The formation of modern Hindi as demonstrated in early ‘Hindi’ dictionaries

BY STUART MCGREGOR

2000 GONDA LECTURE
Eighth Gonda lecture, held on 23 November 2000 on the premises of the Royal Netherlands Academy of Arts and Sciences
The formation of modern Hindi as demonstrated in early ‘Hindi’ dictionaries

BY STUART MCGREGOR

ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCES
Amsterdam, 2001
Modern Hindi has been widely seen as a new language that came into existence at the beginning of the nineteenth century, the product of new political conditions in north India: a language to be distinguished historically from forms of north Indian language current before it. Research on these languages provides various illustrations, however, of the extent to which modern Hindi shares features of lexical repertoire as well as grammatical structure with them, and of the depth and strength of its roots in pre-existing language usage.

It is my intention today to look at lexical and grammatical connections between older language and modern Hindi, making use of evidence from early dictionaries. In the course of working in Hindi lexicography over a considerable number of years I came to realise the great interest of the early dictionaries, not only as documenting the history of Hindi lexicography, but also for the information relevant to my present subject that they proved to contain. The earliest of the dictionaries, dating from around 1700, appeared to be of special interest, while verse thesauruses in vernacular language dating from over a century earlier were also relevant. I will be concentrating on four dictionaries today. It is not only in the contents of these dictionaries that evidence bearing on the formation of modern Hindi is to be found, but also in the prefaces attached to three of the dictionaries by their authors.

These dictionaries are, for obvious reasons, not themselves of modern Hindi. They are of types of language related to modern Hindi: first, the mixed language of Delhi that spread quickly during the early Muslim period as a lingua franca, and came eventually to be known as Hindustani; and secondly Brajbhāṣā, the speech of the Braj district north of Agra, that spread across north India as a literary language, co-existing with the lingua franca from the sixteenth century.

In my title it has been convenient to refer to these dictionaries collectively as 'Hindi' dictionaries, and I have placed the word Hindi in inverted commas. By that I mean that their language bears one or another close linguistic relationship to modern Hindi. The scope for ambiguity in using the word Hindi is almost unending; it rests of course on the meaning of the Persian word Hindi, which is simply 'Indian', so that from an Indo-Muslim perspective the term could apply,
first to the vernacular language of Delhi, but also to Sanskrit, and possibly Apabhraṃśa; much later to Urdu; and later still to modern Hindi; while used as an adjective, the word ‘Hindi’ can refer today, with even more confusing implications, to any or all of the regional or local languages of the Hindi language area. I have tried to make the different senses in which I use the word today sufficiently clear.

Of my four early dictionaries, none, with the possible exception of the materials attached as a word list to Joan Joshua Ketelaar’s grammar of Hindustani, is at all well known. They are:

1. The verse thesauruses Mān-mañjari and Anekārth-mañjari of the late sixteenth century Brajbhāṣā poet, Nanddās (discussed here as a single item). The significance of these works for us is that they indicate the coming into use of considerable numbers of Sanskritic loanwords in Brajbhāṣā poetry in the late sixteenth century, and they illustrate a process of diffusion of knowledge of them under way among the auditors of poetry.

2. The Thesaurus Linguae Indianae of François Marie of Tours, a Capuchin friar based at Surat at the turn of the seventeenth century. The Thesaurus was completed in the year 1703. This large and little known work is a dictionary of Latin and French to Hindustani – Hindustani recorded in Devanagari script, and provided with a roman script transliteration. It has an interesting history, for the manuscript, after being taken from Surat to Rome by its author and deposited in the library of the Congregatio de Propaganda Fide in 1704, was sent on loan to Paris in the 1780s at the request of the Iranianist Anquetil Duperron. A copy made by Anquetil remains available today, although François Marie’s original has been lost. The Devanagari script of this copy, the Sanskritic component of its essentially Hindustani vocabulary, and not least the composition of the original in Surat, give the copy some importance in tracing the antecedents of nineteenth century Hindi.

3. Ketelaar’s grammar, or manual of instruction, in Hindustani and Persian, Instructie over Onderwijsing der Hindoustanse en Persiaanse Talen (1698). The Leiden scholar J.Ph. Vogel discussed this unpublished grammar, of which he had discovered a
manuscript in the Rijksarchief at The Hague, in two articles in 1936 and 1941, with some discussion also of materials from the word list attached to the grammar. Vogel's work on the grammar was an offshoot from his historical study of the journey from Surat to Lahore that Ketelaar made during the years 1711–12 as ambassador of the Dutch East India Company to Jahangir. Ketelaar's language material, especially the grammatical part of it, has been discussed more recently by T. K. Bhatia (1987, 1992) and H.W. Bodewitz (1996).

4. The dictionary materials in a Persian work named *Tuḥfasu’l-Hind 'A Present from India'. *Tuḥfasu’l-Hind is a study of Brajbhāṣā grammar and poetry, probably made for the Mughal prince Azamshah around 1675. The introduction and first part were translated by M. Ziauddin in 1935, with an introduction by Sunitikumar Chatterji. *Tuḥfasu’l-Hind has the interest that it was made use of by Sir William Jones (though not for its language), in Jones's paper of 1784, 'On the musical modes of the Hindoos'. *Tuḥfasu’l-Hind, like Ketelaar's manual, has a lexicographical appendix. This amounts to a Brajbhāṣā-Persian dictionary of about 3,000 words, with some phonetic explanations and cultural commentary, compiled for the benefit of members of Indo-Muslim communities.

**NANDDĀS'S THESAURUSES, AND THE NEW VIABILITY OF SANSKRITIC VOCABULARY (16TH CENTURY)**

Nanddās's thesauruses continue the long tradition of Sanskrit lexicography that descends from Amara, its founder, but in a new spirit. The Sanskrit dictionaries (kośas) had been intended primarily for poets' study as an aid to composition, and their role was seen as crucial in the maintenance and transmission of literary culture. Amara could be seen as equal to Pāṇini in prestige, and his *Nāmaḷīṅgaṇī-śāsana or Amarakoṣa*, according to a Sanskrit epigram, was the 'father of literature'. More commentaries were made on *Amarakoṣa* than on any other piece of Indian literature, and its influence in underpinning the cultivation of Sanskrit continued to the nineteenth century. It acquired a new dimension of interest for Europeans in Colebrooke's English translation of 1808, made as a basis for the first Sanskrit-English dictionary.

In the sixteenth century, the new religious consciousness of bhakti was being
given literary form by Kṛṣṇa vernacular poets. This new devotionalism was in part an assertion of cultural identity within a now mixed society. Kṛṣṇa traditions passed down in Sanskrit would now be selectively modernised in Brajbhāṣā, and with these traditions, elements of the vocabulary in which they had been enshrined would acquire an enhanced viability.

Nanddās is thought to be the first poet to have adapted the Amara-kośa into Brajbhāṣā. He was at once respectful of a venerable tradition, and conscious of updating it: eager, too, to commend its subject matter, to audiences of poetry as well as poets. He follows Amara-kośa carefully in Mān-maṅjari, but not slavishly, editing to remove unwanted material and adding material from other kośas or elsewhere. He imprints Mān-maṅjari with a Braj topicality: in its title with its reference to Rādhā’s mān, or capriciousness towards Kṛṣṇa, and in his choice of headwords to evoke the Rādhā-Kṛṣṇa theme. Occasionally, his comment on a set of Amara’s multiple meanings takes a personal, light-hearted turn, as in a word-play in Anekarth-maṅjari, where he links meanings corresponding to kāma ‘desire’ with the well known Hindi item kām-kāj, meaning ‘work’: the loving worship of Kṛṣṇa is no more than one’s job, Nanddās reminds his audience.

Such adaptation suggests how much Nanddās’s kośas are contemporary documents, how much their composition is motivated by a desire to communicate knowledge of a lexical style now felt as becoming part of living usage. Nanddās composes Mān-maṅjari ‘to the best of his ability’, he tells us, as a ‘garland of synonyms for those who don’t understand Sanskrit, but who want to know its vocabulary’:

\[
\text{samyāhi sakata nabim saṁsakṛta, jñāyau cābata nāma}
\text{tina kājī nanda sumati jathā raci nāma kī dāma.}^1
\]

What the poet is talking about here is a distinction between competence in the Sanskrit language, which is implied to be generally unachievable, and a measure of familiarity with its lexicon, which is feasible, wished for, and recommended. The same idea is expressed in a reference in Anekarth-maṅjari to command of spoken Sanskrit:

At the end of *Anekarth-mañjarī* Nanddās appeals to Kṛṣṇa that this work will be not only recited, but read. The vocabulary he presents in his koṇās is for him both a literary testament handed down from Amara, and an inheritance to be used in a new way, shared in a contemporary community.

It is worth noticing as we leave the topic of koṇās that they demonstrate not just the rise of a Sanskritic style in literary usage, but also its consolidation after Nanddās’s time. Most of the Nanddās koṇa manuscripts we have are from the eighteenth and nineteenth centuries. Their number implies a continuous tradition of copying and dissemination existing from around 1600, and also a growing familiarity with their subject matter. A few manuscripts in Persian script illustrate interest in Brajbhāṣā poetry in Indo-Muslim circles. We are surprised to find that the extant manuscripts of each of Nanddās’s koṇas far outnumber those of any other of his works, including his famous *Rās-pamcāḍhyāyi* ‘Five Cantos’ on Kṛṣṇa’s dance, and despite his reputation as a supreme literary craftsman:

‘*aur sab gariya, nandadas jariya*.

Nanddās’s more elaborate compositions tended to become classics of Brajbhāṣā poetry, and their vogue tended to wane in comparison with later work in the tradition of court poetry, but the koṇas maintained, and even increased their importance into the later period. As manuals of Sanskritic usage in vernacular poetry they were an enduring resource, preserving something from the Sanskrit heritage that was felt to be needed in vernacular poetry, an aspect of Sanskrit which could, in part, still remain alive – its lexicon.

**FRANÇOIS OF TOURS, AND SANSKRITIC USAGE IN THE HINDUSTANI OF GUJARAT**

The Capuchin friars of Surat had become an active presence there by 1660. The traveller Tavernier mentions their ‘commodious house’ and their ‘beautiful

\[2\text{ ibid., p. 99, lines 5–6.} \]
church’, and Capuchin chronicles dating back to 1700 indicate that the community had early acquired a local reputation for good relations both with Christian and non-Christian communities. The Capuchins’ assistance was enlisted by the Council of the French factory when the Council was first established in 1668; they were evidently in a good position to acquire knowledge of local communities and language in the late seventeenth century. They were well received by the local banias, and at this time when Surat was at its peak of prosperity, some banias were generous donors.3

The Italian chronicles of the Capuchins know of François Marie’s reputation as a learned, careful and persevering writer, and of his *Thesaurus Linguae Indiae* as a valued work.4 We know the history of the *Thesaurus* manuscript from several early sources, of which the chief are Beligatti’s *Alphabetum brahmanicum* (1771), a Latin work on non-roman scripts, and a note added by Anquetil Duperron to the copy of the *Thesaurus* he made in Paris. François had presented his manuscript to the administrators of the Capuchin missions in Rome in 1703. Anquetil notes that in the late 1750s he saw the text of a small Hindustani to French dictionary in use among the Capuchins at Surat; later, on learning from the preface to the *Alphabetum* of the existence of the *Thesaurus* and its whereabouts, he realised its probable connection with the dictionary he had seen, and requested permission to borrow the manuscript to copy. He returned the manuscript to Rome in 1784, having had it in his possession eleven months. Anquetil’s copy is now preserved in the Bibliothèque Nationale,1 but François’ original was subsequently lost. George Grierson searched in vain for it in Rome at the end of the nineteenth century, shortly before Caboton’s Catalogue of Indian manuscripts in Paris appeared, and Grierson may have remained unaware of the existence of Anquetil’s copy. Since that time the manuscript has been described in Iftikhar Husain’s catalogue of Urdu, Panjabi and Sindhi manuscripts in Paris.6 Garcin de Tassy had, however, mentioned Anquetil’s copy in his *Rudimens de la langue bin-

---

3 Oriental MS no. 839.
doustanie (1829), and this proved for me some years ago to be a fortunate and quite unexpected, alternative route to finding it in the Bibliothèque Nationale.

It was exciting after tracing this history finally to see Anquetil's material: material repeatedly discussed, and so fortunately copied, but which has remained so little known. It was satisfying, moreover, to find that Anquetil's copy of the *Thesaurus* agrees closely in its external details with the *Alphabetum brahmanicum's* account of it - in its two volumes, number of pages and entries per page, and layout of the text in four columns across facing pages. Occasional annotations by Anquetil confirm the care with which he worked, and indicate that this copy is a remarkable, reliable document.

The expression *lingua indiana* of François' title-page (Pl. 1) denotes a language already known as Hindustani in western India, although the wide-spread speech of Delhi was still chiefly known as 'Hindi' among the Muslim communities of north India in the early eighteenth century. The second line of the title-page states that this language is recorded in its 'own characters', *per proprios caracters*. These are the characters of the Devanagari script, written with some Gujarati features. The idea of the Devanagari script being the proper script of Indian languages, and especially of Hindustani, with a privileged, prior claim to that of Persian, is found in several seventeenth century accounts of visits by European travellers, such as Ovington and Fryer. The early European perception it reflects of Persian language and script as 'non-Indian' was liable to be reinforced, we may assume, wherever European contacts with Hindu communities were close.

The title-page continues that the dictionary glosses are to Latin headwords, and are provided with French equivalents, as well as with a roman script transliteration of the glosses. At the bottom of the page, we have the date of completion, 1703, with the author's initials, his name and Capuchin affiliation.

At more than 10,000 entries the *Thesaurus* has the scope of a substantial reference work. At the same time, it has a contrasting pioneering aspect, for in the manner of Ketelaar and the early British lexicographers, it works into, rather than from its Indian target language. Provision of a Hindustani to French dic-

---

THE SAURUS
LINGUAE INDIANAEE DE

Latine per proprios caracterem
conscriptae, atque in altera ea adverso
paginæ, nostri carborum explicatae.

ad cuius latine, orb majorem
commune, qui latimalem reserunt, com-
meditationem, normatae Gallorum
linguae. Ommæ explicantes.

IHS. MA.

Per F. F. M. Anno MDCCIII.

Id est, per F. Franciscum Mariani de Tourn,
Missionarium Apostolicum
Caputiuinum, id est, p. Leo. "S.

1. Title-page of François Marie of Tours' Thesaurus Linguae Indianae.
   Courtesy Bibliothèque nationale de France.
tionary of the kind that Anquetil saw later at Surat would soon have become essential as a complement to it.

The *Thesaurus* was also, with its carefully worked out transliteration, a work of some scholarship. Anquetil describes François’ system in a two-page note that is his only substantial addition to the author’s text. As the first page of this note illustrates (Pl. 2) François in this pioneering transliteration has arrived at a clear understanding of the structure of the Devanagari script and the phonological analysis underlying it. He uses roman letters with diacritics: superscript dots for most aspirates, tails for retroflexes and accents for most long vowels except finally, and although he does not achieve a complete one-to-one correspondence between transliteration and Devanagari characters he represents the phonology of his target language clearly, in terms of French pronunciation.

The language of the *Thesaurus* is what we might expect from the compilation of the material at Surat in west India: grammatically and in its base vocabulary it agrees quite well with later Hindustani usage, hence also with the Hindi-Urdu usage of today. Various items on the first page-opening of the text (Pl. 3a, 3b), such as *chûp*nâ, glossing *se cacher*, or *chôr*nâ and *chôr* denâ, illustrate the appreciable grammatical similarity of this language to Hindi-Urdu, though both grammatical and phonological divergence from Hindi-Urdu is noticeable. There is a significant component of Persian and Arabic loanwords, sometimes paired with equivalent glosses of Indian origin (as when *chôr* denâ is paired with *tark* karna); but Persian and Arabic loanwords also occur quite frequently alone. Examples are to be found on every page. There is a smaller, though still considerable number of Gujarati words. For instance on this page, alongside present-day Hindi-Urdu *chôta* in *chôta* karna, we find Gujarati *kosar* (itself a word of Arabic origin); also *bâbâr* rather than Hindi-Urdu *bâbar*; and on another page, not Hindi-Urdu *lark*zî but Gujarati *dzkri* glossing *puella, fille*.

Finally we find a proportion of loanwords of Sanskritic origin throughout the text that, though not large, cannot be overlooked. These words belong to very varied categories of meaning: feeling or emotion, religious faith, aspects of everyday experience, or technical terms, as from astronomy or grammar. Many are phonologically simple and so easily acclimatised in vernacular language, while others are more complicated: such as *visvâs* (glossing *fiducia/confiance*), or *sât-vik* (glossing *fidelis*fidele), alongside items like *sôk*, *mûrat*, *graban*, or *drou*. The acclimatisation of Sanskrit loanwords in this language appears to have been a proce-
2. P.1 of Anquetil Duperron's presentation of Devanagari script and François Marie's transliteration as used in Thesaurus Linguae Indianae. Courtesy Bibliothèque nationale de France.
due of some scope and depth. It was, moreover, not an aspect specifically of Gujarati usage. It operated in a Hindustani context, as the author's use of paraphrases that reveal a Hindustani syntax sometimes luckily indicates: for instance in the phrase mot [maut] kā āsvās 'the moment of death', glossing agonía/agonie. This hybrid phrase, with its even-handed drawing on Perso-Arabic and Sanskrit, suggests that Sanskrit loanwords were well integrated in the Hindustani of Surat.

We can make some conjectures with probability about the situation François was in in compiling his material. His sources will have been first and foremost contacts in the Gujarati community of Surat, speakers both of Hindustani and Gujarati. From his view of the status of the Devanagari script it appears that most of those advising him were of Gujarati language identity, while the fact that the language of court and government at Surat was Persian means that a substantial command of Hindustani among many of his sources can be assumed. We can assume from his grasp of the Devanagari script that François will have received some help with its interpretation, as he very probably also did with the writing of the Devanagari glosses in his text; Anquetil (who had seen the text) recorded his opinion that the glosses were written by a Gujarati.

We must ask, as we leave the Thesaurus, whether the Sanskritic aspect of François Marie's Hindustani material was unique, or specific to Surat or Gujarat, or represented a wider usage extending to north India? The circulation in north India of texts in Sanskritised language composed by Gujaratis around 1700 demonstrates that the Sanskritic aspect of François' Hindustani was far from unique or regionally restricted. This is illustrated in the language usage of the religious leader Prānnāth (1614–94) and his followers, notably his biographer, Lāldās (d. ?1694). The Prānnāthis produced extensive texts in Gujarati but also, with an eye to the texts' reception among devotees in north India, in Hindustani versions: poetry in ‘bhākhab hindustām, vāste mominom ke ’, as Lāldās put it. Some of the Hindustani of these texts, especially parts of those by Lāldās, is considerably Sanskritised. The Prānnāthis clearly expected the language of these texts to be viable in the north, and their preservation in the canon of Prānnāthi texts implies that they were justified. It is fair to say that the modern role of Hindi as a language of inter-area communication, a sampark bhāṣā, begins to be anticipated in the Hindustani usage recorded by François Marie around 1700.

<table>
<thead>
<tr>
<th>Latin</th>
<th>Indian</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. De die. Vix 80</td>
<td>Dunqaeq,</td>
</tr>
<tr>
<td>3. Quali.</td>
<td>laba</td>
</tr>
<tr>
<td>4. Auston.</td>
<td>laba</td>
</tr>
<tr>
<td>5. Labe.</td>
<td>laba</td>
</tr>
<tr>
<td>6. Abaya.</td>
<td>laba</td>
</tr>
<tr>
<td>7. Accar.</td>
<td>laba</td>
</tr>
<tr>
<td>8. Abeg.</td>
<td>laba</td>
</tr>
<tr>
<td>9. Abalum.</td>
<td>laba</td>
</tr>
<tr>
<td>10. Abali.</td>
<td>laba</td>
</tr>
<tr>
<td>11. Abaga.</td>
<td>laba</td>
</tr>
<tr>
<td>12. Labaga.</td>
<td>laba</td>
</tr>
<tr>
<td>13. Caster.</td>
<td>laba</td>
</tr>
<tr>
<td>14. Se caster.</td>
<td>laba</td>
</tr>
<tr>
<td>15. Sub.</td>
<td>laba</td>
</tr>
<tr>
<td>16. Sum.</td>
<td>laba</td>
</tr>
<tr>
<td>17. Dinf.</td>
<td>laba</td>
</tr>
<tr>
<td>18. Dipo.</td>
<td>laba</td>
</tr>
<tr>
<td>19. Dini.</td>
<td>laba</td>
</tr>
<tr>
<td>20. Dijner.</td>
<td>laba</td>
</tr>
<tr>
<td>21. Cash.</td>
<td>laba</td>
</tr>
</tbody>
</table>

Ketelaar’s manual (Instructie), exactly contemporary with François Marie’s The­saurus, comprises a grammar of Hindustani and a word list of over 2,000 items, together with a valuable introduction summarising features of both Hindustani and other forms of north Indian speech.9 Ketelaar, who was born in Elbing (El­blag, Poland) and first came to the Netherlands in the early 1680s, became a se­nior officer of the Dutch East India Company and was known in the Company for his proficiency in Hindustani. He first arrived in India in 1683, at Surat, and for at least thirteen years was based in Gujarat, mainly at Surat.10 During that time, however, he evidently spent long periods in, and travelling to and from, Agra. From Ashin Das Gupta’s study of Surat we can conclude that he probably spent four or five years at Agra prior to completion of the Instructie. He is likely to have travelled east on several occasions purchasing for the Agra factory, and on these occasions, in Lucknow and elsewhere, he would have acquired the knowledge of eastern language usage, varying pronunciations, and writing prac­tices shown in his Introduction.11

The wording of the title-page of the Instructie as given in the Hague manu­sript indicates that Ketelaar thought of his materials first and foremost as a lan­guage manual. (Pl. 4) Prominence would be given in it to learning aids such as tables of declensions and conjugations, and ancillary materials would be pro­vided. By contrast to such material the word list goes unmentioned on the ti­tle-page. The Instructie text was copied at Lucknow by a colleague of Ketelaar’s, Isaac van de Hoeve, a Company accountant who was of Utrecht origin.


11 See Ashin Das Gupta, Indian Merchants and the Decline of Surat, Wiesbaden, 1979, p.57.
A second manuscript of the *Instructie* is preserved in the University of Utrecht Library. Its text corresponds closely to that of the Hague manuscript, but with some interesting discrepancies, including one which seems to connect it more closely than the Hague manuscript to Ketelaar's original material. The Utrecht manuscript, which is written partly in Gothic script, seems not to have been produced with quite the same motivation and purpose as the Hague manuscript. For reasons that are unclear, Ketelaar's name is not acknowledged on—perhaps has been deleted from—the title-page, and the word list has been given a privileged position in the text, coming before the grammatical material. (Pl. 5) We conjecture that the Utrecht manuscript may have been copied from Ketelaar's original material for private use by someone already having some acquaintance with Hindustani, and less familiar with its vocabulary than its grammar.

Ketelaar's word list amounts to a large polyglot vocabulary which, since it is provided with an alphabetical index, would also have functioned as a dictionary. Its Dutch headwords and expressions are glossed in both Hindustani and Persian, with the items arranged in sections devoted to categories of meaning: objects, persons and relationships, entities and ideas, actions and activities. The first page of the Utrecht manuscript text (Pl. 6) illustrates *wereld* 'world' glossed as *dunnia*, *hemel* 'sky, heaven' as *asmaan*, *zon* 'sun' as *soerets*, etc. Ketelaar's material is of a similar range of lexical origin to what we found in the Surat *Thesaurus*, comprising colloquial north Indian, Persian and Arabic, and Sanskritic words. His representation of Persian and Hindustani words is a phonological transcription in terms of Dutch pronunciation and spelling: not a transliteration of spellings, which Ketelaar, who perhaps did not read either script, could not have achieved. Ketelaar hears, but does not always clearly distinguish the features of Hindustani pronunciation that are new to him. Hence, in his introduction, his throw-away remark that 'many letters' are pronounced 'in the middle of the mouth' (*binnens mond*). Such a perception, correct as far as it goes but too imprecise to be useful, leads Ketelaar to transcribe various sounds inconsistently. He confuses the processes of aspiration, retroflexion and lengthening of consonants, and details of

---

12 no. 1478. I am indebted to Dr B.J. Slot of the Algemeen Rijksarchief, The Hague for drawing my attention in 1993 to this manuscript.

13 Identical text in the Hague ms.
Title-page of Ketelaar's Instructie (Utrecht M.S.). Courtesy Universiteitsbibliothek, Utrecht.
<table>
<thead>
<tr>
<th>Dutch</th>
<th>Hindustani</th>
<th>Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>God</td>
<td>alla</td>
<td>goda</td>
</tr>
<tr>
<td>Kiev</td>
<td>assai</td>
<td>massa</td>
</tr>
<tr>
<td>Engie</td>
<td>enuia</td>
<td>bétar</td>
</tr>
<tr>
<td>God</td>
<td>gowies</td>
<td>mafhi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dutch World</th>
<th>Hindustani World</th>
</tr>
</thead>
<tbody>
<tr>
<td>De wereld</td>
<td>dinia</td>
</tr>
<tr>
<td>Heinde</td>
<td>abinaan</td>
</tr>
<tr>
<td>Zon</td>
<td>doets</td>
</tr>
<tr>
<td>Maan</td>
<td>spaand</td>
</tr>
<tr>
<td>Zonne Schijn</td>
<td>stemp</td>
</tr>
<tr>
<td>Maand Schijn</td>
<td>spaandon</td>
</tr>
<tr>
<td>Vérre</td>
<td>dore</td>
</tr>
<tr>
<td>De voeten</td>
<td>dertel</td>
</tr>
<tr>
<td>De wessen</td>
<td>abed</td>
</tr>
</tbody>
</table>

his transliterations of individual words (as in soerets) will have been misleading to a greater or lesser extent.

We turn now, with reference to both manuscripts of the Instructie, to Ketelaar’s introduction. Vogel and Bodewitz have both drawn attention to Ketelaar’s description of the mixed, Indo-Persian lexical nature of Hindustani, and the existence alongside Hindustani of a less hybridised, hence more ‘Indian’ style of language. The Hague text (H) of this passage translates to English as:

‘the language of the ‘Moors’ is called the speech of Hindustan, although it might rightly be called a degenerate Hindustani, for the speech of the Moorish Muslims is far different from that of the true Indians or indigenous population, and to make a distinction in names between them the Moors’ speech is called, as mentioned above, hindustān ki boli.’

This is a useful statement, but it is evidently incomplete as it stands, for it omits all reference to the second of the two languages which provides the point of the discussion, that of the indigenous population. The Utrecht manuscript (U) continues at this point as follows, however, in the same sentence:

‘and the language usual among the indigenous population is called hinduki boli’ [Hinduki language].

The Dutch words translated here supply an evident lacuna in the Hague manuscript. Occupying just the length of a line of text, they illustrate that in copying the latter a scribal lapse of a routine kind has taken place. (Pl. 7)

We need therefore to consider the sense of the term hinduki, and its variant hindugi. The term is not common, but can be seen in several sources to refer to an Indian language as contrasted in some way with Persian or Hindustani. The fourteenth century sufi poet Dāūd uses hinduki to refer to the Avadhī language of his Candāyan, distinguishing it from the turki or Persian script in which he

14 Dutch text for final part of the passage translated from H, with the additional line found in U: [H] en om tusschen de moorse en heydense tale in het noemen een verschil te maken, zoo noemt men der mooren taale, gelijk boven gesijd, hindoustanki boelie, [U] en die bij de heydenen gebruikkelyk Hindoukie boelie, [H] [U] ook werd de moorse...
wrote Candāyan down. Later the form hindubi, and the much commoner hindavi or hindui are found, referring to the language of sufi poets of the Delhi region. Towards mid-seventeenth century the Jain poet Banārsidās of Agra refers to Persian and Hindugi together, as the two languages of his father's education. It is clear that by Hindugi, Banārsidās means both the language of the Delhi region and something that would be important in a Jain merchant family – its Devanagari script.

How then did Ketelaar understand the term? We note that his reference to Hinduki boli is linked to his preceding reference to a lexical aspect of Hindustani usage – its ‘degenerate’ or hybrid nature. It seems clear, therefore, that his reference to ‘the true Indians” language should similarly be taken to have a lexical sense. If this is so, we are justified (since Ketelaar’s hinduki language ‘differs greatly’ from that of the Muslim communities) in believing that Ketelaar regarded Hinduki as a style of the Hindi of Delhi showing a larger proportion of colloquial Indian vocabulary, and probably of Sanskritic loanwords, than the Hindustani of his word list.

Ketelaar makes a further reference to language usage which is relevant to his understanding of the word Hinduki. Dasgupta notes his report in the Company’s records, in which he describes the town of Gwalior (which he visited in 1699) as ‘famous for its high Hindustani language in which most of [the] local

poetry and songs are written'.

This reference to a 'Hindustani language' is clearly not to the Persianised Hindustani which is the subject of the Instructie; nor, given that it is a 'high' language, can it be to any local speech. We must, then, consider whether by 'high Hindustani language' Ketelaar understands some form or usage of Brajbhāṣā.

The history of Brajbhāṣā poetry, with its origins partly at Gwalior, and the fame of Mānsingh of Gwalior's court musicians in the fifteenth century, tend to support the above possibility. It is strongly supported also by the fact that Indian Persian poets of the eighteenth century are found to refer to Brajbhāṣā as 'the language of Gwalior'. The poet Banvālidās, for instance, does so, in his Gulzār e ḫāl based on Nanddās's version of the Sanskrit drama, Prabodhacandrodāya. It may seem a paradox that Ketelaar could refer to Brajbhāṣā as 'high Hindustani language'. I am suggesting that this is possible for him, because he is aware of a tendency operating in his time for Brajbhāṣā and Hindukī to draw together in their functioning as language vehicles in informal urban usage.

A model for such a tendency exists in the mixed Brajbhāṣā and Hindukī of the Agra merchant Banarsīdas's autobiographical poem, Arddh-kathānāk 'Half a Tale'. Banarsīdas (whose father as we saw had studied Hindukī) calls this language 'the speech of the madhyadēṣā', madhyadēṣā ki boli. The culturally evocative expression madhyadēṣā means for him not only a home ground of traditional culture, but also the north Indian world of his own day. In Banārsidās's somewhat haphazard mixture of language, Hindukī and Brajbhāṣā are used in apparently free variation, suggesting that both speeches were in constant use in circumstances where little need was experienced to keep them separate. Sanskritic loanwords occur in both the Hindukī and Brajbhāṣā passages of Arddh-kathānāk. Some of these loanwords are quite colloquialised, which suggests that a process of active acclimatisation of loanwords in speech was under way in Banārsidās's time.

It appears that two levels of Brajbhāṣā usage were in operation during the sev-

---

18 Das Gupta, Indian Merchants, p. 52.
191877 lithographed edition, Lucknow, introd.
enteenth century: one the widespread, ‘high’ literary usage, and the other a usage perhaps no less widely spread: loosely linked to Hindustani in urban communities acrosss the madhyadeśa. The distinctive features of modern Hindi, namely its Devanagari script, its Hindustani language base, and its lexicon in which elements of Sanskritic and Persianised language are both represented, can be seen here in process of assembly.

TUHFATU'L HIND; EIGHTEENTH CENTURY PERCEPTIONS OF SANSKRITIC CULTURE AND OF PERSIANISED LANGUAGE

The preface to Mīrzā Khān’s Persian compendium on Brajbhāṣā and its poetry called Tuḥfatu'l Hind illustrates the reception of Brajbhāṣā among the Indo-Mus­lim community of the Agra-Delhi region at a time shortly preceding the rise of Delhi Urdu. Composition of the Tuḥfat within the community throws some light on the swing that would soon follow among its members, away from use of Persian as their literary language towards a more Persianised usage of Hindustani.

Mīrzā stresses that the arts and literature of Braj are of great interest to, and highly valued by the Indo-Muslim community, many of whom will wish to know more about them. He notes in his introduction that the Indian word bha­khā, which in principle means any vernacular language, has now acquired a particular application as the ‘language of the Braj people’, and he describes the bha­khā of Braj as ‘the most eloquent of languages’ ... containing ‘poetry full of colour and sweet expressions and of the praise of Krṣṇa and Rādhā’, and fa­voured among poets and cultivated people.21

The members of the north Indian Indo-Muslim communities would have had as their mother tongues chiefly the mixed Hindustani of general urban use, or the Hindi speeches of the Doab, or Panjabi. A great majority of these people would by the seventeenth century have had some passive knowledge of Brajbhāṣā, if not anything like mother-tongue command of it, but since their liter­ary language was Persian, we imagine that few of them would have had a good

command of the Sanskritic vocabulary of sixteenth and seventeenth Brajbhaṣā poetry. The members of Indo-Muslim society thus represented a particular class of the devotees of Brajbhaṣā poetry who, as Nanḍās had put it long before, could not understand Sanskrit, but wanted to understand the Sanskrit terms used in poetry.

Mīrzā’s relative concentration on Sanskrit and Sanskritic loanwords suggests just how much the Tuḥfat is the product of a bipartite culture. His Brajbhaṣā-Persian dictionary materials show the characteristic feature of a bilingual dictionary of including many headwords that are clearly already familiar to his intended readership — in this case, speakers of the Hindi of Delhi literate in Persian.22 But Mīrzā pays scant attention to many such headwords. In explaining the three homonyms bās, for instance, meaning ‘smell’, ‘clothing’ and ‘house’ respectively, he dismisses the common sense ‘smell’ on his readers’ behalf in a summary way that not all students of, let’s say, Brajbhaṣā as a foreign language might be ready to do. He is content to say ‘there are three senses of this word and the first is ‘smell’, as is well known (ān maḏrīf)’. Mīrzā then goes on to describe the other meanings of bās which are more difficult for him. Elsewhere in his text he attaches the Persian tag maḏrīf ast, meaning ‘this sense is well known’, to many glosses of Brajbhaṣā words, and even of simple Sanskrit words.

We should not, then, see this dictionary as a one of a foreign language. It resembles, rather, the English language dictionaries of what were called ‘hard words’, that were produced in the seventeenth and early eighteenth centuries for the Latin and Greek loanwords borrowed into English, or coined, in the new cultural conditions of the Renaissance and thereafter. The Tuḥfat addresses the challenge posed by the ‘hard words’ of Sanskrit to members of a Hindustani-speaking community oriented to use of Persian language. They were drawn to Brajbhaṣā poetry by its interest and charm, but also, we may suppose, by the sense of growing difficulty attaching to the use of Persian as their literary language which is recognised to have contributed to the rise of Delhi Urdu. But they may have been in a dilemma. If Persian had its problems as a literary language for members of the Indo-Muslim community, what other feasible language was there for them to use? The only alternative was the speech of Braj:

22 The dictionary materials of Tuḥfat u’l-Hind are unpublished; the above paragraph is based on text of a Cambridge ms (Pote Kings no. 119).
so close to their own, yet, as its 'hard words' made painfully clear, not really their own in the last resort.

Mirza perhaps bore in mind in naming the *Tuḥfat* that its readers or users would take the work in the spirit of its title, as a 'gift' to them of knowledge of the wider mixed, modern culture of north India. Yet the gift, attractive as it was, was destined not to be fully 'accepted', and the space given to 'hard words' in the *Tuḥfat* was perhaps a portent of that situation.

When the move towards Delhi Urdu is seen in the wider context of the Mughals' reappraisal of their political situation during the reign of Aurangzebe, it can be connected with a converse development then in progress in Brajbhāṣā poetry. We can trace from the late seventeenth century a new interest among Brajbhāṣā poets in making adaptations from Sanskrit works – *Mahābhārata*, *Bhāgavata-purāṇa* and others. One of the earliest of these adaptations was Navāz's *Sakuntalā* (composed for Āzamshāh, the likely patron of Mīrzā Khān). Composition of these works seems to illustrate a wish arising from Aurangzebe's time for the reassurance that could be found in a redefinition of cultural values during a period of political uncertainty.23

A tendency in the language of these works suggests similarly that poets' and audiences' perceptions of the traditional culture were subtly changing at this time. The language of these works is more explicitly Sanskritised than that of the older poetry. Unmodified (tatsama) forms of the borrowed words now become commoner, displacing vernacularised forms to some extent. This tendency is likely to have been encouraged in the eighteenth century by the rise of the more intensively Persianised Hindustani that was soon to be known as Urdu. There was a parallel impact of Persian vocabulary on Brajbhāṣā at this time, but not all of it was entirely welcome. The court poet Bhikhāridās for instance, commenting on current styles of Brajbhāṣā, accepts, in mid-eighteenth century, what he calls its *sahaj pārsi* style – its 'easy Persianised usage'; but Bhikhārī clearly disapproves of the *jaman bhākhā* 'foreigners' language' in which obscure Persian

words became a burden. The increasing Sanskritisation of Brajbhāsa, contrasting with this Persianised usage, anticipates the nineteenth century emergence of modern Hindi and Urdu as separate styles of language.

There is another eighteenth century portent of the separation of Hindi and Urdu. A few prose texts on Hindu or Jain cultural topics now appear showing all three distinctive linguistic features of modern Hindi in combination: its essentially Hindustani language base, its Sanskritic vocabulary option and its Devanagari script. We have already (in Banāsīdās's Arddh-kathānāk) noticed the presence of these three linguistic features together, in combination with Brajbhāsa, in poetry of the eighteenth century. The present eighteenth century prose texts mostly show some Brajbhāsa, Panjabi or Rajasthani features, but they are, nevertheless, immediate typological ancestors of modern Hindi literary prose style.

We might say, therefore, that the language that began early in the nineteenth century to be recognised under the name of modern Hindi was now already in effective existence, independently of European presence or influence, at both colloquial and literary levels.

SOME DICTIONARIES OF THE YEARS BETWEEN 1780 AND 1850

The above developments up to the late eighteenth century may be linked, in conclusion, to the time from the 1860s onwards, when the rising consciousness of Indian and Hindu identity among north Indians brought a mood favouring use of Hindi in the Devanagari script and Sanskritic language. The work of certain early British lexicographers had some influence up to that time in the direction that would be taken in institutionalising, via the new print culture, the form of Hindūī that would soon take the name Hindi.

William Kirkpatrick, who was a Persian interpreter to the British East India Company, had seen the importance of the Devanagari script and of Hindūī by around 1780. He had drawings made of the script, and produced from them in


\[25\] Some details in McGregor, Hindi Literature, pp. 213–4 and references.
London the first Devanagari font that had been cast in Britain. He planned to use this type in what he would call a ‘New Hindvi Grammar and Dictionary’, which would be a work of several volumes. The scale of this plan suggests that Kirkpatrick had a grammar of Hindustani in the Devanagari script, and a dictionary of colloquial and loanword usage in mind; but beyond the Arabic and Persian vocabulary that would have made up his last volume, nothing of this work was ever published.\(^{26}\) To begin his project with these loanwords was of course Kirkpatrick’s easy option as a Persianist. We imagine that a sense of the loanwords’ typological identity also influenced him, a sense of their manageability as compared with the colloquial, less well known Hindvi. It was as if, in Hindvi, Kirkpatrick saw a language defined not by its own identity, but entirely by loanwords and script.

John Gilchrist, who became head of the Hindustani Department of Fort William College in 1800, was more involved with Hindustani than Kirkpatrick had been. Gilchrist had arrived in India in 1782, and was engaged in language work from the outset. Like Kirkpatrick, he too perceived the importance of Hindvi as a style current among Hindus. As he puts it in the preface to his important English-Hindustani dictionary of 1787, ‘the learned Musulman glories in his Arabic and Persian, the Hindoo is attached to his Sanskrit and Hinduwee’. But practicalities dictated that Gilchrist’s chief concern was with Hindustani. He was uneasy about the fact that both Devanagari and Persian scripts were in use to write Hindustani. To use both scripts for this particular purpose seemed an unnecessary complication to Gilchrist, one to be removed from the study of Hindustani if possible since it would create problems almost impossible to resolve in the future. To Gilchrist’s classicist imagination these problems presented themselves as a ‘Gordian knot’, meriting nothing less than to be decisively ‘cut asunder’.

Thereafter, at Fort William College, with Persian script and language perceived as the criteria of Urdu language and style, it is easy to see how a converse image of Hindi as defined by use of Sanskritised language and Devanagari script could arise. The intentions of Kirkpatrick had been lost sight of amid difficulties of language acquisition and policy formation, as the British went too far in promotion of a lexically uniform Hindi style from which words of Persian origin

would be almost completely excluded. Hindvī materials composed at the College would be specifically Sanskritised. Not only would the more difficult of the Persian words now used in Urdu be avoided—corresponding to Bhikhāridās’s jaman bhākhā style— but the simpler ones also, his sabaj pārsi: words that had been at the colloquial base of Hindustani and in general use for centuries.

The later lexicographers Adam (around 1830) and John Thompson were influenced by the idea of Hinduf as a Sanskritic style (now increasingly called Hindi) that was taking hold in their time. But Thompson was aware, as his preface to his Hindi-English Dictionary of 1846 implies, that there were Persian and Arabic words current in the everyday spoken language that had gone unrecorded in his Hindi materials. At the same time he realised (as Kirkpatrick had also done) the dangers of including loanwords over-liberally from the vast literary repertoires of Persian and Sanskrit. Clearly there was a need for the foreign lexicographer to be guided by Indian usage.

The work of Harīscandra of Banaras, the first major writer of modern Hindi, demonstrates that few difficulties would be experienced from the 1870s onwards in developing Hindi as a versatile prose medium. Hindi now began to be used intensively as a vehicle of prose. The earlier openness of Hinduf to receiving Sanskritic elements, while retaining many of Persian, meant that modern Hindi writers would soon consolidate the flexible usage that was essential, if they were to claim a place for Hindi alongside Urdu and English.

27 The new name was in use in Fort William College publications by the 1820s. M.T. Adam’s Dictionary, English and Hindīwī, Calcutta, 1833 was subtitled in Devanagari script, but Bengali language, Ḫirajī o hindī abhidhan.
The Gonda lectures are organized every year by the J. Gonda Fund. The J. Gonda Fund was established in 1993 under the auspices of the Royal Netherlands Academy of Arts and Sciences and has the task of administering the inheritance bequeathed to the Academy by the Sanskritist and Indologist Professor J. Gonda.

Published in this series:
H.W. Bodewitz: Jan Gonda, 14 April 1901-28 July 1991, 1992
W. Doniger: Masquerading Mothers and False Fathers in Ancient Indian Mythology, 1993 Gonda lecture
G. Obeyesekere: Amerindian Rebirth and Buddhist Karma: An Anthropologist’s Reflections on Comparative Religious Ethics, 1995 Gonda lecture
R.F. Gombrich: Kindness and Compassion as Means to Nirvana, 1997 Gonda lecture
J. Bronkhorst: Why is there Philosophy in India? 1998 Gonda lecture
Janice Stargardt: Tracing Thought through Things: The Oldest Pali Texts and the Early Buddhist Archaeology of India and Burma, 1999 Gonda lecture