



## URDU

---

**URDU**, the national language (*qaumī zabān*) of Pakistan and one of the fifteen officially recognized languages of India. It is spoken, according to recent censuses made in India and Pakistan, by an estimated 53 million people in the South Asian subcontinent (Schmidt, 2004, p. 288). To this we may add the millions of people, both inside and outside the subcontinent, who use Urdu as a primary means of communication. The Panjabi-speaking population of Pakistan, for example, employ Urdu rather than their own language almost exclusively as a written and literary medium. Along with its “sister language,” Hindi, with which it shares a virtually identical grammatical base, Urdu at the most basic, spoken level still functions as a convenient *lingua franca*, and is intelligible to vast sections of the population of South Asia. For this reason it is the preferred medium of the Indian film industry, to which many well-established Urdu writers contribute scripts and especially songs, the lyrics of which frequently follow the conventions of classical Urdu poetry.

*General considerations.* From its earliest stages, Urdu has been strongly influenced by Persian, which, after the Muslim conquests of India in the 12th and 13th centuries C.E., was employed in the administration of the courts of Delhi and elsewhere. It has always been written in an adapted form of the Persian script (often referred to as the “Perso-Arabic script”). The distinctive *nasta’liq* style of writing, which is still employed, reflects the calligraphy found in medieval manuscripts. The *nask* style, now the form most commonly employed in Arabic and Persian printed works, has never been favored by publishers of Urdu, although it has occasionally been experimented with, and



modern computer fonts have been devised to reflect the handwritten form of *nasta'liq*. Indeed, until the advent of the word processor most Urdu books and newspapers were written entirely by hand, and until the last decades of the 20th century the calligrapher (*kāteb*) played a crucial role in society.

Urdu literature, the first substantial works of which date from the middle of the 16th century, has always been heavily influenced by Persian models, and although a small number of its most prominent writers, especially during the twentieth century have been Hindus and Sikhs, the overwhelming majority have been, and still are, Muslims. Until the beginning of the 19th century, the most substantial part of Urdu literature consisted of verse, while Persian, the language of administration, dominated prose writing. The most favoured poetic genres, as elsewhere in the Islamic world, were the *gāzal*, the *matnawī*, and the *qaṣīda*. Even now, in India and Pakistan and in other parts of the world where South Asian communities have migrated, Urdu poetry plays a significant role, and the *muṣā'ira* (Pers. *moṣā'era*), a gathering of poets, who recite their compositions according to strict traditional conventions, can attract audiences of thousands, members of which may or may not have Urdu as their mother tongue. It is undoubtedly because of its poetry that Urdu, which at present, for political and other reasons, faces many difficulties, continues to flourish and is universally regarded as *širin* “sweet,” an epithet granted even by its most vehement opponents.

*History of the language.* Urdu is a member of the Indo-European family of languages; and, like most of those spoken in the northern half of the subcontinent, it belongs to the Indo-Aryan subgroup. Ultimately it is derived from Sanskrit, the classical language of India, or more accurately from its colloquial form, known as Prakrit, which is commonly referred to as “Middle Indo-Aryan” (MIA). The modern languages of north India, such as Bengali, Gujarati, and Panjabi, to which Urdu and Hindi are fairly closely related, emerged from their MIA parents around 1000 C.E., but little evidence exists for their earliest form. Urdu developed from the regional speech of Delhi, which at the time of the Muslim conquests was known variously as *Kharī Bolī* “the upright speech,” *Hindī* or *Hindavī* “Indian” (as distinct from Persian), or simply as *zabān-e dihlavī* “language of Delhi.” *Kharī Bolī* was most closely related to other languages spoken in the vicinity of Delhi, such as Haryani, eastern Panjabi, and western Braj *Bhāshā*, a language found in and around [Agra](#) (Schmidt, 2004, p. 289; Shackle and Snell, 1990, p. 24).

The earliest examples of *Kharī Bolī* date from the period soon after the Muslim

conquests in 1192 (on which see [GHURIDS](#)). They consist mainly of short utterances and stray quotations in the hagiographies of sufi preachers. From these it is obvious that at this early stage a considerable quantity of Persian (and through Persian, Arabic) words were being freely employed in the languages of the native population. The first substantial examples of vernacular writing are widely considered to be the ‘Hindavī’ (Hendavi) verses ascribed to the great Delhi Persian poet, [Amir Kōsrow](#) (d. 1325). There are, however, good reasons to doubt the authenticity of these poems, which may belong to a much later period.

During the first two decades of the 14th century, the rulers of the [Delhi Sultanate](#) subjugated much of the [Deccan](#). With the arrival of their armies and that of the sufis, who followed closely in their wake, the language of Delhi began to assume the role of a *lingua franca* among the peoples of the conquered regions, who spoke a number of diverse languages, both Indo-Aryan and Dravidian. In Gujarat this highly convenient common tongue was referred to as “Gujrī,” and in the Deccan it acquired the local name “Dakanī” or “Dakhinī” (“southern”), by which the inhabitants of Hyderabad still call their distinctive form of Urdu. Little by little, Dakanī and Gujrī acquired a substantial corpus of religious verse literature, mainly adapted from standard Arabic and Persian texts. As literary activity grew, a certain amount of linguistic standardization took place, and with it the number of Arabic and Persian words that could safely be used increased. From the end of the 16th century the rulers of the southern sultanates, which were now virtually independent Muslim kingdoms, in their desire to assert their separate identity from the Mughal dynasty of Agra and Delhi, began to patronize Dakanī poets and writers, even though Persian remained the language of court administration (Matthews, 1994, pp. 91-93)

While Urdu in its distinctive Dakanī form flourished in the kingdoms of the south, the courts of Agra and Delhi clung rigidly to their Persian tradition, but poets, who sought their patronage, became increasingly interested in composing in their own native tongue. At first a style known as *rēkta* (*rikta*) “composite” was developed, in which one line of a couplet was written in Persian and the second line in “the language of Delhi.” The term *rēkta* soon became synonymous with the language itself and was employed as yet another name for it (Schmidt, 2004, p. 289).

Towards the end of the 17th century the Mughal emperor, Aurangzeb (r. 1658-1707), conquered the Deccan; after his death in 1707 poets and writers



from the provinces, always in search of patronage, began to migrate to Delhi, which, in spite of the subsequent decline in Mughal power, was still regarded as the imperial capital. The people of Delhi seem to have welcomed the freshness of the verse brought in by the poets of the south. To some extent the language they brought, which still retained its bewildering variety of names, replaced Persian as the major poetic medium of the Muslim élite. Persian, however, never lost its status, and even at the beginning of the 20th century the renowned poet-philosopher, [Muhammad Iqbal](#) (d. 1938), considered it the most suitable medium for his serious verse.

By the middle of the 18th century, at the hands of the great Delhi poets, the language (still officially unnamed) acquired its classical form, which, give or take a few minor archaisms, differed very little from the Urdu spoken and written at the present time. The finest practitioners of verse gathered around the area of the capital known as the *Urdū-e Mu'allā* (*Ordū-ye Mo'allā*) "The Exalted Army Camp," and the language they were busy refining came to be known as *zabān-e urdū-e mu'allā*. From this it finally acquired its present name "Urdu," which is first attested in a verse of the Delhi poet, Moṣṣāfi, composed somewhere around 1780 (Schmidt, 2004, p. 289). The new name soon gained currency, even though the British, who at that time were nurturing their imperial designs, persisted in calling it "Hindustani," yet another word for "Indian" (Schmidt, 2004, pp. 290-91).

As the British began to consolidate their rule, they soon realised that Persian was no longer the most viable means of communication in India; and at first they chose Urdu, written in the Persian script (the language which had been fostered by the Muslim élite), as the foremost medium of their administration. From the middle of the 19th century, however, increasingly vehement demands for the recognition of Sanskritized Hindi, written in the *devanāgarī* script, came from the Hindus, who formed the overwhelming majority of the population. Hindi was thus basically formed by replacing Persian loans with words taken directly from or derived from Sanskrit. As the Hindi movement gained momentum, Muslim leaders could plainly see that the prestige hitherto enjoyed by Urdu stood in danger of being eroded. At first, however, the British did not seriously question the status of Urdu in the northern Indian provinces, nor did they ignore the patronage it received from the rulers of some of the Princely States, especially from the Nizam of Hyderabad. The language issue nevertheless remained an important factor in the politics and the communal riots, which led up to independence and the eventual partition of India and

Pakistan in 1947. Inevitably, the provinces of northern India, traditionally regarded as the homeland of Urdu, passed from the control of the Muslims into the hands of the Hindu middle class.

Since Independence, in Pakistan there has been relatively little opposition to Urdu being adopted as the “national language” of the country, and in spite of the growing pressure from English as the preferred medium of education, there seems little doubt that Urdu will continue to thrive there. In India, however, where Urdu has been somewhat anomalously relegated to the position of the state language of Jammu and Kashmir, its position seems much more precarious (Matthews, 2003, pp. 61 ff.). Even though, after constant demands from its largely Muslim adherents, Urdu has recently been recognized as the “second official language” of Uttar Pradesh and Delhi, the fact remains that in most of north India all elementary and secondary education is imparted in Hindi. In spite of these impediments, there is still considerable optimism in both India and Pakistan, and no one seriously doubts that Urdu can continue to flourish in both countries for the foreseeable future at least.

*Urdu and Persian.* After the Muslim conquests of India, the languages spoken in and around Delhi rapidly absorbed a large amount of vocabulary from Persian, which became the primary medium of administration and belles lettres. The two major categories from which loanwords were acquired were those of nouns and adjectives, which quickly replaced their native equivalents. Such loans can be found not only in Urdu and Hindi, but also in neighboring languages such as Panjabi, Sindhi, Gujarati, and Bengali. A few examples from Urdu and Hindi will suffice to illustrate the range of vocabulary, which has now become completely naturalized: e.g., *dost* “friend,” *mez* (NPer. *miz*) “table,” *šahr* “city,” *nān* “bread,” *gošt* “meat,” *sabzī* “vegetables,” *garm* “hot,” *tāza* “fresh.” These languages were not only influenced by lexicon, but also by Persian syntax, which bears a strong resemblance to that of the modern Indo-Aryan languages.

The Persian of South Asia has retained a number of archaic and locally determined features, which distinguish Indo-Persian from that spoken in present-day Iran. Although the morphology and syntax of the two styles are virtually identical, the phonology exhibits a number of crucial differences. These are most apparent in the phonology of the vowels, where Indo-Persian has retained the archaic system, which corresponds exactly to that of Urdu and Hindi:



Indo-Persian:	<i>a</i>	<i>ā</i>	<i>i</i>	<i>ī</i>	<i>u</i>	<i>ū</i>	<i>ē ey</i>	<i>ō aw</i>
Modern Persian:	<i>a</i>	<i>ā</i>	<i>e</i>	<i>ī</i>	<i>u</i>		<i>ey</i>	<i>ow</i>

Final -â (i.e., -a followed by “silent h”) as in *tâzâ* (in Mod. Per. pronounced *tâze*) is pronounced -â. Hence *dânâ* “seed” and *dânâ* ‘wise’ are homophonous. Final -ân, -în and -ûn, as in *mardân*, *dîn*, and *čûn* are frequently nasalized: *mardân*, *dîñ*, *cûñ*. All the consonants are identical to those of Modern Persian with the exception of the plosive /q/, which in the best forms of speech retains its Arabic value. Many speakers, however, pronounce /q/ as /k/. Whereas in Modern Persian stress is usually on the final syllable, in Indo-Persian (as in Urdu) it is variable. For example:

Indo-Persian:	<i>‘qalam</i>	<i>ḵub‘şurat</i>	<i>‘muşkil</i>	<i>‘tâza</i>
Modern Persian:	<i>qa‘lam</i>	<i>ḵubşu‘rat</i>	<i>moş‘kel</i>	<i>tâ‘ze</i>

Although many people in the subcontinent are able to speak Persian well, it is largely these phonological differences that prove to be a barrier in communication. The difference between Modern Persian and Indo-Persian has often been likened to that between British English and ‘Indian English’. In the most simple terms it might be said that by and large Urdu-speakers pronounce Persian as if it were Urdu, making little or no concession to Modern Persian phonology or intonation. Recently, however, it has been advocated (by the Iqbal Academy of Pakistan, for example) that efforts should be made to pronounce the verse of Iqbal, and hence that of all Indo-Persian poets, according to the rules of Modern Persian. Not only this, but it has also been suggested that the Iranian style of calligraphy, rather than the traditional “Indian-style” *nasta‘liq* should be employed in future editions of his works.

The phonological system of Indo-Persian has determined the pronunciation of Persian loanwords taken into Urdu. The commonly accepted transcription for Urdu vowels, which incidentally is that adopted for the transliteration of Persian by Steingass in his *Persian-English Dictionary* (repr., 1963), is as follows: *a ā i ī u ū e ai o au*.

Nasality is indicated with -ñ. This transcription will be used for Urdu words cited in the following remarks.

Hence in Urdu we find, e.g., *mez*, *gošt*, *muşkil*, *jâved*, *baččagāñ*, *kohistān*. There are only a few exceptions, largely proper names and geographical terms,



where Modern Persian pronunciation is followed. In contemporary Urdu, the forms *tabrīz* and *čīngīz* have to some extent replaced the expected *tabrez* and *čīngez*. On the other hand more familiar names such as *aurangzeb*, *jāved*, *jahān-ārā* retain their original Indo-Persian pronunciation.

In terms of morphology, Persian loans are treated as if they were Urdu words, inflecting for gender, number, and case,

Urdu nouns are assigned to two genders, masculine and feminine. Gender has been given to loanwords by employing a number of criteria. Nouns denoting males and females are naturally assigned to the masculine and feminine genders respectively: *pisar* “son” (m); *duktar* “daughter” (f). Many categories of nouns which are feminine in Arabic retain the same gender in Urdu. For example, nouns of Arabic origin terminating in *-at* are regarded as feminine in Urdu: *daulat* “wealth,” *kāšīyat* “speciality,” *zarūrat* “necessity”; cf. the hybrid form *šahrīyat* “citizenship.” All Arabic nouns with the form *tafīl* are treated as feminine: *tašvīr* “picture,” *takmīl* “conclusion,” *tadbīr* “plan.” Many other nouns take their gender from an underlying Sanskrit or Hindi synonym. For example, *kitāb* (masculine in Arabic) is feminine in Urdu cf. Hindi *pothī* and Sanskrit *pustak* “book,” which are both feminine.

Urdu possesses two numbers (singular and plural) and three cases (direct, oblique, and vocative), which are marked by special terminations, e.g., *khānā* “food, dish” (sing, direct), *khāne* (pl. direct) *khānoñ* (pl. oblique). The Persian loan *bačča* “child” behaves in the same way: *bačča*, *bačče*, *baččoñ*

Persian plurals are used in more formal styles of speech and writing; but, with the exception of some very common lexically conditioned items, such as *bārḥā* “time and time again,” *hazārḥā* “thousands of,” these are rarely employed in colloquial speech.

In Urdu adjectives precede the noun they qualify, and this also applies to Persian loans: *buzurg mard* “a senior man.” Only in formal writing or verse do they follow the noun joined to it by the *ezāfe* (*izāfat*): *mard-e buzurg*, *masala-e sanglāk* “thorny problem.” In general, Persian comparative and superlative forms function as intensives: *bihtar* “very good,” *bihtarīn* “really excellent.”



The Persian pronoun *kʷod* (Urdu *kuḍ*) and the pronominal adjectives *čand*, *har*, and *baʒ* have been completely assimilated and are of frequent occurrence in both Urdu and Hindi: *čand log* “a few people,” *har baras* “every year,” *ham kuḍ* “we ourselves.” Persian personal pronouns are only found in set expressions, one of the most frequent of which is *jān-e man* “my darling.”

Persian syntax has also had a strong influence on Urdu. Among the most notable features are the *ezāfe* (*izāfat*) and copula constructions, prepositional phrases, phrase verbs and subordinate clauses.

Rules governing the use of the *ezāfe* and the copula /o/ are similar to those in Persian. Phrases such as *barr-e šaḡīr-e hind o pāk* “the subcontinent of India and Pakistan,” *ḥukūmat-e panjāb* “the government of Panjab,” *āb o havā* “climate” are of common occurrence in all styles of Urdu. Strictly speaking, such constructions should only be employed with words of Persian origin, but phrases involving a Persian word linked to a word of Indo-Aryan origin are not uncommon, e.g., *qābil-e bharosā* “worthy of trust,” which is modeled on *qābil-e iʿtimād*.

A number of Persian prepositions, notably *az*, *ba* (*be*), *bar*, and *dar*, occur in a host of set phrases which have been taken into Urdu. Examples are: *az sar-e nau* “afresh,” *kam az kam* “at least,” *roz baroz* “daily,” *bar zabān* “by heart,” *dar aṣl* “in fact.”

Certain prepositional phrases are very common in the modern press, such as *baʿd azān* “afterwards,” *darīn aṭnā* “meanwhile.”

Phrase verbs, which are formed in Persian with a noun or adjective construed with *kardan* or *šodan* are formed in exactly the same way in Urdu with the verbs *karnā* “to do” and *honā* “to be”: *taiyār karnā* “to prepare,” *band o bast karnā* “to arrange,” *kāmyāb honā* “to succeed,” *band honā* “to be closed.”

In the formation of subordinate clauses, the medieval Indo-Aryan languages dispensed with many of the conjunctions derived from Sanskrit. In order to facilitate the writing of official documents, these were later replaced, mainly by borrowing from Persian. Many of the conjunctions found in all styles of



Urdu and Hindi are therefore of Persian origin. Examples are *agar* “if,” *agarče* “although,” *čūnki* “since,” *hālānki* “although,” *ki* “that.”

Throughout the long period of its existence, Urdu has at all levels been greatly indebted to Persian, and this is especially apparent from its literature. Urdu verse is usually replete with ‘Persianisms’ that would neither appear foreign nor would cause any difficulty to the average Urdu speaker. A typical verse taken from a poem of Iqbal, may serve as an illustration of the immense influence that Persian still exerts on the language:

*huā kaimazan kāravān-e bahār / iram ban gayā dāman-e kuhsār*

“The caravan of spring has pitched its tent / The mountainside has become Iram.”

Of the eight words in the verse, only the two verbs *huā* = *buda* (*ast*) and *ban gayā* = *šoda*(*ast*) are of Indo-Aryan origin. The rest, including the metrical form of the verse, remains Persian (see also Thiesen, pp. 179-209).

See also [INDIA xviii. PERSIAN ELEMENTS IN INDIAN LANGUAGES.](#)

## BIBLIOGRAPHY

---

The following works all contain references to the history and development of Urdu and its relationship with Persian, Hindi and other languages. The most comprehensive account is Schmidt, 2004, which also contains an extensive bibliography. The introduction to Shackle and Snell, 1990, contains much information on the influence of Arabic, Persian, Sanskrit and English on Hindi and Urdu.

M. A. R. Barker et al., *Spoken Urdu*, 3 vols., Ithaca, N.Y., 1975.

Ather Farouqi: “Future Prospects for Urdu in India,” *Mainstream Annual*, 1992, pp. 36-43.



D. J. Matthews, *Dakani Language and Literature*, Ph.D. diss., SOAS, University of London, 1976.

Idem, "Eighty Years of Dakani Scholarship," *Annual of Urdu Studies* 9, 1994, pp. 91-107.

Idem, "Urdu Language and Education in India," *Social Scientist* 31, 2003, pp. 57-72.

D. J. Matthews and M. K. Dalvi, *Teach Yourself Urdu*, Sevenoaks, Kent, UK, 1999.

D. J. Matthews, C. Shackle, and S. Husain, *Urdu Literature*, Islamabad, 2003.

R. S. McGreggor, *Outline of Hindi Grammar*, New Delhi, 1972.

John T. Platts, *A Dictionary of Urdu, Classical Hindi and English*, repr., Lahore, 1994.

Ruth Laila Schmidt, *Dakhini Urdu. History and Structure*, New Delhi, 1981.

Idem, *Urdu. An Essential Grammar*, London, 1999.

Idem, "Urdu," in *The Indo-Aryan Languages*, ed. George Cardona and Dhanesh Jain, London, 2004, pp. 286-350.

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

Rupert Snell and Simon Weightman, *Teach Yourself Hindi*, Sevenoaks, Kent, UK, 1992.

F. Steingass, *A Comprehensive Persian-English Dictionary*, repr., London, 1963.

Finn Thiesen, *A Manual of Classical Persian Prosody, with Chapters on Urdu, Karakhanidic and Ottoman Poetry*, Wiesbaden, 1982.