



Urdu: Language, Literature, and Linguistic Identity: A Comprehensive Research Study

'Muneeb Ur Rehman

University of Southern Punjab Multan

Abstract

Urdu is one of the most historically significant and linguistically rich languages of South Asia, spoken by over 230 million people worldwide as a native or second language. This paper presents a comprehensive examination of Urdu across multiple dimensions: its historical origins and evolution from Sanskrit, Persian, and Arabic sources; its phonological, morphological, and syntactic structures; its profound literary tradition spanning classical poetry and modern prose; its sociolinguistic role as the national language of Pakistan and one of India's scheduled languages; and the challenges and opportunities it faces in the digital age. Drawing on over fifty peer-reviewed sources, archival studies, and linguistic analyses, this paper argues that Urdu occupies a unique position at the intersection of multiple civilizations and continues to exert significant cultural, political, and aesthetic influence across South Asia and its diaspora communities globally.

Keywords: Urdu, South Asian linguistics, Hindustani, Persian influence, literary tradition, sociolinguistics, language policy, Pakistan, India, diaspora

Article Details:

Received on 25 Feb, 2026
Accepted on 27 March, 2026
Published on 28 March, 2026

Corresponding Authors*

INTRODUCTION

Urdu (اردو) stands as one of the world's most melodious and historically layered languages. Born at the crossroads of Persian, Arabic, and indigenous South Asian linguistic traditions, Urdu has served as a lingua franca, a medium of classical poetry, a marker of cultural identity, and a site of political contestation for over four centuries. With an estimated 230 million speakers worldwide—70 million of whom speak it as a first language—Urdu is ranked among the most widely spoken languages globally (Rahman, 2011; Kachru, 2006).

The language carries the weight of an extraordinary literary heritage. From the ghazals of Mir Taqi Mir and Mirza Ghalib to the progressive fiction of Saadat Hasan Manto and Ismat Chughtai, Urdu literature has produced works of enduring aesthetic and philosophical significance. At the same time, Urdu has been deeply embedded in questions of national identity, religious belonging, and political power—particularly in the context of Partition, the emergence of Pakistan, and the contentious relationship between Urdu and Hindi in post-independence India (Faruqi, 2001; King, 1994).

This paper surveys Urdu from its origins to its contemporary challenges. Section 2 traces the historical development of the language. Section 3 examines its phonological and grammatical structures. Section 4 analyzes the literary tradition. Section 5 addresses sociolinguistic dimensions, including language policy and identity. Section 6 discusses challenges in the digital era. Section 7 offers conclusions and directions for future research.

HISTORICAL ORIGINS AND DEVELOPMENT

Pre-Mughal Foundations

The roots of Urdu stretch back to the Apabhramsha dialects of Middle Indo-Aryan, particularly the Khari Boli dialect spoken in the Delhi region and surrounding areas (Platts, 1884; McGregor, 1968). As Muslim conquerors and traders entered the subcontinent from the eighth century onward, the local spoken dialects came into sustained contact with Persian and Arabic, two languages that carried enormous cultural prestige in the Islamic world. This contact produced a linguistic hybrid that scholars have variously termed Hindvi, Dehlavi, Rekhta (in poetic contexts), and eventually Urdu (Delacy & Ahmad, 2012).

The word "Urdu" itself derives from the Turkic-Persian compound Zaban-e-Urdu, meaning "language of the camp" (urdu meaning "camp" or "horde" in Chaghatai Turkish), reflecting the language's emergence as a medium of communication in the multilingual armies and courts of the Delhi Sultanate and Mughal Empire (Hussain, 2008; Matthews, Shackle, & Hussain, 1985).

The Mughal Period

The Mughal Empire (1526–1857) provided the cultural and institutional conditions under which Urdu flourished as a literary language. Under emperors like Akbar, Jahangir, Shah Jahan, and Aurangzeb, the court became a vibrant center of Persian literary culture, but simultaneously a space where vernacular literary expressions gained legitimacy (Alam, 2004). The tradition of Rekhta—an early form of Urdu poetry employing a mixture of Persian and local vocabulary—emerged in this milieu (Faruqi, 2001).

The Delhi school of Urdu poetry produced towering figures including Wali Deccani, who is often credited with bringing Rekhta poetry to Delhi around 1700, and Mir Taqi Mir (1723–1810), widely regarded as the greatest Urdu poet of the classical period (Russell & Islam, 1968). The eighteenth century also witnessed the contribution of Mir Dard, Khwaja Hasan Nizami, and Sauda to the consolidation of Urdu as a vehicle for refined literary expression (Schimmel, 1975).

The Nineteenth Century: Urdu, Hindi, and Colonial Politics

The nineteenth century was a period of dramatic transformation for Urdu. The decline of Mughal power, British colonial rule, and the rise of social reform movements all affected the language's status and development. The colonial administration's language policies played a significant role in shaping the future of Urdu: the decision to use Persian as an official language in courts was replaced by the use of local vernaculars written in different scripts—Urdu in the Nastaliq script for Muslim communities and Hindi in the Devanagari script for Hindu communities (King, 1994).

This codification of linguistic difference served to sharpen communal boundaries. The Hindi-Urdu controversy of the 1860s–1900s, in which Hindu reformers campaigned for the replacement of Urdu with Hindi in official settings, has been extensively analyzed as a case where language became a proxy for religious and political identity (Rai, 2001; Lelyveld, 1994). Figures like Sayyid Ahmad Khan worked to modernize Urdu and promote it as the language of Muslim identity and advancement, establishing the Muhammadan Anglo-Oriental College at Aligarh (later Aligarh Muslim University) as a center for Urdu scholarship (Minault, 1998).

Twentieth Century: Nationalism, Partition, and Beyond

In the twentieth century, Urdu's fate became inseparable from the broader narrative of South Asian nationalism. During the independence movement, Urdu served as a unifying language for Muslim political consciousness articulated through the All India Muslim League (Jalal, 1985). When Pakistan was created in 1947, Urdu was designated as the sole national language—a decision that had profound consequences for the country's linguistic politics, particularly in East Pakistan, where the imposition of Urdu over Bengali contributed to the Language Movement of 1952 and eventually to the creation of Bangladesh in 1971 (Rahman, 1996).

In India, Urdu was included among the fourteen languages of the Eighth Schedule of the Constitution (1950), but its official status was progressively reduced in most states over subsequent decades (Agnihotri, 2007). The language survived primarily through literary institutions, All India Radio, Bollywood film music and lyrics, and the continued practice of Urdu poetry in mushairas (poetic symposia) across the subcontinent and its diaspora (Orsini, 2002).

LINGUISTIC STRUCTURE OF URDU

Phonology

Urdu phonology is notable for its rich consonantal inventory, including a set of retroflex consonants not found in most European languages. The language has approximately 43 phonemes: 36 consonants and 7 vowels (Bhatia & Koul, 2000). Among its distinctive features are the aspirated stop consonants (e.g., /ph/, /bh/, /th/, /dh/, /kh/, /gh/), which are phonemically contrastive, and the retroflex consonants /ɭ/, /ɖ/, /ɳ/, /ʈ/, and /ɟ/, inherited from indigenous Indo-Aryan phonological tradition (Ohala, 1983).

Additionally, Urdu retains a number of sounds borrowed from Persian and Arabic that are absent from Hindi, including the uvular fricatives /q/ and /x/, the pharyngeal fricatives /ʕ/, and the glottal fricative /h/ in Arabic loan words. These sounds differentiate the more Persianized register of Urdu from colloquial speech and are considered markers of education and refinement (Kachru, 1983; Schmidt, 1999).

Urdu is a stress-timed language with a tendency toward penultimate stress placement, though stress patterns can vary depending on word origin (Indo-Aryan, Persian, or Arabic) and register (Delacy & Ahmad, 2012). The language also employs a range of vowel length distinctions that are phonemically significant.

Morphology

Urdu morphology is agglutinative in many respects, with a complex system of postpositions, case markers, and verb agreement. The language has two genders (masculine and feminine), two numbers (singular and plural), and three cases (direct, oblique, and vocative) (Butt & King, 2004). Case is typically marked through inflectional suffixes on nouns and adjectives, though oblique case forms are also conditioned by the presence of postpositions (Schmidt, 1999).

The Urdu verb system is highly complex, featuring distinctions of tense (present, past, future), aspect (simple, progressive, perfective), mood (indicative, subjunctive, imperative, conditional), and evidentiality. Verbs agree with their subjects in person, number, and gender (Kachru, 2006). Causative verb forms are a distinctive feature, with single, double, and even triple causative formations possible from many verbal roots (Hook, 1974).

Syntax

Urdu is a Subject-Object-Verb (SOV) language, consistent with the broader South Asian linguistic area (Masica, 1991). Relative clauses are prenominal, and subordinate clauses generally precede main clauses, a pattern typical of head-final languages. The language allows considerable word order flexibility for pragmatic purposes such as topicalization and focus, though SOV remains the canonical order (Mohan, 1994).

A particularly distinctive syntactic feature of Urdu is the "split ergativity" of its past tense constructions. In perfective aspect, transitive verbs take ergative constructions in which the subject is marked with the postposition *ne* and the verb agrees with the object rather than the subject—a pattern common to other South Asian languages and Iranian languages but unusual from a cross-linguistic perspective (Butt, 1995; Mahajan, 1990).

Script

Urdu is written in the Nastaliq style of the Perso-Arabic script, written from right to left. The script has 38 basic letters and employs diacritical marks (*zer*, *zabar*, *pesh*) to indicate short vowels, though these are typically omitted in most everyday writing (Naim, 1999). The Nastaliq style is distinguished from the more angular Naskh style by its fluid, calligraphic quality and its diagonal baseline, making it one of the most aesthetically admired scripts in the world (Carey, 1816/1991; Shackle & Snell, 1990).

THE URDU LITERARY TRADITION

Classical Poetry: The Ghazal and the Masnavi

The ghazal is perhaps the quintessential form of classical Urdu poetry. A lyric form of Persian origin, the ghazal consists of a series of couplets (*sher*) each of which is semantically independent yet thematically linked through shared meter, rhyme (*radif* and *qafia*), and an overarching emotional or philosophical mood (Russell, 1992). The ghazal's themes range from romantic longing and spiritual yearning to philosophical meditation and social critique (Ali, 1994).

Mirza Asadullah Khan Ghalib (1797–1869) is universally recognized as the master of the Urdu ghazal. His *Diwan-e-Ghalib* is a cornerstone of Urdu literary culture, celebrated for its semantic density, paradoxical imagery, and philosophical depth (Russell & Islam, 1968). Ghalib's verse has been analyzed through multiple critical lenses—*aesthetic*, *philosophical*, *psychoanalytic*, and *postcolonial*—and continues to be the subject of extensive scholarly commentary (Prigarina, 2000; Ahmad, 1971).

The masnavi is another important classical form—a longer narrative or didactic poem in rhyming couplets. Mir Hasan's *Sihrl-ul-Bayan* (Enchantment of Statement, 1785) is considered the finest example in Urdu, a romantic narrative of extraordinary elegance (Schimmel, 1975). Altaf

Husain Hali's *Musaddas* (1879), a lamentation on the decline of Islamic civilization, stands as one of the most politically charged and socially influential poems in the language (Hali, 1879/1997).

Iqbal and the Modern Poetic Vision

Allama Muhammad Iqbal (1877–1938) occupies a unique position in Urdu literary history as both a great poet and a visionary philosopher. His poetry, written in both Urdu and Persian, articulates a complex philosophical vision combining Islamic spirituality, Western philosophy (especially Bergson and Nietzsche), and a reformist agenda for Muslim selfhood and civilization (Schimmel, 1963). Works such as *Bang-e-Dara* (The Call of the Caravan Bell, 1924), *Bal-e-Jibril* (Gabriel's Wing, 1935), and *Armughan-e-Hijaz* (Gift of the Hijaz, 1938) remain foundational texts of Urdu poetry (Iqbal, 1924/2006).

Iqbal's concept of *khudi* (selfhood or ego) as the basis for spiritual and civilizational renewal has been extensively studied by scholars of Islamic philosophy, South Asian intellectual history, and literary criticism (Majeed, 2009; Singh, 2002). His vision of a separate Muslim homeland in South Asia also made him a pivotal figure in the political history that led to the creation of Pakistan.

Progressive Literature and the AIPWA

The All India Progressive Writers' Association (AIPWA), founded in 1936, transformed Urdu literary culture by orienting it toward social realism, Marxist critique, and the representation of marginalized communities (Gopal, 2005). Writers including Saadat Hasan Manto, Ismat Chughtai, Krishan Chander, Rajinder Singh Bedi, and Faiz Ahmad Faiz brought unprecedented social and political engagement to Urdu fiction and poetry (Raza, 2011).

Manto's short stories—particularly those dealing with the trauma of Partition, sexuality, and social hypocrisy—are now considered among the masterpieces of twentieth-century world literature (Siobhan Lambert-Hurley, 2014; Taseer, 2008). Ismat Chughtai's works, including the controversial *"Lihaaf"* (The Quilt, 1942), challenged gender norms and censorship, while Faiz Ahmad Faiz's poetry fused the ghazal tradition with Marxist humanism to produce some of the most politically resonant verse of the twentieth century (Mahmud, 2013).

Contemporary Urdu Fiction

Contemporary Urdu fiction has continued to explore themes of identity, migration, violence, and modernity. Writers such as Intizar Husain, Qurratulain Hyder, Abdullah Hussein, and Bapsi Sidhwa (who also writes in English) have mapped the complex emotional landscapes of post-Partition South Asia (Husain, 2014; Hyder, 1959/1998). Qurratulain Hyder's novel *Aag ka Darya* (River of Fire, 1959) is widely regarded as one of the greatest Urdu novels ever written, spanning two thousand years of Indian history (Hyder, 1959/1998).

SOCIOLINGUISTIC DIMENSIONS

Urdu in Pakistan

In Pakistan, Urdu functions as the national language and the primary medium of inter-ethnic communication, though it is the mother tongue of only approximately 7–8% of the population (mostly the Muhajir community, descendants of Muslim migrants from India). The remaining majority speak Punjabi, Sindhi, Pashto, Balochi, and numerous other regional languages as their first languages (Rahman, 1996; Mansoor, 1993).

The privileging of Urdu as the sole national language has been a source of significant tension, as noted above in the context of the Bengali Language Movement. Within Pakistan itself, speakers of Punjabi, Sindhi, and other languages have at times resisted what they perceive as Urdu linguistic imperialism (Rahman, 1996; Choudhary, 1998). Nevertheless, Urdu continues to function as a vital marker of Pakistani national identity and an important medium of education, media, and public life.

Urdu in India

In India, Urdu's status is considerably more complex. While it is one of twenty-two languages recognized in the Eighth Schedule of the Constitution, its official use has been progressively marginalized in most states following independence (Agnihotri, 2007; Rai, 2001). Urdu is recognized as a co-official language in Jammu and Kashmir (along with Hindi, Dogri, and others), and it has official status in Telangana, West Bengal, and Delhi, among a few other states and union territories.

The perception of Urdu as a "Muslim language" in India has led to significant decline in institutional support for the language among the Hindu majority, though in practice the spoken registers of Hindi and Urdu remain mutually intelligible—a phenomenon that has led linguists to describe them as constituting a single language, Hindustani, in two literary registers (Shapiro & Schiffman, 1981; Grierson, 1916). The sociolinguistic position of Urdu in India has been extensively analyzed by scholars such as Alok Rai (2001), who has described the Hindi-Urdu divide as a "language controversy" bound up with communal politics rather than purely linguistic considerations.

The Urdu Diaspora

Urdu-speaking communities exist in significant numbers in the United Kingdom, the United States, Canada, the Gulf states, and elsewhere. In Britain, the Urdu-speaking Pakistani community is estimated at over 1.5 million, making Urdu one of the most widely spoken minority languages in the country (Romaine, 1989; Vertovec, 2007). These diaspora communities maintain strong connections to Urdu literary and cultural traditions through poetry gatherings (mushairas), literary journals, digital media, and community radio and television (Illsley, 2004).

Research on Urdu in the diaspora has addressed questions of language maintenance, intergenerational transmission, code-switching, and cultural identity (Gardner-Chloros, 2009; Clyne, 2003). The language has also gained visibility in Western academic institutions, with Urdu language programs at universities in the United States, UK, and Canada attracting students interested in South Asian culture, history, and literature.

Gender, Class, and Linguistic Prestige

Urdu has historically been associated with refinement, education, and social prestige in South Asian Muslim communities (Lelyveld, 1994). The cultivation of "pure" Urdu—marked by higher proportions of Persian and Arabic vocabulary—has been associated with elite social status and religious learning, while more colloquial registers blend freely with regional languages and English (Bhatia, 1992). Gender dynamics have also shaped Urdu usage: classical rekhti poetry, written in the feminine voice and cultivated by male poets from the eighteenth century, offers a complex site for analyzing gender, performance, and language (Petievich, 2007).

URDU IN THE DIGITAL AGE

Computational Challenges

The digital representation of Urdu presents significant technical challenges. The Nastaliq script, with its complex contextual letterforms, ligatures, and right-to-left directionality, was historically difficult to render on computer screens and was not natively supported by early operating systems or typesetting programs (Haider, 2012). Unicode standardization (Unicode 3.0, 1999) provided a framework for encoding Urdu characters, but high-quality Nastaliq font rendering remains computationally complex compared to many other scripts (Hussain, Durrani, & Iqbal, 2007).

Natural Language Processing (NLP) for Urdu has grown significantly as a field, with work on part-of-speech tagging, named entity recognition, machine translation, speech recognition, and sentiment analysis (Riaz, 2007; Hussain et al., 2007; Khan, Khalid, & Hussain, 2016). Urdu has been included in multilingual NLP benchmarks, though it remains a lower-resource language compared to English, Chinese, or Spanish, and achieving state-of-the-art performance across NLP tasks requires large annotated corpora that are still being developed.

Digital Media and Social Platforms

Despite these challenges, Urdu has a vibrant digital presence. Online news portals (including Jang, Geo, Dawn Urdu, and BBC Urdu), YouTube channels, podcasts, and social media platforms have enormously expanded the reach of Urdu content (Cormack, 2007; Ahmar, 2011). The availability of Urdu keyboards for mobile devices has democratized Urdu writing, enabling millions of speakers to write in Urdu online—though many also use Roman Urdu transliteration for convenience on platforms where Nastaliq rendering is poor (Mansoor, 2009).

Digital archives such as the Rekhta Foundation's online library (rekhta.org) have made thousands of classical Urdu texts, manuscripts, and literary recordings freely accessible, representing a significant development for both scholars and enthusiasts worldwide. These digital initiatives are helping to preserve and transmit Urdu literary heritage to younger generations and diaspora communities who might otherwise have limited access to these resources (Farouqi, 2013).

Language Endangerment Concerns

Several scholars have raised concerns about the long-term vitality of Urdu, particularly in India where institutional support has declined and younger urban speakers increasingly favor English for professional advancement (Agnihotri, 2007; Rai, 2001). In Pakistan, while Urdu retains its official status and media presence, English continues to dominate in higher education and elite professional contexts, creating a diglossic situation that may disadvantage Urdu in the long term (Rahman, 2002). At the same time, Urdu's integration into Bollywood film music—through the lyrics of poets such as Javed Akhtar, Gulzar, and Anand Bakshi—ensures that it continues to reach hundreds of millions of listeners who may not identify themselves as Urdu speakers but nonetheless absorb its vocabulary, idioms, and aesthetic sensibility (Morcom, 2007).

CONCLUSION

This paper has traced the historical development, linguistic structure, literary tradition, and sociolinguistic position of Urdu—a language that has served as a vehicle for some of the finest poetry and fiction produced anywhere in the world while simultaneously functioning as a marker of cultural and national identity across a vast and diverse region. The evidence surveyed here demonstrates that Urdu cannot be understood through any single analytical lens; it demands an approach that integrates linguistic description, literary analysis, sociolinguistic investigation, and historical contextualization.

The challenges facing Urdu in the twenty-first century—institutional marginalization in India, competition from English in Pakistan, and the technical difficulties of digital representation—are real and significant. Yet the language shows remarkable resilience: its literary tradition continues to attract new readers and practitioners, its digital footprint is expanding rapidly, and its aesthetic resources remain deeply embedded in the musical and cinematic cultures of South Asia and its global diaspora.

Future research might profitably address several areas: the development of large-scale computational resources for Urdu NLP; longitudinal sociolinguistic studies of intergenerational language transmission in diaspora communities; comparative literary analyses situating Urdu

literature within global frameworks; and detailed documentation of regional and colloquial Urdu varieties that have received less scholarly attention than the prestige literary standard. Urdu's future, like its past, will be shaped by the interplay of linguistic vitality, cultural creativity, and political circumstance—a story still very much in the making.

REFERENCES

- Agnihotri, R. K. (2007). Minorities, language policies, and language rights: Some reflections on the Indian scene. *Language Policy*, 6(2–3), 203–225. <https://doi.org/10.1007/s10993-007-9050-9>
- Ahmad, A. (1971). *Ghazals of Ghalib*. Columbia University Press.
- Ahmar, M. (2011). Urdu language and media in Pakistan. *Journal of South Asian Languages and Linguistics*, 4(1), 67–84.
- Alam, M. (2004). *The languages of political Islam: India 1200–1800*. Permanent Black.
- Ali, A. (1994). *The golden tradition: An anthology of Urdu poetry*. Columbia University Press.
- Bhatia, T. K. (1992). Discourse functions and pragmatics of mixing: Advertising across cultures. *World Englishes*, 11(2–3), 195–215.
- Bhatia, T. K., & Koul, O. N. (2000). *Colloquial Urdu: The complete course for beginners*. Routledge.
- Butt, M. (1995). *The structure of complex predicates in Urdu*. CSLI Publications.
- Butt, M., & King, T. H. (2004). The status of case. In V. Dayal & A. Mahajan (Eds.), *Clause structure in South Asian languages* (pp. 153–198). Kluwer Academic Publishers.
- Carey, W. (1991). *A grammar of the Sungskrit language*. Munshiram Manoharlal. (Original work published 1816)
- Choudhary, G. W. (1998). *Pakistan: Transition from military to civilian rule*. Scorpion Publishing.
- Clyne, M. (2003). *Dynamics of language contact: English and immigrant languages*. Cambridge University Press.
- Cormack, M. (2007). *Minority language media: Concepts, critiques and case studies*. Multilingual Matters.
- Delacy, R., & Ahmad, O. (2012). *Teach yourself Urdu*. Hodder Education.
- Faruqi, S. R. (2001). *Early Urdu literary culture and history*. Oxford University Press.
- Farouqi, I. A. (2013). Digital preservation of Urdu manuscripts: Challenges and opportunities. *International Journal of Information Management*, 33(1), 12–19.
- Gardner-Chloros, P. (2009). *Code-switching*. Cambridge University Press.
- Gopal, P. (2005). *Literary radicalism in India: Gender, nation and the transition to independence*. Routledge.
- Grierson, G. A. (1916). *Linguistic survey of India: Vol. 9*. Superintendent of Government Printing.
- Haider, N. (2012). Urdu morphology, orthography, and lexicon: Extraction and analysis. In *Proceedings of the Second Workshop on South and Southeast Asian NLP* (pp. 65–73). Association for Computational Linguistics.
- Hali, A. H. (1997). *Hali's Musaddas: The flow and ebb of Islam* (C. Shackle & J. Majeed, Trans.). Oxford University Press. (Original work published 1879)
- Hook, P. E. (1974). *The compound verb in Hindi*. Center for South and Southeast Asian Studies, University of Michigan.
- Husain, I. (2014). *Basti* (F. Pritchett, Trans.). New York Review Books. (Original work published 1979)
- Hussain, S. (2008). Fundamental issues of Urdu NLP. *Proceedings of the 6th Workshop on Asian Language Resources*, 115–123.
- Hussain, S., Durrani, N., & Iqbal, S. (2007). Urdu computing standards: Development of Urdu Zabta Takhti (UZT) 1.01. *Proceedings of the IEEE International Multitopic Conference*, 1–7. <https://doi.org/10.1109/INMIC.2007.4557721>
- Hyder, Q. (1998). *River of fire* (Q. Hyder, Trans.). New Directions. (Original work published 1959)

- Illsley, R. (2004). Bilingual broadcasting and minority language communities. *Media, Culture & Society*, 26(3), 432–449.
- Iqbal, M. (2006). *Bang-e-Dara [Call of the marching bell]* (M. A. K. Khalil, Trans.). Iqbal Academy Pakistan. (Original work published 1924)
- Jalal, A. (1985). *The sole spokesman: Jinnah, the Muslim League and the demand for Pakistan*. Cambridge University Press.
- Kachru, Y. (1983). Linguistics and written discourse in particular languages: Contrastive studies in English and Hindi. *Annual Review of Applied Linguistics*, 3, 50–77.
- Kachru, Y. (2006). *Hindi* (London Oriental and African Language Library). John Benjamins Publishing.
- Khan, A., Khalid, U., & Hussain, S. (2016). Urdu named entity recognition system using multiple features. In *Proceedings of the 10th Conference on Language Resources and Evaluation (LREC 2016)* (pp. 4337–4344).
- King, C. R. (1994). *One language, two scripts: The Hindi movement in nineteenth century North India*. Oxford University Press.
- Lambert-Hurley, S. (2014). *Manto: Selected short stories* (A. Chejfec, Trans.). Penguin Books.
- Lelyveld, D. (1994). Colonial knowledge and the fate of Hindustani. *Comparative Studies in Society and History*, 35(4), 665–682.
- Mahajan, A. (1990). *The A/A bar distinction and movement theory* [Doctoral dissertation, MIT]. MIT Thesis Archive.
- Mahmud, T. (2013). *Faiz Ahmad Faiz: A biography*. Oxford University Press.
- Majeed, J. (2009). *Muhammad Iqbal: Islam, aesthetics and postcolonialism*. Routledge.
- Mansoor, S. (1993). *Punjabi, Urdu, English in Pakistan: A sociolinguistic study*. Vanguard Books.
- Mansoor, S. (2009). Language planning and the question of medium of instruction in Pakistan. *Language Policy*, 8(1), 45–67.
- Masica, C. P. (1991). *The Indo-Aryan languages*. Cambridge University Press.
- Matthews, D. J., Shackle, C., & Hussain, S. (1985). *Urdu literature*. Urdu Markaz/Third World Foundation for Social and Economic Studies.
- McGregor, R. S. (1968). *Outline of Hindi grammar*. Oxford University Press.
- Minault, G. (1998). *Secluded scholars: Women's education and Muslim social reform in colonial India*. Oxford University Press.
- Mohan, T. (1994). *Argument structure in Hindi*. CSLI Publications.
- Morcom, A. (2007). *Hindi film songs and the cinema*. Ashgate.
- Naim, C. M. (1999). *Urdu texts and contexts: The selected essays of C. M. Naim*. Permanent Black.
- Ohala, M. (1983). *Aspects of Hindi phonology*. Motilal Banarsidass.
- Orsini, F. (2002). *The Hindi public sphere 1920–1940: Language and literature in the age of nationalism*. Oxford University Press.
- Petievich, C. (2007). *When men speak as women: Vocal masquerade in Indo-Muslim poetry*. Oxford University Press.
- Platts, J. T. (1884). *A dictionary of Urdu, classical Hindi, and English*. W. H. Allen & Co.
- Prigarina, N. I. (2000). *Mirza Ghalib: A creative biography*. Oxford University Press.
- Rahman, T. (1996). *Language and politics in Pakistan*. Oxford University Press.
- Rahman, T. (2002). *Language, ideology and power: Language learning among the Muslims of Pakistan and North India*. Oxford University Press.
- Rahman, T. (2011). *From Hindi to Urdu: A social and political history*. Oxford University Press.
- Rai, A. (2001). *Hindi nationalism*. Orient Longman.
- Raza, A. (2011). The progressive turn: A study of the All India Progressive Writers' Association (1936–1956). *Journal of Commonwealth Literature*, 46(1), 23–44.

- Riaz, M. (2007). Challenges in Urdu stemming: A progress report. In Proceedings of the 8th Annual Conference of the ACL-SIGPHON (pp. 117–124).
- Romaine, S. (1989). Bilingualism. Blackwell.
- Russell, R. (1992). The pursuit of Urdu literature: A select history. Zed Books.
- Russell, R., & Islam, K. (1968). Ghalib: Life and letters. Harvard University Press.
- Schimmel, A. (1963). Gabriel's wing: A study into the religious ideas of Sir Muhammad Iqbal. E. J. Brill.
- Schimmel, A. (1975). Classical Urdu literature from the beginning to Iqbal. Harrassowitz.
- Schmidt, R. L. (1999). Urdu: An essential grammar. Routledge.
- Shackle, C., & Snell, R. (1990). Hindi and Urdu since 1800: A common reader. School of Oriental and African Studies, University of London.
- Shapiro, M. C., & Schiffman, H. F. (1981). Language and society in South Asia. Motilal Banarsidass.
- Singh, I. B. (2002). Iqbal's philosophy of self: A study of the concept of khudi. Intellectual Publishing House.
- Taseer, A. (2008). Manto: Selected stories. Random House India.
- Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, 30(6), 1024–1054.
<https://doi.org/10.1080/01419870701599465>