



# INDIA XVIII. PERSIAN ELEMENTS IN INDIAN LANGUAGES

---

## INDIA

### xviii. PERSIAN ELEMENTS IN INDIAN LANGUAGES

Some Persian elements are present in most of the modern languages of the subcontinent of South Asia, as a consequence of the prolonged cultivation of Persian associated with pre-modern Indo-Muslim culture. In those languages and registers in which the impact of the classical Indo-Muslim civilization is most clearly discernible the presence of Persian elements is a very substantial one. In order to provide a clear if simplified picture of the quite complex factors which need to be considered in an overall mapping, the following treatment begins with a survey of the historical and cultural processes governing the differential presence of Persian elements in Indian languages. This is followed by descriptions of the principal elements which may be discerned, arranged in terms of broad linguistic categories.

*Persian and South Asian languages.* The vast region of the Indian subcontinent is characterized by great linguistic diversity. The variety of local spoken languages has, however, always been balanced by the parallel use of trans-regional standard languages in royal administration and in religious and



secular literature. In the earlier period of Indic civilization, the standard language was Sanskrit, carefully preserved by the Brahminical elite as a learned classical tongue, a function which it continues to preserve in the religious context of orthodox Hinduism. Elsewhere Sanskrit came to be replaced by Persian as the major language of administration and literary culture, first in the Ghaznavid kingdom of Lahore in the northwest, then over much of northern and central India after the substantial Muslim conquests of the 13th century onwards, with subsequent reinforcement thereafter through the continuing immigration of Muslims from the larger Persian world of Iran and Central Asia.

The core position of Persian in the pre-modern Indian education system ensured its wide diffusion as a pan-Indian standard language amongst the indigenous elite and service classes, including significant groups of Hindus as well as the Muslims. Itself somewhat distinguished from the Persian of Iran both by its greater conservatism and by the influence of Indian languages, this Indo-Persian remained culturally dominant until the consolidation of British rule in the 19th century led to its substantial replacement by English, the transnational elite language of all the countries of modern South Asia.

Both geography and cultural history have conditioned the levels of Persian elements variously present in the different languages of the subcontinent. Linguistic contact has been greatest between Persian and the Indo-Aryan languages (see Masica, 1991; Cardona and Jain, 2003) of the northern plains which are linguistically cognate with Iranian, particularly the languages of the northwest like Sindhi and Panjabi, which are closest to the Indo-Iranian linguistic border (for Dardic languages see [DARDESTĀN ii. LANGUAGES](#)). Generally speaking, the currency of Persian elements diminishes with progressive distance from that border, whether towards Bengali at the eastern end of the vast Indo-Aryan area, or towards the south, where Persian elements are markedly less prominent in Marathi than in Gujarati. Going still further south across the major linguistic divide of the region, Persian elements are predictably much less well represented in the Dravidian languages spoken in peninsular India.

This geographical spectrum is, however, crosscut by socio-political factors. The earliest centuries of Muslim rule in South Asia saw the rise of a lingua franca based upon an amalgam of Persian elements with an Indian linguistic base, which has been identified as a mixture deriving from the local dialects of the Lahore-Delhi region. This mixed lingua franca is conveniently given the



traditional label “Hindustani,” in distinction from “Urdu” (q.v.), defined as the highly Persianized language of the elite which evolved from it as a literary language written in the Persian script, first in Hyderabad and the Muslim kingdoms of the Deccan and then in Delhi, Lucknow, and other centers of Muslim courtly culture in northern India (see further Shackle, 2000).

In the colonial setting of the 19th century, the increasing competition between different groups in Indian society came to be pursued in terms of language as well as of religion. This was particularly the case in northern India, where the position granted by the British to Urdu as a convenient vernacular standard for use alongside English was challenged with increasing success by proponents of a de-Persianized “Hindi,” structurally identical to Urdu but written in the Devanagari script associated with Sanskrit and characterized by the exclusion of Persian words in favor of learned Sanskritic equivalents (called *tatsama* as opposed to the inherited Indo-Aryan *tadbhava* vocabulary). The opposition between Urdu and Hindi came to be enshrined in their respective status as official languages in Pakistan and in India, and although the total identification between language and religion sought by zealots on both sides is far from being achieved, an important touchstone is provided by the presence of Persian elements (see further Shackle and Snell, 1990).

These elements are most prominent in formal Urdu, notably present in languages of the north and west such as Punjabi or Sindhi (see separate articles) or Gujarati, especially in the usage of Muslim speakers, and are also well represented in everyday Hindustani, which has itself acted as a medium for the transmission of Persian loans into other Indo-Aryan languages, besides the Dravidian Telugu in the Deccan (Steever, 1998, p. 237). Persian elements are, however, rather sparse in other Dravidian languages (other than the special case of Brahui; cf. Rossi, 1979) and are most rigorously excluded from the modern Hindi, which is most closely identified with the cause of Hindu nationalism. Although the picture is thus a dynamic one, it remains true that, along with English, Persian still provides the Indo-Aryan languages with their most substantial set of non-Indian elements.

*Persian elements in South Asian languages.* Against this general background, the following survey first describes the Persian influence on the phonology and the morphology and syntax of South Asian languages, before concluding with a survey of the major Persian impact on vocabulary. For the sake of conciseness, detailed parallel examples from different languages have been omitted here in favor of a broad representation through citations of Urdu



words and forms, with distinction where appropriate between the modern standard language, taken as the exemplar of elite usage which is also partially reflected in other Pakistani languages, and the spoken “Hindustani” lingua franca, whose Persian inventory is to a greater or less degree reflected in many other Indian languages.

*Phonology.* The assimilation of Persian loans into South Asian languages has been eased by phonological similarities. Indo-Persian (cf. Qureshi, 1965) is phonologically conservative, retaining the *majhul* vowels *ē* and *ō* in some 10-vowel system as Indo-Aryan, here written *a ā e i ī u ū ē ai ō au* (versus modern Persian *a ā e i o e [i] ey [u] ow*). At the level of word phonology, too, modern Indo-Aryan shares the Persian preference for simple patterns of the CVCVC type. In the consonantal inventory fricatives are historically less well represented in Indo-Aryan than in Iranian, and it is the influence of Persian which largely accounts for the presence in modern South Asian languages of *z*, *š*, *f* (for which everyday speech commonly substitutes *j*, *s*, *ph*). The careful preservation of the Persian phonemes *k*, *g̃*, *q* is a mark of elite speech in Urdu and other languages, as opposed to the realizations *kh*, *g*, *k* characteristics of common speech.

*Morphology and syntax.* The Indo-Aryan languages are of the same subject-object-verb (SOV) type as Persian, with the verb placed finally in the sentence. Within the construction of the noun phrase, however, Indo-Aryan has the order modifier-noun-postposition versus the Persian order preposition-noun-modifier, with elements typically linked by the *eżāfe*—thus, e.g., Urdu *mazīd imdād kē ṭaur par* “by way of additional assistance” versus Indo-Persian *ba-ṭaur-i imdād-i mazīd*. Modern Urdu quite freely admits noun phrases of the Persian type, e.g., *makānāt barā-ē farōkt*, translating English “houses for sale,” or *aqwām-ē muttaḥida* “United Nations” versus Hindi *samyukt rāšṭr*, with preceding adjective. The use of inflected Persian nominal forms, e.g., noun plurals in *-ān* or (less commonly) *-hā* or superlative adjectives in *-tarīn* is typically restricted to phrases of this type. The use over many centuries of Indo-Persian as a formal standard language characterized by a more elaborate syntax than the typical parataxis of Indo-Aryan is reflected in the ubiquitous borrowing of Indo-Persian *ki* “that” to introduce reported speech, as well as in numerous other conjunctions, illustrated by such Urdu examples as *lēkin* “but,” *bā īn hama* “nevertheless,” *čūnki* “because,” *bā-wujūdē-ki* “in spite of the fact that,” for all of which Sanskritic substitutions are plentifully represented in other languages. More speculatively, it has been suggested (Masica 1981,



1991, p. 367) that the marking of direct objects by a dative-accusative postposition (Urdu *kō* and its equivalents) may have been influenced by the example of Persian *-rā*.

*Vocabulary.* Loanwords unambiguously constitute by far the most prominent Persian elements in South Asian languages. These naturally include Perso-Arabic and Turkish words borrowed through Persian as well as words of Persian linguistic origin. The long history of Persian in India is reflected in well-established borrowings across a large range of semantic categories (Masica, 1991, pp. 71-73), with nouns constituting by far the largest grammatical category of loans. Selective inventories of such loans, in many cases attested from texts of the 16th century onwards may be found in the principal histories of the main languages (e.g., Chatterji, 1970, vol. I, pp. 200-214; Saksena, 1971; Shackleton, 1978, 1995). As would be expected, one core class of such loans is formed by words connected with Islam, including both religious vocabulary like *namāz* “prayer,” *masjid* (> popular *masīt*) “mosque,” *pīr* “Sufi master,” and words connected with distinctive Muslim practices like the eating of meat, as in *gōšt* “meat,” *qaṣāī* (probably) “butcher,” or the preference for tailored clothes, e.g., *qamīz* (< *qamīṣ*) “shirt,” *pājāma* “trousers,” *qaincī* “scissors,” *darzī* “tailor.” The urban orientation of Islam is reflected in numerous common loans like *šahr* “city,” *bāzār* (> popular *bazār*) “market,” and its role in education by words like *kāgāz* “paper,” *qalam* “pen,” *siyāhī* “ink.” A very considerable further class of core loans is made up of words connected with administration, e.g., *bādšāh* () “king,” *zila’* “district,” *qānūn* “law,” *fauj* “army,” although this last class of words has naturally been subject in many Indian languages to conscious replacement by Sanskritic *tatsama* equivalents. Conversely, the number of such Persian loanwords has been vastly increased in modern Pakistani usage by fresh borrowings and coinages (cf. Barker et al., 1969) added to the already enormous Persian abstract vocabulary which characterizes formal Urdu (cf. Chaudhuri, 1975; Khweshgi, 1989), e.g., *barr-ē ṣaḡīr* “subcontinent” (where the contrast with Persian *nim-qārre* points to the divergence of Urdu from the nativizing tendencies of modern Persian coinages), or *taraqqī-paḍīr mamālik* “developing countries.”

Besides the numerous syntactic markers of Persian origin noted above for Urdu, Persian loans also provide many core adjectives, e.g., Urdu *kūbsĀūrāt* “beautiful” versus Persian *zibā*, or *siyāh* “black” alongside native *kālā*. The adjectival stock is further expanded by the free use of Persian compounding formations, for which there are few Indo-Aryan equivalents, e.g., the negative



prefix *bē-*, found not only in loans like *bē-šabr* “impatient” but also with Indo-Aryan elements as in *bē-čain* “restless” (< *čain* “repose”). Modern coinages include many more extended examples, e.g., *nāqābīl-ē bardāšt* “intolerable.” Well-established loans in other grammatical categories include pronouns, e.g., Urdu *k̄ud* “self,” alongside native *āp*, and numerals, including Urdu *awwal* “first” alongside *pahlā* and the ubiquitous *hazār* “thousand.”

The indigenous stock of native verbs, itself much more numerous than in Persian, is enhanced by only a few loan-formations, e.g., Urdu *qabūlnā* “to accept.” As in Persian itself, however, loans are most commonly used to form nominal compounds with simple formants like *karnā* “to do” for transitives and *hōnā* “to be” for intransitives, e.g., Urdu *šurū’ karnā* “to begin (trans.),” *šuru’ hōnā* “to begin (intrans.)” The influence of Persian is often reflected in the choice of marker, e.g., Urdu *khānā* “to eat” in *šikast khānā* “to be defeated,” reflecting Persian *šekast kardan*, and in such common honorific formations as Urdu *far-māiyē* “please say” or *tašrif rakhiyē* “please sit down,” which continue to typify the role of Persian in South Asia as the linguistic symbol of the refinement associated with the courtly traditions of Indo-Muslim culture.

## BIBLIOGRAPHY

---

- M. A. R. Barker et al., *An Urdu Newspaper Word Count*, Montreal, 1969.
- G. Cardona and D. Jain, *The Indo-Aryan Languages*, London and New York, 2003.
- S. K. Chatterji, *The Origin and Development of the Bengali Language*, London, 1970.
- S. Chaud-huri, *Farhang-e vāžehā-ye fārsi dar zabān-e ordu*, Tehran, 1375.
- M. A. Khweshgi, *Farhang-e āmera*. Islamabad, 1989.
- C. P. Masica, “Identified Object Marking in Hindi and Other Languages,” *Topics*



in *Hindi Linguistics* 2, 1981, pp. 123-46.

Idem, *The Indo-Aryan Languages*, Cambridge, 1991.

W. Qureshi, "The Indian Persian," in A. S. Dil, *Studies in Pakistani Linguistics*, Lahore, 1965.

A. Rossi, *Iranian Lexical Elements in Brāhūī*, Naples, 1979.

B. Saksena, *The Evolution of Awadhi*, Delhi, 1971.

C. Shackle, "Approaches to the Persian Loans in the *Ādi Granth*," *BSOAS* 41, 1978, pp. 73-96.

Idem, *A Gurū Nānak Glossary*, New Delhi, 1995. Idem, "Urdū," in *EI* <sup>2</sup> IX, pp. 873-81.

C. Shackle and R. Snell, *Hindi and Urdu since 1800*, London, 1990.

O. Spies, "Türkisches Sprachgut im Hindūstānī," in W. Kirfel, *Studia Indologica*, Bonn, 1955, pp. 321-43.

S. B. Steever, *The Dravidian Languages*, London and New York, 1998.