MIDDLE INDO-ARYAN "ASPIRATE" CLUSTERS REVISITED[1]

Hans Henrich Hock

The issue of the fate of Sanskrit clusters with sibilant + stop (whether oral or nasal) and with [h] + sonorant has been revived through a paper by Palaschke & Dressler (1999). Focusing on the developments of Sanskrit sibilant + stop clusters (see e.g. (1)), they propose a two-step process through which these become Middle Indo-Aryan geminated stops with postaspiration.

(1) Skt.   asti   >   MIAr.   att^{h}i   'is'

The first step in the development (see (2) below), which they consider a "natural process", namely a lenition, backgrounding, or weakening, could have resulted in preaspirated stops. However, preaspiration is "prevented [...] in terms of system adequacy" by a "paradigmatic prelexical process, within the framework of Natural Phonology [... which] fits the universal scale of naturalness presented inter alia in Hurch (1988: 61–63)". See example (3).

(2) s ≤ h

(3) C^h > ^{h}C
'Postaspirated consonants are more likely to be expected than preaspirated ones.'

In support of their account they refer to Hurch (1986: 62) who cites examples of a supposedly similar process in the Spanish of Seville, where coda *s* becomes *h*; see (4). According to Palaschke & Dressler (62–63),

> Hurch assumes that a process of metathesis converts preaspiration to postaspiration. In this context, it is important to emphasize that this process is

paradigmatic, i.e. a process restricting the aspirated consonants of a language to postaspirated consonants. Thus as soon as the sibilant is debuccalized, this process automatically moves the aspiration to the plosive release of the following consonant [...] This does not justify an intermediate diachronic step of preaspiration as assumed in Vaux (1992) [for Middle Indo-Aryan].

(4) estos casos   [etʰokʰaso]   'these cases'

Even on the face of it, this argument is problematic, since a fair amount of literature exists to show that preaspiration is more usual than postaspiration in Andalusian Spanish. Thus, in their account of the Middle Indo-Aryan fate of Sanskrit sibilant + (oral) stop clusters, Wetzels & Hermans (1985) cite data such as the ones in (5) from Whitley (1978), where at least the nasals are preaspirated, and where oral stops and fricatives are geminated and, if applicable, devoiced.

(5) Castilian   Andalusian
     estúpiðos   ettupiðo(h)   'stupid'
     laz βotas   laφφotæ(h)   'the boots'
     laz manos   lafimanɔ(h)   'the hands'

Lapesa (1959: 321–322) gives the interactions in (6) between what he calls a "pharyngeal aspiration" reflex of *s* (which he symbolizes as [ɦ]), and following consonants. The outcomes include preaspiration before stops, (voiced) fricatives, and nasals, as well as 'casi' (almost) gemination in forms such as *obippo*.

(6) obispo   obiɦpo, obippo   'bishop'
     las gallinas   laɦ gayinaɦ > laɦ ɦayinaɦ / lax xalinaɦ   'the hens'
     las bolas   laf folaɦ   'the balls'
     tiznar   tiɦnar/tiṇnar   'make grimy'

An even larger variety of outcomes is given in Galindo's survey (1977: 58–59, see also 92 and passim); and for the dialect of Montejicar, which he investigates in full detail, he comes to the conclusions summarized as in (7).

(7) a. The most common outcomes are geminates without aspiration.
     b. Aspiration, if present, is preaspiration and usually voiced.
     c. Voiced fricatives usually stay voiced; they are rarely devoiced.

The overall evidence thus suggests that the Andalusian (etc.) dialects of Spanish offer a large variety of different ways of treating clusters of *h* + consonant. Most of these are different from the postaspiration that Palaschke & Dressler consider

natural and automatic; and preaspiration, which Palaschke & Dressler consider less natural, is by far more common.

The question of how to account for the Middle Indo-Aryan postaspiration in (1) thus cannot be simply answered on a-priori grounds of naturalness or the like, but must be decided on the basis of language-specific evidence.

Wetzels & Hermans (1985)[2] present an argument similar to Palaschke & Dressler's, but in a different framework. Their approach can be summarized as in (8).

(8) a. Basic assumption: *s* has an inherent non-segmental *ʰ* attached to it.
     b. "*s*-deletion" leaves that non-segmental *ʰ* stranded.
     c. The stranded *ʰ* spreads to a neighboring consonant.
     d. The outcome, by "convention", is a geminate aspirated consonant with, in the case of Pali, postaspiration.

One problem with this account is the assumption that *s* contains an inherent non-segmental *ʰ*. True, *s* frequently changes to an *h*-sound; but so do other fricatives (such as Span. *f* > *h* in *filius* > *hijo*). Moreover, languages such as Burmese have aspirated *sʰ*, which under Wetzels & Hermans's account would have to be considered an *s* with **double non-segmental** *ʰ*. Most important, the assumption that the outcome of *s*-weakening is always non-segmental is problematic. True, in Greek there is ample evidence that at a certain stage, the spiritus asper, normally transliterated as *h*, is non-segmental or "does not count" on the segmental tier,[3] see (9). Example (9a) shows that contraction takes place across *ʰ* and that *ʰ* fuses with a voiceless stop preceding the first of the two contracting vowels; (9b) demonstrates that word-internally, the outcome of such fusion is a single segment, and not a geminate, as would be expected if *ʰ* were a full segment. On the latter count, see (10), where a similar fusion in Sanskrit yields an aspirated geminate

---

2   Not cited by Palaschke & Dressler.

3   At an earlier stage, however, it apparently did, to judge by the fact that *sN* and *Ns* clusters changed to geminate *nn*, presumably via *hN/Nh* > *ṆṆ*; see Hock 2004 with discussion and references. – Interestingly, a similar change from earlier segmental to later non-segmental status seems to have taken place in the history of Indo-Aryan, from Sanskrit to Modern Hindi. As example (10) illustrates, Sanskrit *h* was segmental and hence required compensatory gemination when fusing with a preceding stop. Hindi does not have such gemination, as seen in examples like *kabhī* 'whenever' < *kab* + *hī*; and like Greek it offers examples of vowel contraction across *h*, as in *kahiṁ* 'wherever' < *kahāṁ hī*. (The parameters under which these developments take place in Hindi still need to be fully worked out.)

(with aspiration on the release phase), showing that *h* here is a full segment, requiring the fused consonant to preserve the mora count of the input structure.[4]

(9) a. kaì ho   > khō   'and the ...'

    b. kat(a) hédra > kathédra, not katthédra*   'seat'

       ap(o) hédra   > áphedros, not ápphedros*   'sitting aside; menses'

(10)   CVC CV   CVC C V

         t a d h i   > t a d dh i   'for/because that'

          μ μ   μ μ

Similarly, in the case of Andalusian there is evidence that the weakening of final *s* in the first instance yielded a segmental *h*. In part the evidence consists of the pervasive gemination of following consonants; in part it comes from the fact that the loss of absolute-final *h* commonly results in compensatory lengthening of the preceding vowel, as in (11) (Alarcos Llorach 1965: 280; see also Galindo 1977: 122–137, with acoustic measurements).

(11) los ves [lɔs βɛs] > lɔh βɛh   > [lɔφφɛ:]   'you see them'

The question of whether the Middle Indo-Aryan development in (1) involved an intermediate stage with non-segmental *ʰ* or with segmental *h*, thus, cannot be answered on a-priori grounds, but must be decided on the basis of language-specific evidence.

   In a paper that Palaschke & Dressler argue against, Vaux (1992) makes the claims summarized in (12).

(12) a. Consonants are geminated after Skt. /s/.

    b. /s/ is debuccalized to /h/.

    c. "Since /h/ lacks all supralaryngeal features, it copies the supralaryngeal features from the consonant. The resulting preaspirated geminate *htt* ends up as a postaspirated geminate *TTh* according to Bartholomae's Law."

---

4 Skt. *h* is phonetically voiced [ɦ]; the aspiration of voiced aspirates, too, is voiced [ɦ]; and there is reason to believe that the *h* resulting from coda weakening of *s* in pre-Middle Indo-Aryan was voiced [ɦ] next to voiced segments. In other contexts, *h* and aspiration are voiceless. These distinctions are ignored in this paper, except where voicing is relevant.

Palaschke & Dressler object that Bartholomae's Law (see the traditional formulation in (13a)) applied to pre-Sanskrit clusters of aspirated stop + obstruent and that "there are no clusters like *hT-* in OIA".

(13) a. *bhudh-to- > Skt. buddha   'awakened'
    b. *bʰudʰto   >   *bʰudtʰo
                    >   bʰuddʰo   (with assimilation of [t] to voiced [ʰ])

Now, it is true that the prehistoric process called Bartholomae's Law did not operate on segmental *h* + consonant clusters, and it is also true that Sanskrit had no clusters with segmental *h* + stop. However, this does not mean that there could not have been such clusters in the transition from Old to Middle Indo-Aryan and these could have undergone what I take to be the motivation behind Bartholomae's Law, namely to move aspiration – or segmental *h* for that matter – from the coda to the following onset (see the more appropriate reformulation in (13b)). This is, in fact, the spirit behind Palaschke & Dressler's tendency in (3), repeated here for convenience.

(3) $C^h > {}^hC$
'Postaspirated consonants are more likely to be expected than preaspirated ones.'

In fact, such a development is precisely what I proposed in Hock 1991 for the fate of Sanskrit *s* + nasal clusters in Prakrit. The developments I postulated and their justification can be summarized as follows.

First, coda *s* was weakened to *h*[5] – a common and natural development; see (14a,b) first stage. Second, since *h* (and aspiration) crosslinguistically is preferred in onset position, rather than in coda, the *h* that resulted from *s* was metathesized; see (14a,b) second stage. Independent evidence for this metathesis is found in original Sanskrit clusters of *h* plus sonorant,[6] which also underwent the change; see (14c). In the case of oral stops, the preexistence of aspirated voiceless stops made it difficult to maintain a distinction between the resulting clusters of stop + *h* and original aspirated stops (see Hock 1986 with further examples); the *h*

---

5   The *h* of (14c) was clearly segmental, since in poetic metrics, the *h* "makes position", i.e. creates a heavy syllable and thus is mora-bearing. Given the parallel second-stage development of (14c) and (14a), it is legitimate to assume segmental status for intermediate *ahmi* and, by extension, for intermediate **ahti*, as well.

6   In the following I ignore the development of Skt. *hy*, *hv* [hw] or [hᵝ] clusters, which develop into *jjh*, *bbh* in all of Middle Indo-Aryan. (What might be noted, however, is that in both cases the outcome is a geminate, reflecting the fact that *h* was segmental in Sanskrit.)

therefore fused with the preceding stop, and as in the case of *tadd^hi* (see (10), repeated for convenience), the result was a geminate aspirate, with preservation of mora count; see (15). No such development took place in the case of nasal + *h* clusters, since there were no preexisting aspirated nasals. As argued in Hock 1991, "The *h* of such Prakrit forms is solidly segmental."

(14) a. asmi   >   *ahmi   >   amhi   'am'
     b. asti   >   *ahti   >   *athi   'is'
     c. brahman =   brahma-   >   bamha-   'Brahman'

(10)   CVC CV   CVC C V

        t a d h i   > t a d d^h i   'for/because that'
          μ μ   μ μ

(15)   *athi   > att^hi   (with fusion of *t* and *h* and mora preservation)

What I did not address in my 1991 paper is the fate of clusters of this type in initial position; see the examples in (16). If we accept my 1991 arguments, these examples can be accounted for by observing that Pali (and Middle Indo-Aryan in general) has a constraint against initial consonant clusters or geminates.[7] Hence the ultimate outcome in (16b) can be explained as resulting from degemination; and those in (16a, c) as employing anaptyxis to avoid violating the constraint.

(16) a. snā- > *nhā-   > n^ahā-   'bathe' (with anaptyxis)
     b. stana > *thana > *tt^hana > t^hana   'breast' (with degemination)
     c. hrada > *rhada   > r^ahada   'lake, pond' (with anaptyxis)

As it turns out, one of the crucial assumptions of my 1991 analysis is called into question by the following facts, noted by von Hinüber (1986: 116, 117). First, Prakrit writes initial <nh> in forms like *nhā-* 'bathe', from Skt. *snā-*. Second, the evidence of poetic metrics shows this <nh> to be a single consonant, not a cluster. Third, the same holds true even for orthographic <nah> in early Pali poetry. Von Hinüber therefore concludes that the interpretation of orthographic <n(a)h> (and the like) must be as in (17a), entirely parallel to the outcomes of Sanskrit *s* + oral stop clusters (17b).

---

7   Some exceptions occur, especially in western dialects which tend to preserve initial stop + *r* clusters. But these exceptions do not affect the present argument.

(17)   Initial position   Medial position
   a.   nh   nnh
   b. t^h tt^h

According to von Hinüber, it is only in later Pali texts that poetic metrics shows initial orthographic <nah-> to be phonetically/phonologically *nah-*, with a genuine vowel *a*.

Von Hinüber's analysis of the Middle Indo-Aryan situation thus appears to place serious, perhaps insurmountable obstacles in the path of my 1991 analysis.

In the remainder of this paper I present arguments to show that von Hinüber's analysis is not likely to cover the entire range of Middle Indo-Aryan and that beside more western varieties which had the distribution in (17), there was another, eastern group of dialects that had the distribution in (17').

(17')   Initial position   Medial position
   a.   nah   nh
   b. t^h tt^h

The first argument concerns the fact that, as von Hinüber observes, later Pali texts have phonetic *nah-* for orthographic <nah>, as attested by poetic metrics. The question which von Hinüber does not address must be how this *nah-* could have arisen, if the earlier stage only had aspirated n^h. One could possibly operate with the assumption that the feature of aspiration, ^{h}, was segmentalized to *h* and that this, in turn required anaptyxis in initial position. While segmentalization of this type seems to have happened in the case of Icelandic preaspiration (see (18)), it would be a highly unusual development for postaspiration. I am not aware of any attested case of segmentalization of postaspiration anywhere.[8] An analysis in terms of an earlier cluster *n* + *h*, plus anaptyxis in initial position, would avoid this problematic assumption.

(18) Icel. ætta [ay^{h}tta]   >   [ayh$^{(t)}$ta]   'owned'

---

8   In the discussion at the SALA meeting, Elena Bashir suggested that such a change seems to be found in varieties of Pakistani Urdu, where *Bhārat*, the official name of the Republic of India, is pronounced as [bahārat]. But as she also noted, this development is limited to voiced aspirates; voiceless aspirates do not exhibit the change. Now, there is a strong Panjabi and a lesser Kashmiri substrate in Pakistani Urdu; neither Panjabi nor Kashmiri has voiced aspirates; but both of them have voiceless aspirates. This makes it likely that the pronunciation [bahārat] reflects an attempt of Panjabi (and Kashmiri-)dominant speakers to pronounce [b^h] by rendering it as a sequence of [b] + [h], similar to, say, the Indonesian rendition of Skt. *bhāṣā* 'language' as *bahasa* (as in *Bahasa Indonesia*).

On the other hand, the geminate sonorant aspirates in von Hinüber's corpus (simplified in initial position) can be explained as secondary developments from earlier clusters of sonorant + *h* by extending the fusion of *h* with preceding oral stops, for which see again the discussion of example (15b). Similar extensions of processes that are natural in one context to contexts in which they are marked can be observed elsewhere; see e.g. (19a) vs. (19b) (from Hock 1985).

(19) a. Common Iranian metathesis

>caxra   >   caxr >   carx   'wheel'
>(to avoid sonority hierarchy violation)

→   b. Ossetic metathesis

>(*)tray- >   rtä- (> ärtäk)   'three'
>(in initial position, violating the sonority hierarchy, subsequently remedied by prothesis)

Second, examination of the Aśokan inscriptions[9] reveals that there is indeed a difference between the western inscriptions of Girnar and the more eastern inscriptions of Dhauli, Jaugada, Kalsi, etc. While the inscriptions do not offer any examples of relevant clusters – or aspirates – in initial position, they do suggest a difference medially. Where western Girnar consistently has *mh*, which (given the nature of the writing system) could be interpreted as a geminate aspirate *mmʰ*, the eastern inscriptions have forms containing an oral stop ± preceding nasal; see example (20). (The alternation between *aṁbh* and *ābh* follows from general Middle Indo-Aryan tendencies to treat sequences of this sort as equivalent.)

(20)	Sanskrit	Western	Eastern	
		Girnar	Kalsi	Dhauli/Jaugada
	brāhmaṇa	b(r)āmhaṇa	baṁbhana	bābhana
			once: baṁhmane	once: baṁbhana

If the starting point for the eastern outcomes had been an aspirated *mmʰ*, it would be difficult to account for the eastern forms, since unaspirated *mm* (as in *dhamma* < Skt. *dharma*) remains unchanged. If however we assume a cluster *mh*, then we can account for the eastern forms as involving consonant epenthesis between nasal and (non-nasal) *h* (21a), comparable to the epenthesis between nasal and (non-nasal) liquid in (21b).

9   Data from Bloch 1950 and Hultzsch 1924.

(21) a. bamhana   >   *bam^bhana >   bamb^hana (with fusion of stop and *h*)

    b. Skt. amla   >   *am^bla   >   Pali ambila (with anaptyxis)

      Skt. tāmra   >   *tām^bra   >   Pali tamba (with cluster reduction)

As it turns out, the comparative Modern Indo-Aryan evidence assembled by Turner (1962–1966)[10] supports a similar east : west distinction, if we make allowances for some cross-regional borrowing. (The relative dearth of Sindhi data makes it difficult to classify the language.)

As the representative data in the Appendix show, eastern varieties have a much higher incidence of oral stop insertion in original *s* or *h* + nasal clusters, such as Old Bangla *bāmbhaṇa* or Oriya *āmbhe*. In addition, they also regularly show anaptyxis in initial position.[11] In addition, note Hindi forms such as *anhānā* 'bathe', *hanāṇ* 'bath', and Avadhi *hanāb*, with metatheses that presuppose earlier *nah-*.

Western dialects, in contrast, have only a few instances of oral stop insertion, and these could be attributed to borrowing. (Note for instance that Mar. *bābhurḍā* is not the ordinary word for 'brahmin', but has a specialized, pejorative meaning, possibly reflecting negative attitudes to "eastern" brahmins.) Moreover, (Old) Gujarati, Marathi, and Panjabi have forms in which initial orthographic <nh> is not broken up by anaptyxis. On this count, Hindi forms a transition area, with <nh> in western varieties such as Braj Bhāṣā and <nah> in Standard Hindi and in eastern varieties.[12] As for the later Gujarati and alternative Marathi forms of the type *nahāṇ*, *nāhṇ*, *nāhāṇ* from Skt. *snā-*, these can be explained as resulting from what Bloch (1914 (1970): 124) calls aspirate transfer; see also Masica (1991: 120), who suggests that this may really involve a spread of the breathy voice or murmur of the original aspiration into the following vowel, producing murmured vowels.

Most of the western languages cannot be expected to preserve evidence for the original geminate aspirated sonorants in medial position, because of the widespread simplification of Middle Indo-Aryan geminates (with compensatory lengthening). However, Panjabi offers a few forms with geminates such as Eastern Panjabi *hummh* from Skt. *ūṣman* 'heat, steam', and so does Pahari, as in West

---

10 Supplemented by McGregor 1993, Rhys Davids & Stede 1921–1925.

11 Bangla *nāoyā* 'to bathe' could go back to earlier *nahā-* because of the widespread loss of *h* in non-initial position; see also the Bangla outcomes of medial *nh* etc.

12 What complicates matters is that initial aspirated nasals in western Hindi may result from secondary changes, as in *mhaĩs* 'water buffalo' < Skt. *mahiṣa*. (Interestingly, the eastern and standard Hindi counterpart is *bhaĩs*, presumably because of the lack of initial *mh* or *ṃh* of independent origin – or as the result of a "replay" of oral-stop epenthesis in the cluster *ṃh*, hence *mbh*, followed by cluster simplification.)

Pahari *cìnn*[13] from Skt. *cihna* 'sign'. In this context the Kalsi form *baṁhmane*, occurring once in the Aśokan inscriptions (see (20) above), may be relevant, since the best way to interpret the highly unusual orthographic representation <ṁhm> would be as an aspirated geminate *mm^h*.[14] Note that although Kalsi generally shows eastern forms (e.g. *baṁbhana* with oral stop epenthesis), its location in the present-day Pahari area, at the border between east and west, makes it a likely candidate for exhibiting both eastern and western features.

The evidence of Middle and Modern Indo-Aryan thus does support a division between eastern dialects with sonorant + *h* clusters and western dialects with originally geminate, aspirate sonorants; see (22). And as mentioned earlier, the aspirated sonorant pattern of the western dialects can be explained as an extension of the development of Skt. *s* + oral stop clusters.

(22)   Dialectal division of Middle and Modern Indo-Aryan

	Eastern	Western
Initial	nah	n^h
Medial	nh	nn^h

The only remaining problem might be that in literary Middle Indo-Aryan, the "western" pattern is found not only in clearly western Maharashtri, but also in early Pali and Ardha-Magadhi – even though these languages are commonly considered eastern in origin and, in fact, contain a fair amount of "easternisms", especially in their earliest stages. However, Pali and Ardha-Magadhi are dialectally composite and, in their overall structure, closer to western Aśokan Girnar than to the east (see e.g. von Hinüber 1986: 38–39 in reference to Pali). The presence of more "eastern" features in later Pali (as in *nah-* with full vowel; see above) can possibly be attributed to the influence of Sinhala which, to judge by the inserted oral stop in *baṁba* 'brahman' and *uṁba* 'you (pl.)', was affiliated with the eastern dialects as regards the development of Sanskrit *s* or *h* + sonorant clusters. The evidence of literary Middle Indo-Aryan, thus, does not seem to be in conflict with the hypothesis of this paper.

Those familiar with the Aśokan inscriptions may have noted that the developments accounted for so far do not cover the complete range of data.

---

13 The tone indicated by the grave accent suggests that the form goes back to earlier *cinnh*, with aspirated geminate nasal.

14 Curiously, this form and its phonetic interpretation seem to have been generally neglected in Indo-Aryan linguistics.

Forms of the type (23) appear to be exceptions to the ordinary developments of *s* or *h* plus *m*.

(23) Exceptional developments involving m … n

Sanskrit	asmān	yuṣmān	tasmād	-asmin
Dhauli etc.	aphe	tuphe		-as(s)i
Kalsi			taphā	-a(s)si
Expected	amhe	tumhe	tamhā	-amhi
	ambhe	tumbhe	tambhā	-ambhi

Von Hinüber (with references) provides a likely explanation, namely dissimilation of *m* in the context *sm … n*, whence e.g. *asmān > aspān*, which by the usual changes will result in *a(p)pha-*.

What is problematic is that loc. sg. *-asmin* yields *-a(s)si* in the same dialects, something which von Hinüber does not seem to account for explicitly. An indirect explanation would suggest itself in his idea of different syllabications (see (24)), which he proposes in order to account for the different developments of Skt. *tm* into *pp* or *tt*. But given what we know about Sanskrit (and its Prakrit descendants) such a variation in syllabication is highly unlikely (the syllable boundary is always in the middle of the cluster, as in (24a)). It is more likely that the solution is to be sought in postulating two different types of dissimilation; see (25). Note that dissimilation is not necessarily a regular process (in fact, Girnar does not show these developments); it is therefore not necessary to assume that the same development must have taken place in the case of *a(p)phe* and *-a(s)si*.

(24) a.   as.mi
      b.   a.smi

(25) a.   sm … n > sp … n   (hence *a(p)phe* etc.)
      b.   sm … n > sv … n   (hence *-a(s)si*, like *aśva > assa*)
      cf.   Pali *bhassanta* 'ending in ashes', *aśmamuṣṭika > * Pali *assamuṭṭhika* 'with a stone in his fist' (v. Hinüber 119)

Forms of the type (23) are not found in the literary Prakrits, except for some apparently not localizable forms recognized by Hemacandra (see Pischel §§312, 313). Interestingly, however, similar forms do occur here and there in the Modern Indo-Aryan languages, in words where sibilant + *m* was originally followed by *n*. Compare the examples in (26), from Turner (1962–1966).

(26) Modern Indo-Aryan counterparts

    a. Skt. *śreṣman 'phlegm'   Bangla *chep*

                                                          Oriya *chepa*

                                                           Nepali *sep*

                                                           Sindhi *sīpho*

    b. Skt. *pakṣman* 'eyelashes'   Maith. *papani*

                                                           Mar. *pāpṇī*

                                                           Panj. *bhapphan*

                                                           Lah. *pippṇī*

                                                           W Pah. *pimphṇi*

As Turner points out in reference to *pakṣman* (s.v.), a large number of other developments appear to have taken place in the prehistory of Modern Indo-Aryan, including contamination with other words. Nevertheless, it is remarkable that outcomes very similar to Aśokan *aphe*, *tuphe*, *taphā* can be observed in Modern Indo-Aryan. Moreover, evidently these do not reflect regular sound changes, but sporadic ones, affecting a particular word in a particular variety of Indo-Aryan, but not others, in other varieties. This is entirely consonant with my earlier claim that the dissimilatory processes in (14) need not have been regular, since dissimilation in general is usually an irregular phenomenon.

    Ignoring the complications in combinations containing the sequence *m ... n*, the findings of this paper can be summarized as follows. The development of Sanskrit *s* + stop or nasal and *h* + sonorant clusters into Middle Indo-Aryan proceeded via a stage with clusters of *h* + stop or sonorant. This *h* was segmental *h* and metathesized from the coda to the following onset. Because of the preexistence of aspirated oral stops, the *h* fused with preceding oral stops. It did so with mora preservation, yielding geminates, which were simplified in initial position. In the case of sonorants + *h*, eastern Middle Indo-Aryan retained the cluster configuration, while western Middle Indo-Aryan extended the *h*-fusion of the stops. To relieve initial sonorant + *h* clusters, the eastern dialects resorted to anaptyxis; moreover, medially they tended to insert an epenthetic oral stop between nasal and *h*. The western dialects simplified the geminate aspirated sonorants in initial position. These developments are summarized in (27).

(27)   Sanskrit   Intermediate Stages   After metathesis (and further developments)

   a. asti   > *ahti   > *athi   > att^hi (with fusion and mora preservation)

   a'. stana   > *htana   > *thana   > $*tt^hana$ > t^hana

   b. asmi   > *ahmi   > *amhi   >
            >   >   > > W amm^hi
            >   >   > > E (*)amb^hi (not in all forms)

   b'. snā   > *hnā   > *nhā   >
            >   >   > > W $nn^hā$ > $n^hā$
            >   >   > > E nahā

   c. brahman   = brahma-   > *bamha   >
            >   >   > > W $bamm^ha$
            >   >   > > E $bamb^ha$ (not in all forms)

   c'. hrada   = hrada   > *rhada   >
            >   >   > > W rr^hada > r^hada
            >   >   > > E rahada

APPENDIX

Sanskrit *h* and *s* + sonorant clusters in Middle and Modern Indo-Aryan (representative examples).

Sanskrit	hrada[15] 'pond'	snā- 'bathe'	snapana[16] 'bathing'	brāhmaṇa 'brahmin'	brāhmī (a fish)	cihna 'sign'
Pali	rahada	n(a)hā-		(brāhmaṇa)[17]		cinha
Prakrit		ṇhāi		bamhaṇa		cinha
Apabhr.				**bambhaṇa**		**ciṁdha**
Bangla		nāoyā		OB **bāmbhaṇa** / bāman		cin
Oriya				**bāmbhuṇa** / **bābhuna** / bāman		cina
Maithili		nahāb		**bābhan** / bāman		cenh
Nepali	raha			bāman / bāhun	bām	cinu
Avadhi		nahāb / hanāb		**bābhan** / bāmhan		cīnh
Hindi		nahānā / anhānā / hanāṇ / *nhānā*		bāmhan / bābhan[18]	**bāb** / bām	cinh
Pahari		WP nihān		Ku bāman		WP *cìnn*
Panjabi		*nhāuṇā*	nauṇ	bāmhan	bām	

---

15 For ease of exposition, the Sanskrit input forms given here in many cases stand for a whole "family" of related forms. For instance, *snā-* 'bathe' covers both the verbal root *snā-* and its various verbal forms, as well as nominal forms such as *snāna* 'bath'.

16 Obviously related to *snā-*, but with an irregular short root vowel.

17 A borrowing from Sanskrit and/or an archaic/archaizing form comparable to Girnar *b(r)āmhaṇa*?

18 From Pathak 1989.

Gujarati	nāhvū nahāṇ	OG *nhavaṇa* MG nahāvaṇ	bāmaṇ	bām	OG **ciṁdha** cin
Marathi	nāhṇē nhāṇ nǎhāṇ		bāmaṇ **bābhurḍā**[19]	bām	cinh
Sindhi			**b̲ābhaṇu**		cinhu

	tṛṣṇā 'thirst'	kṛṣṇa 'Krishna'	ūṣman- 'heat'	grīṣma 'summer'	asmad 'we' (obl.)	yuṣmad 'you' (obl.)
Sanskrit						
Pali	taṇhā	kaṇha		gimha	amhē	tumhē
Prakrit	taṇhā	kaṇha	umha	gimha	amhē	tumhē
Apabhr.				**gimbha**		
Bangla		kāna	umāna		OB ambhe āmhe āmi	OB tumhe tumi
Oriya		kānha	**umbhāi,** u(h)māi		āmbhe āme	tumbhi
Maithili		kānha			ham	
Nepali					hāmro	timi
Avadhi		**kādhaiyā**			ham	OAv tumhahiṁ
Hindi		kānh			ham	tum tumhe
Pahari					Jaun ām	Ku tum
Panjabi		kānh	EP *hummh*			
Gujarati					OG āma ame	OG tumhe tame
Marathi	tānh tahān	kānhū		gīhm gimh-	āmhī	tuhmī
Sindhi		kāno		ghima		

---

19 Turner: 'term of reproach for a Brahman'.

REFERENCES

Alarcos-Llorach, Emilio 1965. *Phonología española*, 4th revised edn. Madrid: Editorial Gredos.

Bloch, Jules 1914. *La formation de la langue marathe*. Engl. transl. 1970, *The formation of the Marāṭhī language*, by Dev Raj Chanana. Delhi/Patna/Varanasi: Motilal Banarsidass.

Bloch, Jules 1950. *Les inscriptions d'Asoka*. Paris: Les Belles Lettres.

Galindo, Miguel Angel 1997. *A study on eastern Andalusian dialectology: A phonemic description of the syllable-final consonant system of the dialect spoken in Montejicar (Granada)*. UIUC PhD dissertation.

von Hinüber, Oskar 1986. *Das ältere Mittelindisch im Überblick*. 2nd edn 2001. Wien: Verlag der Österreichischen Akademie der Wissenschaften. (Page references are to the first edition.)

Hock, Hans Henrich 1985. Regular metathesis. *Linguistics* 23: 529–546.

Hock, Hans Henrich 1986. *Principles of historical linguistics*. Berlin: Mouton de Gruyter.

Hock, Hans Henrich 1991. Dialects, diglossia, and diachronic phonology in early Indo-Aryan. In: W.G. Boltz & M.C. Shapiro (eds), *Studies in the historical phonology of Asian languages*: 119–159. Amsterdam: Benjamins.

Hock, Hans Henrich 2004. Fish, push, and Greek R + y clusters: A return to Danielson 1903. In: Karlene Jones-Bley et al. (eds), *Proceedings of the Fifteenth UCLA Indo-European Conference, Los Angeles, November 7–8, 2003* (Journal of Indo-European Studies Monograph Series, 49): 155–167. Washington DC: Institute for the Study of Man.

Hultzsch, E. 1924. *Inscriptions of Aśoka* (Corpus Inscriptionum Indicarum, 1). Repr. 1991, New Delhi: Archaeological Survey of India.

Hurch, Bernhard 1988. *Über Aspiration: Ein Kapitel aus der Natürlichen Phonologie*. Tübingen: Narr.

Lapesa, Rafael 1959. *História de la lengua española*. 4th edn. Madrid: Escelicer.

McGregor, R.S. 1993. *The Oxford Hindi-English dictionary*. Oxford: OUP.

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

Palaschke, Mario & Wolfgang U. Dressler 1999. Middle Indic aspirate formation: Syllable structure vs. natural processes. *The Yearbook of South Asian Languages and Linguistics* 1999: 53–68.

Pathak, R.C. 1989. *Bhargava's standard illustrated dictionary of the Hindi language: Hindi-English edition*. 5th edn. Varanasi: Bhargava Book Depot.

Rhys Davids, T.W. & William Stede (eds) 1921–1925. The Pali Text Society's Pali–English dictionary. London: Luzac.

Turner, Ralph L. 1962–1966. *A comparative dictionary of the Indo-Aryan languages*. Oxford: OUP.

Vaux, Bert 1992. Gemination and syllabic integrity in Sanskrit. *Journal of Indo-European Studies* 20: 283–303.

Wetzels, W. Leo & Ben Hermans 1985. Aspirated geminates in Pāli. In: Hans Bennis & Frits Beukema (eds), *Linguistics in the Netherlands*: 213–223. Dordrecht: Foris.

Whitley, M. Stanley 1978. *Generative phonology workbook*. Madison: The University of Wisconsin Press.