is therefore justified to transcribe GBactr. <γ> from **g* as /γ/ in every attested stage of the language.⁹⁸

In case of a cluster -γδ-, GBactr. <γ> can also derive from OIr. **x*, as in e.g. αλωγδο ‘fully privileged’ < **hada-uxta*-. There are no words of this type in the Manichaean fragment;⁹⁹ but seeing that <γ> results here from a fricative, chances are high that -γδ- from **-xt-* was pronounced as /γδ/ and not as **/gd/ (SFBG: 47).¹⁰⁰

In short, to determine the most likely value of GBactr. <γ> in a given word, one can use the following rule of thumb: when occurring after a nasal or representing the outcome of OIr. **k*, <γ> can be assumed to stand for /g/; if the result of OIr. **g* or **x*, <γ> is to be transcribed as /γ/.

2.2.3.3 -β- /b/ and /v/

A first source for GBactr. <β> is OIr. **b*, as can be illustrated by words such as βιζαγο ‘sin’ < **bazdya-kā-*, αβιρ- ‘to obtain’ < **abi-ār(a)ya-*, and λαδοβαρο ‘judge’ < **dāta-bāra-*. In Manichaean orthography, these words are spelled as βγζγ, βγρ-, and λδβρ, so there is proof that OIr. **b* developed to Bactr. /v/. As it is likely that this development occurred simultaneously with the lenition of **g* to /γ/, i.e. before Bactrian literacy started (cf. supra), one can phonologically transcribe GBactr. <β> from **b* as /v/ also in words attested in Kuṣāṇa-period texts.

The only exception concerns the position after a nasal. As briefly discussed under <γ>, a preceding nasal probably prevented the lenition of voiced stops. As a result, one can safely assume that, whatever its

⁹⁸ The chronology of Bactrian lambdacism, i.e. the change of OIr. **d* to Bactr. λ /l/, could provide circumstantial evidence that the change of **g* to /γ/ occurred even much earlier than Kuṣāṇa times. One can approximately date Bactrian lambdacism to the 4th or 3rd century BC (de Blois 2013; Kreidl 2021; see §3.2.2 s.v. *milima*), which would be a terminus ante quem for OIr. **g* becoming a voiced fricative in Bactrian—if we accept as premises that (i) the change of **d* to Bactr. λ passed through an intermediary stage **δ* (so e.g. Morgenstierne 1970: 126; Sims-Williams 1989: 233; 2004a: 542; Kümmel 2007: 78) and (ii) the hypothesised change of **d* to **δ* coincided with the fricativisation of **g* (and **b*, see below).

⁹⁹ The -γδ- in *myγdyg* ‘fruit’ < **migda-ka-* comes from **-γl-* < **-gd-* and not from **-xt-* (Sims-Williams 2009a: 263; 2016b: 278).

¹⁰⁰ In dialectal Bactrian, initial γ- /γ/ can also derive from OIr. **w-*, e.g. γολο /γυλə/ < **wadū-*. See §3.2.1 s.v. *guśura*.

etymological origin, -μβ- stood for /mb/ rather than **/mv/ (so also Sims-Williams 1989b: 233; 2011a: 170 *contra* SFBG: 45).¹⁰¹

Two further environments in which <β> may be assumed to stand for /v/ are the clusters -βδ- from *-ft- and -λβ- from *-dw-.¹⁰² Examples of the former are ταβδο ‘sealed’ < *tafta- and πιδοροβδο ‘received’ < *pati-grfta-. The fact that <β> continues a fricative is sufficient reason to assume that -βδ- from *-ft- was realised as /vd/ rather than as **/bd/ (cf. SFBG: 47 and *-xt- > -γδ- /γd/). The change of *-dw- to Bactr. -λβ- can be exemplified by αλβαρο (also αλβαργο in Uv 2f.) ‘court’ < *dwara(-ka)- and the PN βαραλβαγο, possibly from *bara-dwāga- ‘standard bearer’ (BPN: no. 73). In this case, the assumption that -λβ- stands for /lv/ is supported by the parallel development of *-θw- to Bactr. -λφ- /lf/ (Sims-Williams 2004a: 542f.; SFBG: 53).¹⁰³

Besides, GBactr. <β> can stem from OIr. *p. This can be the case when *p stood (originally) in intervocalic position, as in αβαβγο ‘waterless’ < *apāpa-ka- (BD II: 181); -δηβγο ‘sealer’ (attested in μολροδηβγο) < *tāpa-ka- × *τηβ- (pret. ταβδο) < *tāpaya- (BD II: 235); or χοαβο ‘dream’ < *hwāpa- (BD II: 245). If procope happened, voicing of *p can also be observed word-initially, e.g. βαμδδιγο ‘loan’ < *apāmita-ka-.¹⁰⁴

The change of OIr. *k to Bactr. /g/ and that of OIr. *t to Bactr. /d/ leads one to expect that <β> from *p stands for /b/. Yet, in the Manichaean fragment, the outcome of intervocalic *p is consistently written with <β>; compare, for instance, β ~ αβ(β)ο ‘water’ < *āp-; nβyxyg ~ ναβιχτιγο ‘written’ < *ni-pixšta-ka-; xw’β ~ χοαβο ‘dream’ < *hwāpa-; and xwβ ~ χοβο ‘one’s own, property’ < *hwa(i)-paθya- (Sims-Williams 2011a: 170). *b < *p thus seems to have progressed further to Bactr. /v/, even though “it is not clear ... whether this development was already complete in the Kushān

¹⁰¹ The only instance of <mb> in the Manichaean folio occurs in a loanword, i.e. in *zmbwd’ng* ‘wordly, earthly’ < *zmbwd** ← Skt. *jambudvīpa-* (with adaptation) + suffix -’ng < *āna-ka-.

¹⁰² There is no Manichaean evidence for these clusters. A third environment could be -ζβ- from *-zw- in case the sequence εζβαγο in h4 4 means ‘tongue’ from *hizwā-kā- (Sims-Williams 2006: 115).

¹⁰³ The development of *dw- may have been context-sensitive. When followed by *i, *dw- is continued as Bactr. β- in βιδδιγο ‘second’ < *dwitīya-ka- and βαυιασπανο ‘messenger’ or a PN < *dwiy-asp-āna- (BPN: no. 66). I find it difficult to decide if in these words β- stands for /b/ or /v/.

¹⁰⁴ See above under <δ> /d/ for the fact that voicing predates the occasional procope of initial vowels.

period” (Sims-Williams 1989b: 233 fn. 25).¹⁰⁵ The Bactrian loanwords in Niya Prakrit are consistent with an early date for this sound change, although they cannot prove it (see §4.3.8). Therefore, I assume—albeit with due caution—that the outcome of postvocalic OIr. **p* was already /v/ in Kuṣāṇa Bactrian.¹⁰⁶

A more tricky question is whether initial GBactr. <β> from **p*, as in βαμδδιγο < **apāmita-ka-*, stands for /v/ or /b/. In this case, the value of <β> depends on the chronological order of the relevant sound changes, on which decisive evidence appears to be lacking. If procope preceded the lenition of **b* < **p* to /v/, βαμδδιγο would write /bāmddigə/. If the lenition happened first, the phonological realisation would be /vāmddigə/. This uncertainty affects one of the Bactrian loanwords in Niya Prakrit proposed in Chapter 3, i.e. *badho* ‘pledge, mortgage’ ← **βαμδαοο* ← **βαμδο* ‘loaned’ < **apāmita-* + suffix *-αο* (see §3.2.3 s.v.). This etymology requires the initial β- in **βαμδαοο* to have been a /b/, which is only possible if procope happened before the lenition of **b* < **p* to /v/.¹⁰⁷ Therefore, I have hesitantly chosen to transcribe GBactr. <β> as /b/ when it continues **p* in initial position after procope.

To summarise, I will transcribe <β> as /v/, except when a nasal precedes or when procope happened. In the latter cases, I take <β> to be /b/.

2.2.4 The phonological value of ζ, σ, and þ

With <ζ>, <σ>, and <þ>, one should decide (i) how many sounds they represent in the first place; (ii) whether Bactrian phonology included only sibilants or also affricates; and (iii) if, as suggested by the Manichaean fragment, sibilants and affricates could be affected by processes of palatalisation. I will start the discussion of these questions with the least contentious of the three graphemes, i.e. <þ>.

¹⁰⁵ Morgenstierne (1970: 127) compares the same development of intervocalic **p* to /v/ in neighbouring Munji-Yidgha and Sanglechi-Ishkashimi.

¹⁰⁶ The interchange between -β- and -φ- /f/ seen in the PN σπαλβο / σπαλφο < **spāda-pā-* ‘army-chief, general’ (or shortened from σπαλοβιδο ‘idem’ < **spāda-pati-*, cf. BPN: no. 449) provides indirect evidence that <β> from postvocalic **p* also stands for /v/ when the labial no longer stands in intervocalic position synchronically, due to syncope of the relevant vowel.

¹⁰⁷ Although this does not prove its phonemic status, initial /b/ also occurs in loanwords into Bactrian, e.g. MBactr. *b’gyg* ‘having a share’, most likely a Parthian loanword, and *bwt* ‘Buddha’ ~ βοτο (also βοδδο / βουδο) ← Skt. *buddha-* or Gandh. *budha* (Sims-Williams 2009a: 258; 262).

2.2.4.1 -p- /š/

GBactr. <þ> corresponds to the Manichaean grapheme <š>, as exemplified by words such as *rštyg* (cf. Rab. 1 ραρθογο) ‘righteous, just’ < **rašta-ka-* or *pydkšt* ~ πιδοκιπτο ‘restrained, prevented’ < **pati-kṛsta-* (Sims-Williams 2011a: 171). <þ> is therefore commonly transliterated as /š/ (cf. e.g. Sims-Williams 1989b: 233; SFBG: 17).¹⁰⁸

When ignoring loanwords and ad hoc assimilations,¹⁰⁹ the etymological sources of Bactr. <þ> /š/ can be divided into four categories.

First, <þ> is used for the outcome of OIr. *š in intervocalic position and in certain consonant clusters;¹¹⁰ compare, for instance, νοβαλμο /nəšalmə/ (also archaic νιβαλμο) ‘seat’ < **ni-šadma(n)-* or αβιπταδο /avištādə/ ‘master’ < **abi-štāta-*.

Second, <þ> can result from a simplified cluster involving *š, i.e. *-xš- and *-rš- (*-rš-). Examples of this category are þαυρο /šahrə/ ‘city’ < **xšaθra-*; ιαμθο /yamšə/, divine name < **yama-xšāwan-*; φροχοαþ- /frəxwaš/ ‘to withdraw’ < **fra-xwarša-*; and κιþαγο /kišāgə/ ‘plough-ox’ < **kṛš-āka-*.

A third possibility is that <þ> comes from an OIr. *s (< PIr. *ś) that was retracted by a neighbouring *r (*r) in clusters such as *-sr-, *-str-, or *-ršt-.¹¹¹ This category can be exemplified by words like þουο /šūhə/, a late form of *þογο /šūgə/ ‘fingernail’ < **srū-kā-*; χοοþο /xwušə/ ‘mother-in-law’ < **hwasrū-*; οαþο /wāšə/ ‘hay’ < **wāstra-* (MBactr. wš); and πιδοκιπτο /pidkištə/ ‘restrained, prevented’ < **pati-kṛsta-* (MBactr. *pydkšt*).

Finally, <þ> can stem from *čy, be it initially, e.g. þαδο /šādə/ ‘pleased’ < **čyāta-*, or intervocalically, e.g. αβαþο /avāšə/ ‘back, in return’ < **apāčyā-*.

¹⁰⁸ See further §4.3.18, where based on loanword evidence I argue that <þ> may have been phonetically realised as a retroflex [ʂ].

¹⁰⁹ For ad hoc assimilations, cf. ζιþτο /zištə/ ‘requested’ < **ḡasta-*, i.e. *s > *š due to assimilation to the preceding *ḡ (BD II: 212), and ναυαþτο /nəhāštə/ ‘settled’ < **ni-šāšta-* < **ni-šāsta-*, i.e. *s > *š due to assimilation to the preceding *š (BD II: 239; Michetti 2024: 175–177).

¹¹⁰ After back vowels, *š develops to Bactr. <u> /h/, e.g. in ασνωουο /əsnōhə/ ‘daughter-in-law’ < **snaušā-* (Michetti 2024). This point is of no concern for the present argument.

¹¹¹ *r only causes retraction of *s in clusters in which *r is lost, not in the case of intervocalic *-rs- (*-rš-); cf. e.g. ταρσο /tarsə/ ‘fear, respect’ < **tarsV-* and πορσ- /purs/ ‘to ask’ < **pṛsa-* (SFBG: 56).

2.2.4.2 -σ- /s/, /ś/, /ts/, and /tś/

The value of GBactr. <σ> depends on the etymological source and phonetic environment of the sound which this letter represents. To begin with, <σ> is used to represent the default outcome of OIr. *s (< Plr. *č) in words such as σαδγο ‘levelled, flat’ < *sāta-ka-, λασο ‘ten’ < *dasa, and πωστογο ‘document’ < *pawasta-ka-. In such cases, the corresponding grapheme in Manichaean script is <s>, e.g. *srd’nyg* ‘pertaining to years’ ← *srd** ~ σαρδο < *σαρλο ‘year’ < *sardV-. Hence, we can safely conclude that at least in an unconditioned environment, <σ> from *s writes /s/.¹¹²

GBactr. <σ> can also be the result of OIr. *č. Two such examples are σανο ‘pleasure’ < *čanah- and ρωσο ‘day’ < *raučah-. In the Manichaean fragment, the outcome of *č is written as <c> in the following nine examples:¹¹³

- (i) *’cyd* (‘*cyyd*; *cyd*) ~ (α)σιδο ‘which’ < *čim + *uti
- (ii) *c’’k’myg* ‘whosoever’ (cf. MP *čegām-iz* ‘whatever’, *čegāmag* ‘a bit’)
- (iii) *c’sq* ~ (α)σασκο ‘over, upon, against’ < *hačā + *uskād
- (iv) *cγwwng* ~ *σαγωγγο / *σαγογγο (later σογγο) ‘as’ < *čiyāt-gauna-ka-
- (v) *hrc* ~ υαρσο ‘absolutely, at all, just’ < *harwa- + *čim
- (vi) *qyc* ~ κισο ‘anyone’ < *ka- + *čim
- (vii) *pyc* ~ πισο ‘before, in the presence of’ < *patīča
- (viii) *rwc’nyg* ~ *ρωσανιγο ‘pertaining to days’ ← *rwc** ~ ρωσο ‘day’ < *raučah-
- (ix) *-yc* ~ -(ι)σο ‘also, too’ < *čim

At the same time, there is one word in which the result of *č is spelled with MBactr. <s>, i.e. ‘s ~ ασο ‘from’ < *hačā.

The double treatment of *č in the Manichaean folio is helpful to assign a phonological value to MBactr. <c>. This grapheme probably stood for /ts/ because, as Sims-Williams convincingly argues (2011a: 171), the spelling of the proclitic preposition ‘s ~ ασο with <s> is then easily explained as due to a sandhi development of word-final /ts/ to /s/ before consonants.¹¹⁴ In the

¹¹² See below for a possible palatalisation of *s to /ś/ when next to *ī.

¹¹³ Partially preserved and etymologically unclear words are not included.

¹¹⁴ Gholami (SFBG: 36f.; 50) also considers the possibility that <σ> from *č stands for /ts/, but her account of the sibilants and affricates suffers from internal contradictions, as pointed out by Ciancaglini (2016: 269).

case of *c'sq* ~ (α)σασσο < **hačā* + **uskād*, this sandhi development would not have occurred because **hačā* was here followed by a vowel.

An additional argument for the depalatalisation of *č in Bactrian is BSogd. *sxrh* 'wheel', which is most likely a loanword from Bactr. *σαχρο < *čaxra- (Sims-Williams 1979: 133 fn. 7; 1996a: 50).¹¹⁵ The orthography of MBactr. *c'k'myg*, *c'sq*, and *cywwng* suggests that word-initially, /ts/ from *č was not simplified to **/s/. The initial *s-* in BSogd. *sxrh* may therefore be due to a sound substitution of Bactr. /ts/ with Sogd. /s/ caused by the lack of a Sogdian cluster /ts/ in initial position.

Further evidence that /ts/ is one of the values of GBactr. <σ> comes from the observation that the syncopated outcome of OIr. **tVs-* can also be written with <σ>. A few instances of such spellings are πισηβ- 'to slander, to accuse' < **pati-sāpaya-*, υαλοπισαγο 'mutual agreement' < **hada-pati-sākā-*, and πισαγ(α)δο 'agreed' < **pati-saxta-*. The latter has a variant πιδοσαχτο /pid(ə)saxtə/, which makes it plausible that the <σ> in πισαγ(α)δο (and thus also in πισηβ- and υαλοπισαγο) stands for /ts/.

A final piece of evidence that the Bactrian outcome of *č is an affricate /ts/ comes from some recently published birchbark letters (Sims-Williams 2022–23). These new documents employ (not entirely consistently) a slightly different orthography than the one used in the Bactrian text corpus known until now: the outcome of **s* is spelled, as normal, with <σ>, but the outcome of *č can be written as <σσ>, e.g. ασσο 'from' < **hačā*, ρωσσο 'day' < **raučah-*, and ναμωσσο 'homage' (Sims-Williams 2022–23: 139).¹¹⁶ This orthographic peculiarity is plausibly explained by Sims-Williams as a way to disambiguate /ts/ from /s/.

So far, we have seen that GBactr. <σ> can correspond to MBactr. <s> and <c>. There is yet a third possibility: some words which in the Graeco-Bactrian script are (or would be) written with <σ> are in Manichaean orthography spelled with <ś>. This grapheme occurs in the following six words:

¹¹⁵ Sogd. *cxr-* 'wheel' is the inherited word.

¹¹⁶ The spelling ασσο suggests that the simplification of /ts/ to /s/ witnessed by MBactr. 's ~ ασο is a later development. The etymology of ναμω(σ)σο (also ναμασο / ναμοσο) 'homage' still awaits a definitive solution (see BD II: 238; Albino 2019 for some proposals). Yet, its Middle Persian and Sogdian cognates, i.e. *namāz* and *nm'cyw*, at least confirm that <σσ> / <σ> derives from earlier *č.

- (i) *hβyśznng* ‘of many kinds, manifold’ ← *hβyś* ~ *υαβισο* ‘many, much’¹¹⁷ + *znng** ~ *ζαγγο* ‘manner, way’ < **zana-ka*-
- (ii) *nyšt* ~ *νιστο* ‘is not’ < **nai(d) asti*
- (iii) *rymś* ‘blame’ (cf. Sogd. *rymysš*)
- (iv) *śwwh-*, a verb used with the Indian loanword *nrh* ‘hell’ as its direct object and approximately meaning ‘to escape, to avoid’¹¹⁸
- (v) *tyśygyg* ‘emptiness’, an abstract in *-yg* from *tyśyg** ‘empty’ < **tusya-ka-*
- (vi) *wyśp* (*wyś*) ~ *οισπο* (*οισο*) ‘all’ < **wispa-*; also attested in the compound *wyśpz’nyndyg* ‘all-knowing’

Apart from *śwwh-* (see fn. 118), every instance of MBactr. <ś> is found after <y>, which implies that we are probably dealing with a secondary palatalisation of *s to /ś/ after *ī (Sims-Williams 2011a: 171; SFBG: 38).¹¹⁹

Before using these Manichaean data to posit a Bactrian phoneme /ś/, one needs to address two questions: (i) had this palatalisation already been completed in Kuṣāṇa Bactrian, and (ii) from which sound combinations could a phonemic /ś/ have arisen in Bactrian?

Regarding the first question, one has to choose between two approaches. Option one is to emphasise the late date of the Manichaean fragment and put

¹¹⁷ The etymology of this word is debated. Sims-Williams (BD II: 271), building on work by Gershevitch, suggested the following three possibilities (admitting none of them to be obviously correct): (i) from **ha-bai-sah-*, comparable to Khot. *hambīsa-* ‘heap’; (ii) from **ha-bus-ya-*, with NP *ambōh* ‘multitude, much, many, etc.’; and (iii) from **ha-paisa-* ‘variegated > varied > much’. Durkin-Meisterernst (*apud* SFBG: 38f.) proposed to derive *υαβισο* / *hβyś* from **harwa-wispa-* (cf. Khot. *harbiśśa-* ‘all’, DKS: 469). This etymology is phonologically problematic, however. There is no evidence for *-*rw-* becoming -β- /v/ (contrast (υ)αρουγο / (h)arūgə/ ‘all’ < **harwa-ka-*), and the assumed simplification of -σπ- to -σ- as in *οισο* < *οισπο* ‘all’ < **wispa-* is not attested before 671 AD (SCBD: 64), whereas *υαβισο* already occurs in doc. ba from about 350 AD (SCBD: 82) and as *αβισοι* ← **αβισο* + enclitic -οι < **čim* already in Rab. 21 (3x, cf. Sims-Williams 1998: 87f.).

¹¹⁸ The etymology of MBactr. *śwwh-* is considered unclear by Sims-Williams (2009a: 258; 2011a: 171). In Schoubben (fthc.), I argue that *śwwh-* is a loan from Gandh. **śoh(e)-* < OIA *śodhaya-* ‘purify; remove (something noxious); clear off; exculpate’, but for now this word can be left aside.

¹¹⁹ In the case of *nyšt* ~ *νιστο* ‘is not’ < **nai(d) asti*, the *ī goes back to earlier **ai* (via **ē*). Phonetically, Bactr. /ś/ could, for instance, have been a post-alveolar [ʃ] and /ź/ the respective voiced counterpart [ʒ].

aside the observed palatalisation of *s as a late (or dialectal) feature irrelevant for the phonological interpretation of earlier Bactrian. A second approach would be to highlight the absence of evidence proving these palatalisations to be comparatively recent developments; after all, there would be no problem with assuming GBactr. <σ> to write both /s/ and /ś/.

Two observations favour, in my view, the latter approach. First, there is relatively solid evidence for a triple contrast between /z/, /ž/, and /ǰ/ (see under <ζ> below), making the hypothesis of a corresponding threefold distinction in the voiceless series plausible. In addition, one secure Bactrian loanword into Gāndhārī, i.e. *guśura* ‘prince’ ← *γoσβορο /γyśvūrə/ or γoσoβοργo /γyśvürgə/ < **wisah-puθra(-ka)*-, corroborates an early palatalisation of *s to /ś/ after *i. To account for the palatal -ś- in *guśura*, one must namely assume that the *-i- in **wisah-* first palatalised the *-s- to /ś/ before itself being coloured to /u/ by the initial *w-. See in more detail §3.2.1 s.v. and §4.3.16.¹²⁰

Was *s also palatalised to /ś/ when *ǰ followed rather than preceded the sibilant, i.e. in words such as σνδ- ‘to wish, be pleased’ < **sandaya-*? Phonologically, such a palatalisation would be sufficiently plausible, and one could even draw a parallel with Sims-Williams’ argument (1999: 197; also BD II: 38f.; SCBD: 60) that the Post-Kuṣāṇa Bactrian use of a geminate <δδ> in front of palatal vowels finds its origin in a sound change of **dī* over **dyī* to <δδ> /ddī/ (and **dē* over **dyē* to <δδη> /ddē/).¹²¹ Manichaean evidence is unfortunately lacking, as no relevant word happens to be attested in M 1224.¹²²

¹²⁰ In loanwords from Indic into Bactrian, the Indo-Aryan palatal ś is also regularly substituted with Bactr. <σ> (Michetti 2024: 186f. fn. 84); cf. e.g. βησραμavo ← *Vaiśramaṇa-*, a king of the *yakṣas*; πιασο ← *piśāca-*, a type of demon; and βοδδοσαστρο ← *buddhaśāstra-* ‘teaching of the Buddha’ (TiS 7, Lee & Sims-Williams 2003: 170). This type of sound substitution can, however, not count as compelling evidence for a phonemic value /ś/ of Bactr. <σ>, as one could alternatively hypothesise that IA ś was substituted with the closest sound available, i.e. Bactr. /s/.

¹²¹ See also below for potential evidence in support of a parallel change of **zī* to /žī/. In the loanword corpus of Chapter 3, there is no example to show if *s would also be palatalised by a following *ǰ (or *ē).

¹²² There is one instance of a sequence <sy>, i.e. *c’p’syh*. Yet, no conclusions can be based on this etymologically unclear word, whose intervocalic -p- suggests it to be a loanword.

There could be one more input for a Bactrian phoneme /ś/, i.e. *sy, since the Bactrian loss of *y after sibilants, as in σπισ- ‘to serve, to worship’ < *spasya-, conceivably resulted in the palatalisation of the preceding sibilant (cf. Fattori 2022: 388 fn. 25). Yet, decisive evidence is again missing.¹²³ The only potentially relevant form in Manichaean Bactrian is *tyśygyg* ‘emptiness’ ← *tyśyg** ‘empty’ < *tusya-ka-, which cannot be used as compelling evidence for a change of *sy to /ś/ given that the <ś> in *tyśyg** is both preceded and followed by <y>.¹²⁴ Thus, it remains uncertain if <ś> directly results from *sy or if the observed palatalisation was rather induced by the neighbouring palatal vowel(s). To confirm conclusively that *sy yielded Bactr. /ś/, we would need a Manichaean example in initial position or in a sequence such as *-ausy-.

If one accepts that *s could be palatalised to /ś/ in Bactrian, it seems conceivable that *tsī < *čī also yielded /tśī/. Indirect support for this hypothesis comes from MBactr. *y’wyd’nzyg* ~ ιασηδανζιγο (kb 1) ‘for ever’ < *yāwaitāna-čiya-ka-, where the voiced counterpart *dzī has apparently been palatalised to /dźī/ (see below). In a few Bactrian borrowings into Niya Prakrit discussed in subsequent chapters, Bactr. <σ> from *čī is moreover rendered as Niya -ci-, i.e. with a palatal affricate (see §4.3.17).¹²⁵ On this admittedly slender basis, I tentatively assume GBactr. <σ> to stand for /tś/ in words such as (α)σιδο ‘which’ < *čim + *uti, αλιχτογο ‘inclined, desirous’ < *hadā-čixšta-ka-, and ηησινδ- ‘to approve’ < *pati-sandaya-.¹²⁶

Still, none of the etymologies proposed below crucially depends on the assumption of a phonological value /tś/. Etymologies that require Bactr. <σ>

¹²³ See also under <ζ> for examples such as MBactr. *βyžg* ~ βιζαγο /vižagə/ ‘sin’ < *bazdya-kā-.

¹²⁴ The apparent palatalisation of *u to <y> (perhaps standing for an unlauded /ü/?) may, however, be a quite late development, as there are no traces of such a change in the Graeco-Bactrian materials.

¹²⁵ Indo-Aryan *c* can also be rendered as Bactr. <σ>; cf. e.g. ασαριαριο ← *ācāryārya- ‘reverend doctor’; πιασο ← *piśāca-*, a type of demon; and σανδαροβανο ← *Candrabhānu-*, name of a Buddha. These loanwords are no conclusive proof that <σ> had a phonological value /tś/, given that IA *c* could also have been approximated with Bactr. /ts/. In the Rabatak inscription, the city name *Śrī-Campā* is written in Bactrian as ζιριτ[ι]αμβο (or ζιριταμβο), so here IA -c- would correspond to Bactr. -τ(ι)- /t(y)/ (cf. Sims-Williams 2008c: 61).

¹²⁶ If the interpretation of (α)σιδο ~ MBactr. ‘cyd (also ‘cygd; cyd) ‘which’ as /tśidə/ is correct, MBactr. <c> would have two phonological values, i.e. /ts/ and /tś/.

from *čř̄ to be rendered as Niya *ci* also work if one posits an allophonic pronunciation [tsĩ] or if one accepts that Bactr. /tsĩ/ was substituted with Niya *-ci-* as the closest sound combination available (§4.3.17).

In summary, GBactr. <σ> from OIr. *s (<PIr. *č) normally indicates /s/, but when preceded by ι /ĩ/, perhaps also when followed by it, a palatalisation to /š/ may have occurred. When it stems from OIr. *č or *-tVs-, <σ> most likely represents /ts/, although in front of ι /ĩ/, one may again want to reckon with a palatalised counterpart /tś/.

2.2.4.3 -ζ- /z/, /ž/, /ẓ̌/, /dz/, and /dž/

According to Sims-Williams (1989b: 233), the Graeco-Bactrian grapheme <ζ> would have three phonological values, i.e. /z/, /ž/, and /dz/ (taken over by SFBG: 17). When later discussing the phonology of the Manichaean fragment, Sims-Williams (2011a: 172–174) observed that <ζ> could also correspond to MBactr. <ẓ̌> or <j>, thus supporting an additional phonological value of <ζ> as /ẓ̌/. The first part of this section will summarise the conditions in which <ζ> can be assumed to stand for /z/, /ž/, and /dz/ respectively, while at the same time also briefly considering the possibility of an additional value /dž/. The second part will treat the slightly more contentious issue whether and, if so, when <ζ> could represent /ẓ̌/.

There is no doubt that GBactr. <ζ> writes /z/ when it represents the default outcome of OIr. *z (<PIr. *j), as in e.g. ζαμῖγο /zamīgə/ 'land' < *zamī-kā-, αζο /azə/ 'I' < *azam, γαζνο /γaznə/ 'treasury' < *gazna-, or ανδαρζο /andarzə/ 'instruction' < *han-darza-. In Manichaean script, the phoneme /z/ is written as <z> in words such as hz'r ~ υαζαρο /hazārə/ 'thousand' < *hazahra-, hvrz- ~ υιρζ- /hirz/ 'to let go' < *hrz-, or zwwz- 'to rise' ~ *(α)ζοαζ- / (ə)zwaz/ < *uz-waza-.

A phonological value /ẓ̌/, i.e. the voiced counterpart of /ṣ̌/,¹²⁷ can be assumed in the following cases. First, <ζ> probably represents /ẓ̌/ when it stems from OIr. *ž, as in μοζδοοαβο /muždwanə/ 'a divinity' < *miždwan- 'benevolent' (Sims-Williams 1997b) or χο(υ)ζο /xūžə/ 'good' < *xwṛžū-. Another input for Bactr. /ẓ̌/ is provided by clusters whose voiceless counterpart yields /ṣ̌/. The latter can, for instance, come from *-sr- in þovo

¹²⁷ Given that /ṣ̌/ was written with <þ>, one could expect there to have been a separate grapheme for /ẓ̌/ as well. The absence of such a grapheme is likely due to the low number of instances of /ẓ̌/. The rarity of this phoneme may likewise explain why there are no Manichaean data for this phonological value of <ζ>.

/šühə/ < *ḥoγə /šühə/ ‘fingernail’ < *srū-kā- and χοοβο /xwušə/ ‘mother-in-law’ < *hwasrū- (§2.2.4.1). Consequently, *zr- most likely became Bactr. /ž/ in ζονο /žūnə/ ‘Time (as a divinity)’ < *zruwan-.¹²⁸ Another source for Bactr. /š/ is *čy-, e.g. ῃαδο /šādə/ ‘pleased’ < *čyāta-. It therefore seems plausible that also a <ζ> derived from *jy- writes /ž/. One such example would be ζαο(ο) /ζαοι /žāwə/ ‘life’ < *jyāwa-. In addition, there is αζανο /āžānə/ ‘worthy’ < *arjāna-, whose /ž/ is confirmed by its being borrowed with retroflex sibilants; compare Khot. āšana-, TA āšām, TB ašām, and the Niya PN Ašamna (Weber 1985; Dragoni 2023b: 41).¹²⁹

In at least two types of words, one expects <ζ> to stand for /dz/. First, there is reason to believe that <ζ> from *-tVz- spells /dz/, seeing that the outcome of *-tVs- can be written in Bactrian with <σ> /ts/ (cf. supra). Examples of this kind would be πιζαγγο /pidzāŋgə/ ‘knowledge (?)’ < *patizāna-ka- and (α)σπαζαγγο /(ə)spādzāŋgə/ ‘on bended knee, kneeling’ < *spāta-zānu-ka- (cf. the variant σπαδοζαγγο). A second, slightly more common, source of <ζ> /dz/ is the voicing of <σ> /ts/ < *č after a nasal (Sims-Williams 1989b: 234; SFBG: 46).¹³⁰ This type of voicing occurred in lexemes such as πανζασο /pandzāsə/ ‘fifty’ < *pančāsət-.

Voicing after a nasal also affects the <σ> of the appurtenance suffix -σ(ε)ιγο < *čiya-ka-, after v /n/ written as -ζ(ε)ιγο (BD II: 263; cf. also §5.2.2). One example of this kind is ιασηδανζιγο (kb 1) ‘for ever’ < *yāwaitāna-čiya-ka-, a word attested in Manichaean orthography as y’wyd’nžyg (see also under <σ>). The spelling of the latter with <nž> probably points to a palatalisation of *-dz- < *ts- to /dž/ before *i̯ (Sims-Williams 2011a: 173). As nothing proves this phonological process to be a late phenomenon,¹³¹ I cautiously posit /dž/ as a further value of GBactr. <ζ>, an

¹²⁸ ζονο is attested in an unpublished inscription (cf. BPN: no. 484). In ζορ(ρ)ιγο ‘time’ < *zrunaka-, an early metathesis to *zurnaka- must have occurred to account for the Kušāna Bactrian geminate -ρρ- /rr/. Therefore, the <ζ> in this word probably stands for /z/ (Skjærvø 2005: 314; BD II: 213; BPN: no. 160).

¹²⁹ Another example of this sound substitution is TB koško (later koškīye) ‘pit’ ← *κωζγα /kōžgǎ/ (later *κωζγο /kōžgə/) < *kaužda-kā- ‘opening, hole’ (Bernard & Chen 2022). Bernard and Chen reconstruct the Bactrian word as *κωḥκα /kōškǎ/ (later *κωḥκο /kōškə/), but there is no basis in Bactrian phonology for the devoicing of *ždg- > *žg- > *šk- assumed by them; the devoicing can simply be attributed to Tocharian.

¹³⁰ On the development of *č in Bactrian, see in more detail above under <σ>.

¹³¹ Sims-Williams (2011a: 173) adduces the spelling variation between αστιγανσειγι (SK4M 22) and υαστιγανζειγι (SK4B 25f.) ‘native of Hastilgan’ as

assumption which has no repercussions for the loanwords treated in subsequent chapters.

The Manichaean evidence supports GBactr. <ζ> also to stand for a voiced palatal /ž/, although the existing literature (e.g. SFBG) does not always acknowledge its existence or confuses it with /ž/. The Manichaean Bactrian script includes two graphemes <ž> and <j>, the latter being distinguished from the former by adding two dots on top of the character.¹³² The two letters tend to interchange and may have indicated the same phonological value (cf. Sims-Williams 2011a: 172–174).¹³³

Mostly, MBactr. <ž> and <j> are used for the outcome of OIr. *ǰ, i.e. in *žyn'g* ~ ζιναγο 'body' < *ǰinā-kā- or *ǰayanā-kā- (cf. also *žyn'gyng* 'bodily'); *wjyd* 'killed' < *awa-ǰata-; *wjy'n* (alternatively *wžy'n*) 'killing' < *awa-ǰayana-; and *šyǰg* (alternatively *šyžg*) ~ βιζγο 'good' if < *xšijā-ka-. These examples suggest that *ǰ did not develop into **/dz/, parallel to the change of *č to /ts/, but became /ž/ instead.

As corroborating evidence that *ǰ remained palatal in Bactrian, Sims-Williams (2011a: 173) points to the fronting of *a to -i- /i/ next to *ǰ seen in ζιν- /žin/, pret. ζιδο /židə/ 'to strike' < *ǰan-, pret. *ǰata- and ζιπτο /žištə/ 'requested' < *ǰasta- (cf. Khot. *jista*- 'asked'). The easiest way to account for the -i- /i/ in these words is assuming that *a was palatalised in contact with the preceding ζ- /ž/ < *ǰ.¹³⁴ In the case of ζιπτο, also the observed change of

evidence for a non-palatal affricate in Kušāna Bactrian, on the assumption that the <σ> in αστιλγανσειγι stands for /ts/. In my view, the <σ> in -σ(ε)ιγο can also stand for /tš/, in which case Sims-Williams' argument loses its force.

¹³² <ž> has alternatively been transliterated as <j> (e.g. Gershevitch 1983: 274) and <ǰ> (cf. SFBG: 37 with fn. 28).

¹³³ In the following, I leave aside the unclear sequences *βdyj*- 'to require (?)' and *zβr'j* and the loanwords *j'w* 'tide' (~ Sogd. *c'w* 'idem') ← Chin. *cháo* 潮 ((E)MC *drjew*) and *žmbwd'ng* 'wordly, earthly' < *žmbwd** ← Skt. *jambudvīpa*- (with adaptation) + suffix *-ng* < *āna-ka-. I also will not discuss *yyž(y)*- 'to be born'. This verb could be derived from *ā-zāya- (cf. Sogd. *''jy-*) (cf. Sims-Williams 2009a: 262), but the treatment of *ā- to *yy-* would be exceptional and inconsistent with GBactr. αζι- 'to be born' < *ā-zāya- (BD II: 189).

¹³⁴ Gholami (SFBG: 65) writes that *a can also develop to *i in front of *n, giving οανινδο /wanində/ 'victorious' < *wananta- as an example. In the case of ζιν- < *ǰan-, *a would also have been followed by *n. Yet, this is probably not the reason for the vowel change, since, as noted by Sims-Williams (2017c: 111), the development of *a to *i is only regular before the clusters -vδ- and -vζ- or when there is a palatal vowel or consonant before or after *-an-.

*-st- to -ḫτ- /št/ can be attributed to assimilation to the initial palatal (cf. BD II: 212).

Two words whose <ž> / <j> does not derive from *j are *prdyjg* / *prdyžg* ‘orchard’ and *βyžg* ~ βιζαγο ‘sin’. *prdyjg* / *prdyžg* derives from **pari-daiza-ka-* (Sims-Williams 2009a: 264; 2011a: 172). This word thus exhibits a palatalisation of *z to /ž/ after /ē/ < *-ai-; and we may assume that *z was also palatalised to /ž/ after /ĩ/ (cf. supra on *s to /ś/ after /ĩ/). With *βyžg* ~ βιζαγο ‘sin’ < **bazdya-kā-*, one faces the same uncertainty as with *tyšyg** ‘empty’ < **tusya-ka-*. Did *-z(d)y- result directly into /ž/ or did the <y> /i/ from unlauded *a palatalise *z to /ž/?

None of the words attested in Manichaean script tells us whether *z was also palatalised when followed by *ĩ. This assumption would not be too far-fetched, however, given the parallel afforded by the palatalisation of *dz to /dž/ before *ĩ in *y’wyd’nžyg* ‘for ever’ < **yāwaitāna-čiya-ka-*. A change from *zĩ to /žĩ/ may moreover be witnessed by a few Bactrian loanwords into Niya Prakrit (see §4.3.20; §4.3.26).¹³⁵

I conclude that GBactr. <ζ> could represent five phonological values. When <ζ> serves to write the outcome of OIr. *z (< PIr. *j) in an unconditioned environment, we can be confident that it stands for /z/. Secondly, <ζ> can be assumed to represent a voiced palatal /ž/ if from OIr. *j, conceivably also from a palatalised *z. A third possibility is that <ζ> depicts /ž/, namely when <ζ> derives from OIr. *ž or a cluster whose voiceless counterpart is known to yield /š/. In addition, <ζ> may spell the affricate /dz/, whose sources are syncope of *-tVz- and voicing of /ts/ after a nasal. Finally, there may also have been a palatalised affricate /dž/, resulting from palatalisation of /dz/ in front of *ĩ.

¹³⁵ If TA/TB *senik* ‘care, pledge’ is a direct loan from Bactr. *ζηνιγο (attested ζιιιγο) < **zainiya-kā-* (cf. §3.2.1 s.v. *jheniga*), this suggests that *ζηνιγο stands for /zēnīgə/ and thus that /ē/ < *-ai- did not palatalise a preceding *z. For the suggestion that Niya *jheniga* is a loan from Bactr. *ζηνιγο, it may not matter much if the <ζ> in *ζηνιγο stands for /z/ or /ž/.

TABLE 2 Bactrian phonology

The following table gives a synopsis of the phonological value(s) of the letters of the Graeco-Bactrian alphabet, each value being exemplified by at least one word. The values assumed for β, γ, δ, ζ, σ, and ϐ are justified in §§2.2.3–2.2.4. Letters only used as numbers, i.e. ϑ' 6, ξ' 60, and ϕ' 90, have been left out.

LETTER	VALUE	EXAMPLE(S)
α	/a/ /ā/ /ə/	ανδαρζο /andarzə/ 'instruction' < *han-darza- αζαδο /āzādə/ 'noble' < *āzāta- αχβονο /əxšunə/ 'calendar/regnal year', earlier χβονο ναβιχτο (νιβιχτο; νοβιχτο) /nəvixtə/ 'written' < *ni- pixšta-
β	/v/ /b/	βαγο /vəgə/ 'lord, god' < *baga- αβαρ- /āvar/ 'to bring' < *ā-bara- αβ(β)ο /āvə/ 'water' < *āp- (MBactr. 'β) αβαβγο /avāvɡə/ 'waterless' < *apāpa-ka- βαμδδιγο /bāmddigə/ 'loan' < *apāmita-ka- σταμβο /stambə/ 'injury, ill-treatment' < *stambV-
γ	/g/ /ɣ/	μαρηγο /marēgə/ 'slave, servant' < *maryaka- γαμβο /ɣambə/ 'damage' < *gamba- λρωγο /lrōgə/ 'falsehood' < *drauga-
γγ	/ŋg/	ιωναγγο /yōnāŋgə/ 'Greek [adj.]' < *yaun-āna-ka-
δ	/d/	δαρο /dārə/ 'thither (to you)' < *aitaθra πιδο /pidə/ 'father' < nom.sg. *pitā ← *pitar-
ε	/e/ /ə/	ιεσ- /yes/ 'to take (?)' < *ā-yasa- εβκαρ- (αβκαρ-) /əškar/ 'to chase' < *škara-
ει	/i/	λαδειγο /lādīgə/ 'righteous' < *dāt-iya-ka-
ζ	/z/ /ž/ /ž/ /dz/ /dž/	ζαμιγο /zamīgə/ 'land' < *zamī-kā- αζο /azə/ 'I' < *azam ζινο /žinə/ 'woman' < *jani- βιζαγο /vižagə/ 'sin' < *bazdya-kā- (MBactr. βγζg) ζονο /žūnə/ 'Time (a divinity)' < *zruwa(n)- αζανο /āžānə/ 'worthy' < *arjyāna- πιζαγγο /pidzāŋgə/ 'knowledge (?)' < *pati-zāna-ka- πανζασο /pandzāsə/ 'fifty' < *pančāsat- ιασηδανζιγο /yāwēdāndžīgə/ 'for ever' < yāwaitāna- čiya-ka- (MBactr. y'wyd'nžyg)
η	/ē/	ηβ- /ēv/ 'to dispute' < *āpaya- σηλ- /wēl/ 'to lead' < *wādaya-

		<p>σιδο /t̥sídə/ 'which' < *čim + *uti (MBactr. 'cyd) αλσιχτογο /alt̥sɪxtəgə/ 'inclined, desirous' < *hadā- čixšta-ka- πησινδ- /pēt̥sɪnd/ 'to approve' < *pati-sandaya-</p>
τ	/t/	<p>ταβο /tāvə/ 'seal' < *tāpa- αματο /āmatə/ 'broken' < *āmašta-</p>
υ	/h/	<p>υιρζ- /hirz/ 'to release' < *h̥r̥z- (MBactr. h̥r̥z-) λαυ- /lah/ 'to give' < *daθa- (MBactr. lh-)</p>
φ	/f/	<p>φριο /friyə/ 'dear' < *friya- σοφορο /tsufursə/ 'forty' < *čaθw̥rsat-</p>
χ	/x/	<p>χανο /xānə/ 'house' < *xānā- ταχμο /taxmə/ 'stream' < *taxman-</p>
ω	/ō/	<p>ωζο /ōzə/ 'power' < *aujah- ρωσο /rōtsə/ 'day' < *raučah-</p>
β	/š/	<p>βαυρο /šahrə/ 'city' < *xšaθra- οαβο /wāšə/ 'hay' < *wāstra- (MBactr. wš)</p>