

Arabic Loanwords in Urdu and Their Sanskrit/Prakrit Counterparts: A Linguistic Bridge Study

Dr. Nilofar Nadiyah

Assistant Professor
Department of Arabic
University of Mumbai, Mumbai, Maharashtra.

Abstract:

This research paper investigates the intersection of Arabic loanwords in Urdu with their corresponding Sanskrit and Prakrit counterparts, framing the study within the broader context of the Indian Knowledge System (IKS). Urdu, as a contact language that evolved on Indian soil, uniquely hosts two major lexical streams: one rooted in the Indo-Aryan tradition through Sanskrit and Prakrit, and another borrowed from Arabic (often via Persian) through centuries of cultural, religious, and political exchange. The paper systematically maps synonym pairs across thematic domains including philosophy, ethics, the human body, nature, and social life. It analyzes the phonological adaptation of Arabic words to Indian pronunciation patterns, examines semantic shifts and register differences between the two lexical sets, and applies frameworks from contact linguistics, diglossia theory, and the Sapir-Whorf hypothesis. The findings demonstrate that far from being linguistically opposed, Arabic-origin and Sanskrit-origin words in Urdu represent complementary registers of a deeply syncretic linguistic tradition. The paper argues that Urdu itself is a living monument to India's history as a civilizational meeting point, and that its dual lexicon offers valuable insights into the co-existence and mutual enrichment of Arabic-Islamic and Sanskrit-Indic knowledge traditions.

Keywords: Arabic loanwords, Urdu, Sanskrit, Prakrit, contact linguistics, diglossia, Indian Knowledge System, Indo-Aryan languages, linguistic borrowing, semantic shift.

1. INTRODUCTION

Language is never merely a system of sounds and symbols; it is a repository of cultural memory, intellectual exchange, and civilizational encounter. Urdu, one of the most poetic and expressive languages of the Indian subcontinent, stands as a singular example of what happens when two great linguistic and intellectual traditions meet, negotiate, and ultimately fuse into something greater than either alone. On one side lies the ancient Indo-Aryan heritage — Sanskrit, the language of the Vedas, the Upanishads, the epics, and the vast corpus of Indian philosophy and science; and its successors Pali, Prakrit, and Apabhramsha, which gave birth to the vernacular tongues of northern India. On the other lies Arabic, the language of the Quran, of Islamic theology, jurisprudence, philosophy, mathematics, and medicine — a language that, through the vehicle of Persian, found its way into the Indian subcontinent during the medieval period and profoundly shaped its literary, religious, and intellectual landscape.

Urdu emerged as a contact language in the garrison towns and courts of medieval India, absorbing Arabic and Persian vocabulary at the formal and literary level while retaining the grammatical skeleton and basic lexicon of Khari Boli, itself derived from Sanskrit through Apabhramsha. The result is a language of extraordinary richness where two parallel lexical streams often run side by side, offering near-synonyms for the same concept — one drawn from Arabic-Persian and the other from Sanskrit-Prakrit. This

phenomenon is not merely a curiosity of etymology; it reflects a deeper historical reality of two knowledge systems coexisting, interacting, and enriching each other on Indian soil.

This paper undertakes a systematic study of Arabic loanwords in Urdu alongside their Sanskrit and Prakrit counterparts. It examines the historical pathways through which Arabic vocabulary entered Urdu, the phonological transformations these words underwent, the semantic shifts and connotative differences between the two lexical sets, and the broader sociolinguistic implications of maintaining two parallel registers within one language. The paper situates this study within the framework of the Indian Knowledge System (IKS), arguing that the dual lexicon of Urdu is itself a form of linguistic evidence for India's ancient tradition of intellectual pluralism and civilizational synthesis.

2. HISTORICAL BACKGROUND

2.1 Arabic and Sanskrit: Two Great Knowledge Traditions

Arabic and Sanskrit are among the most venerable and systematically studied languages in human history. Sanskrit, codified by the grammarian Panini in his *Ashtadhyayi* (c. 4th century BCE), served as the lingua franca of Indian intellectual and religious life for millennia. Its phonological precision, grammatical sophistication, and vast vocabulary made it the medium for the Vedas, Puranas, Buddhist sutras, Jain agamas, mathematical treatises, astronomical texts, medical compendia, and philosophical dialogues. Sanskrit was not merely a language but an epistemological framework — its very structure encoding a particular way of organizing and understanding knowledge.

Arabic, similarly, is not merely a language of everyday communication but a vehicle for a distinctive civilizational heritage. Classical Arabic, stabilized through the Quran and the hadith literature of the 7th century CE, became the medium for Islamic theology, jurisprudence (*fiqh*), philosophy (*falsafa*), mathematics, astronomy, alchemy, and medicine. During the Abbasid Golden Age (8th-13th centuries CE), the translation movement centered at Bayt al-Hikmah (House of Wisdom) in Baghdad saw Arabic scholars translate and synthesize works from Greek, Persian, Sanskrit, and Syriac traditions — crucially including Indian mathematical and astronomical texts such as the *Brahmasphutasiddhanta* and works attributed to Aryabhata. This meant that Arabic and Sanskrit knowledge traditions were already in intellectual dialogue centuries before they met directly on Indian soil.

2.2 Contact and Convergence in India

Arab traders had contact with the western coast of India as early as the pre-Islamic period. The rise of Islam and the expansion of Arab trade networks in the 7th-8th centuries CE brought a new intensity to this contact, as evidenced by the Arab merchant communities of Kerala and the Malabar coast, which gave rise to the Mappila Muslim community — perhaps the earliest indigenous Muslim community in India — whose language, Arabi-Malayalam, itself represents a fascinating early instance of Arabic-Indian linguistic hybridization.

The major pathway for Arabic linguistic influence in India, however, was through Persian. When Muslim political power extended into the subcontinent — first through the Ghaznavid and Ghurid dynasties in the 10th-12th centuries, and then through the Delhi Sultanate (1206-1526) and the Mughal Empire (1526-1857) — the court language was Persian, which had itself absorbed a massive Arabic vocabulary through its conversion to Islam. Persian served as the prestige language of administration, poetry, philosophy, and Sufi mysticism. It was in this context that the mixed vernacular we now call Urdu/Hindi gradually emerged as the spoken language of everyday life in the cosmopolitan cities of northern India, blending the native Khari Boli base with Persian-Arabic vocabulary.

Sufi saints and their khanqahs (hospices) played a particularly significant role in this linguistic exchange. Figures such as Amir Khusrau (1253-1325), who himself wrote in Khari Boli, Persian, and Arabic, consciously celebrated the mixing of languages as a metaphor for spiritual and cultural synthesis. The Sufi tradition of using vernacular Indian languages to express Islamic mystical themes created a space where Arabic-Islamic and Sanskrit-Indic spiritual vocabularies met and were sometimes found to be addressing identical concerns — a point that would later become significant in comparative religious philosophy.

2.3 Emergence of Urdu as a Contact Language

By the 17th and 18th centuries, a stable literary language had emerged that would come to be called Urdu — from the Turkic phrase Zaban-e-Urdu-e-Mualla (the language of the exalted camp) — or, in its nearly identical form when written in Devanagari, Hindi. Urdu used the Nastaliq script inherited from Persian and employed an elevated vocabulary drawn heavily from Arabic and Persian for literary, formal, and religious purposes, while retaining the Sanskrit-derived grammatical structure and basic vocabulary of everyday life. The great poets of the Urdu tradition — Mir Taqi Mir, Mirza Ghalib, Allama Iqbal — deployed the tension between these two lexical registers as a poetic resource, playing on the resonances and dissonances between Arabic-origin and Sanskrit-origin words to create layered meanings.

3. THEORETICAL FRAMEWORK

3.1 Contact Linguistics

The foundational framework for this study is contact linguistics, which examines the linguistic consequences of long-term interaction between speakers of different languages. Thomason and Kaufman's (1988) model distinguishes between borrowing (where a language acquires elements from another under conditions of social prestige or utility) and shift (where a population adopts a new language). In the case of Urdu, what occurred was primarily borrowing: the native Indo-Aryan speech community borrowed Arabic-Persian vocabulary extensively while maintaining the basic grammatical structure of their native tongue. This resulted in a language that is structurally Indo-Aryan but lexically enriched by the Semitic tradition.

3.2 Diglossia

Charles Ferguson's concept of diglossia (1959) — later extended by Joshua Fishman to broader multilingual contexts — describes situations where two varieties of a language (or two related languages) serve different social functions, with one considered higher in prestige and used in formal contexts, and the other considered lower and used in informal or intimate contexts. Urdu presents a fascinating case of lexical diglossia: Arabic-Persian vocabulary functions as the H (high) register, deployed in formal speech, literary writing, religious discourse, and contexts requiring social prestige, while Sanskrit-Prakrit vocabulary functions as the L (low) register, used in everyday conversation, folk literature, and intimate family contexts. This study will demonstrate how this register split maps onto the Arabic vs. Sanskrit loanword division in specific thematic domains.

3.3 The Sapir-Whorf Hypothesis

The weak version of the Sapir-Whorf (or linguistic relativity) hypothesis suggests that the language we use influences our conceptual categories and habitual ways of thinking. Applied to this study, the question becomes: do speakers who choose the Arabic-origin word rather than its Sanskrit counterpart thereby align themselves (even unconsciously) with a particular conceptual framework or cultural orientation? For example, does the choice between *ilm* (Arabic: systematic, empirical knowledge) and *gyan* (Sanskrit: intuitive, spiritual knowledge) reflect not merely different words but different epistemological orientations? This paper will explore this dimension without making strong causal claims, but will examine whether the semantic fields of the two sets of synonyms show systematic differences that might reflect underlying conceptual differences between the two knowledge traditions.

3.4 Indian Knowledge System (IKS) Framework

The Indian Knowledge System (IKS), as conceptualized in contemporary Indian academic discourse and formalized in the National Education Policy (NEP) 2020, refers to the body of knowledge, epistemological frameworks, and intellectual traditions that developed on the Indian subcontinent across millennia, including not only the Sanskrit-Vedic tradition but also Buddhist, Jain, Sikh, and Sufi-Islamic contributions. Situating the present study within an IKS framework means recognizing that the Arabic-origin vocabulary in Urdu is not a foreign intrusion into an otherwise pure Sanskrit linguistic landscape, but rather one strand of India's own multilayered intellectual heritage. The synthesis of Arabic and Sanskrit lexical traditions in Urdu is itself a product of the Indian Knowledge System's historical openness to external intellectual influence and its capacity for integration and transformation.

4. PHONOLOGICAL ADAPTATION OF ARABIC LOANWORDS

When Arabic words passed into Urdu — usually via Persian — they underwent significant phonological transformations. These transformations reflect the differences between the Semitic phonological system of Arabic and the Indo-Aryan phonological system underlying Urdu. Understanding these adaptations is essential for tracing the etymological pathways of loanwords and for recognizing disguised borrowings.

4.1 Consonant Modifications

Arabic has several consonants with no equivalents in Indo-Aryan languages. The pharyngeal consonants 'ayn (a voiced pharyngeal fricative) and ha (a voiceless pharyngeal fricative) are frequently simplified or replaced. The 'ayn, which gives words like 'ilm (knowledge) and 'aql (reason) their characteristic opening sound, is typically either dropped or pronounced as a plain vowel by most South Asian speakers. The uvular consonant qaf, found in words like qalb (heart) and qalam (pen), is often pronounced as a plain velar 'k' in colloquial speech, though educated speakers maintain the distinction. The interdental fricatives th (as in mathala, example) and dh are typically replaced by 't' and 'd' respectively. Conversely, Indo-Aryan retroflex consonants (sounds made with the tongue curled back) have sometimes been introduced into Arabic loanwords when they fit patterns familiar to Indian speakers.

4.2 Vowel Shifts

Arabic is a language rich in long vowels (alif, waw, ya) and short vowels (fatha, damma, kasra). These were frequently modified in their transition through Persian into Urdu. Arabic's short 'u' vowel (damma) often became 'o' in Urdu — thus Arabic kitab (book, with short 'i') became the same in Urdu but its plural kutub became kutub. The Arabic definite article al- was often dropped or assimilated, and Arabic broken plurals (an irregular plural system using vowel patterns rather than suffixes) were sometimes preserved but sometimes replaced with Persian or Indo-Aryan plural formations.

4.3 Preservation of the Trilateral Root

One of the most distinctive features of Arabic morphology is the trilateral (three-consonant) root system, in which most words are derived from a root of three consonants by inserting different vowels and adding prefixes and suffixes. For example, the root K-T-B relates to writing: kataba (he wrote), kitab (book), katib (writer), maktaba (library), maktub (letter/written). In Urdu, these related words often entered as individual borrowings, with their etymological relationship partially obscured. However, educated speakers and scholars recognize these root connections, and the trilateral root system provides a powerful tool for etymological analysis. This is analogous to the Sanskrit dhatu (verbal root) system in the Indo-Aryan tradition, where most words can be traced back to a monosyllabic verbal root. The parallel between Arabic trilateral roots and Sanskrit dhatupatha (root lists) is itself a fascinating structural similarity between the two knowledge traditions.

5. DOMAIN-WISE LEXICAL ANALYSIS AND WORD PAIRS

The following section presents a systematic mapping of Arabic loanwords in Urdu alongside their Sanskrit and Prakrit counterparts across five major thematic domains. For each pair, we note the Arabic origin (with triliteral root where applicable), the Sanskrit/Prakrit origin (with dhatu where applicable), and brief notes on semantic and connotative differences.

5.1 Philosophy and Knowledge

The domain of philosophy and epistemology reveals perhaps the most interesting synonym pairs, as both the Arabic-Islamic and Sanskrit-Indic traditions developed rich vocabularies for discussing the nature of knowledge, reality, consciousness, and time.

Concept	Arabic-Origin (Urdu)	Root	Sanskrit-Origin	Root
Knowledge	Ilm (علم)	'-L-M	Gyan (ज्ञान)	jna
Intellect/Reason	Aql (عقل)	'-Q-L	Buddhi (बुद्धि)	budh
Truth/Reality	Haq (حق)	H-Q-Q	Satya (सत्य)	sat
Soul/Spirit	Rooh (روح)	R-W-H	Atma (आत्मा)	an (breath)
Time	Waqt (وقت)	W-Q-T	Kaal (काल)	kal
World/Universe	Duniya (دنیا)	D-N-W	Sansar (संसार)	sar (flow)
Creation	Khalq (خلق)	KH-L-Q	Srishti (सृष्टि)	srj
Wisdom	Hikmat (حکمت)	H-K-M	Prajna (प्रज्ञा)	jna
Thought/Reflection	Fikr (فکر)	F-K-R	Chintan (चिंतन)	cit
Consciousness	Shaoor (شعور)	SH-'-R	Chetan (चेतन)	cit

The pair ilm / gyan illustrates a particularly rich semantic divergence. Arabic ilm, from the root '-L-M (to know, to mark), denotes systematic, verifiable, empirically grounded knowledge — the kind of knowledge that can be transmitted, codified, and tested. It is the same root that gives Arabic the word 'alam (world, i.e., that which is known) and 'alama (sign, mark). In Islamic intellectual tradition, ilm encompasses both religious sciences (theology, jurisprudence, hadith) and rational sciences (philosophy, mathematics, medicine). Sanskrit gyan or jnana, from the root jna (to know), also means knowledge, but in the Hindu philosophical tradition — especially in Advaita Vedanta — it carries a specifically spiritual and intuitive dimension: gyan-yoga is the path to liberation through direct, non-mediated knowledge of the Self (Atman) as identical with ultimate reality (Brahman). The Gita distinguishes between different types of knowledge, with para-vidya (higher knowledge, i.e., knowledge of Brahman) ranked above apara-vidya (lower knowledge, i.e., empirical knowledge of the world). Thus while both ilm and gyan translate as 'knowledge,' they carry subtly different emphases: ilm points toward systematized, transmissible learning; gyan points toward intuitive, liberating insight.

5.2 Ethics and Morality

The domain of moral vocabulary shows how two distinct ethical frameworks — one rooted in Islamic jurisprudence and Sufi spirituality, the other in Dharmic ethics — contributed their terminologies to the Urdu moral lexicon.

Concept	Arabic-Origin (Urdu)	Root	Sanskrit-Origin	Root
Justice	Adl / Insaaf (انصاف)	'-D-L / N-S-F	Nyaya (न्याय)	ni + ya
Patience/Endurance	Sabr (صبر)	S-B-R	Dhairya (धैर्य)	dhr
Compassion	Rahm (رحم)	R-H-M	Karuna (करुणा)	kr
Forgiveness	Maafi (معافی)	'-F-W	Kshama (क्षमा)	ksam
Gratitude	Shukar (شکر)	SH-K-R	Kritagyata (कृतज्ञता)	krt + jna
Honesty/Integrity	Imaandaari (ایمانداری)	'-M-N	Satyanishtha (सत्यनिष्ठा)	sat
Virtue/Goodness	Nek (نیک - Persian)	—	Satgun (सद्गुण)	sat + guna
Sin/Transgression	Khataa / Gunah (خطا)	KH-T-'	Paap (पाप)	pa
Duty/Obligation	Farz (فرض)	F-R-DH	Dharma (धर्म)	dhr
Mercy	Rehmat (رحمت)	R-H-M	Daya (दया)	day

The pair farz / dharma offers a particularly profound point of comparison. Arabic farz (from the root F-R-DH, to cut, to impose) denotes an obligation or duty imposed by divine command in Islamic jurisprudence — one of the five categories of legal acts in fiqh. Sanskrit dharma (from the root dhr, to hold, to support) denotes a cosmic principle of righteous order that pervades the universe and prescribes appropriate conduct for each being according to its nature and position. Both words translate roughly as 'duty' or 'obligation,' but their metaphysical grounding is entirely different: farz is a duty because God has commanded it; dharma is a duty because it sustains the cosmic order (rta). Remarkably, both concepts converge on the idea that individual ethical conduct is not arbitrary but has a transpersonal grounding — whether in divine command or in cosmic law. This convergence, even across very different metaphysical frameworks, may help explain why the two vocabulary streams were able to coexist so comfortably in Urdu without one displacing the other.

5.3 Nature and Cosmos

Natural and cosmological vocabulary in Urdu shows an interesting pattern: Arabic-Persian terms tend to dominate in literary and poetic registers, while Sanskrit-derived terms persist more strongly in folk and rural usage.

Concept	Arabic-Persian (Urdu)	Sanskrit-Origin	Notes
Sky/Heaven	Aasman (آسمان) / Falak (فلک)	Aakash (आकाश)	Falak also means fate in Urdu poetry
Sun	Aftaab (آفتاب) / Shams (شمس)	Surya (सूर्य) / Ravi	Shams is Arabic; Aftaab is Persian
Moon	Mah (ماه) / Qamar (قمر)	Chandra (चंद्र)	Qamar is Arabic; used in names
Water	Aab (آب) / Maa (ماء)	Jal (जल) / Paani	Aab is Persian; used in place names
Earth/Soil	Zameen (زمین) / Arz (ارض)	Dharti / Prithvi (पृथ्वी)	Arz is Quranic Arabic
Wind/Air	Hawa (ہوا) / Bad (باد)	Vayu (वायु) / Pawan	Hawa is Arabic; widely used in Urdu
Night	Raat / Shab (شب)	Ratri (रात्रि)	Raat from Skt; Shab from Persian
Fire	Aag / Aatish (آتش)	Agni (अग्नि)	Aatish is Persian; Agni is Vedic
Sea/Ocean	Darya (دریا) / Bahr (بحر)	Samudra (समुद्र) / Sagar	Bahr is Quranic Arabic
Star	Sitara (ستاره) / Najm (نجم)	Nakshatra (नक्षत्र) / Tara	Tara is common in both traditions

5.4 Body and Medicine

The domain of body and medicine is especially significant for the study of the Indian Knowledge System, as both the Arabic-Islamic Unani medical tradition and the Indian Ayurvedic tradition developed comprehensive vocabularies for the human body and its diseases. Both traditions coexisted and influenced each other in Mughal India, producing a rich hybrid medical lexicon.

Concept	Arabic-Origin (Urdu)	Sanskrit-Origin	Notes
Body	Jism (جسم)	Sharir (शरीर)	Jism from J-S-M; Sharir from sru (to flow)
Heart	Qalb (قلب) / Dil (دل)	Hridaya (हृदय)	Qalb is Arabic; Dil is Persian
Blood	Dam (دم) / Khoon (خون)	Rakta (रक्त)	Dam relates to breath; Khoon is Persian
Breath/Life Force	Saans (سانس)	Prana (प्राण)	Saans is common Urdu/Hindi word

Head	Sar (سر) / Dimagh (دماغ)	Mastishk / Sir (सिर)	Dimagh from Greek via Arabic
Eye	Aankh / Chashm (چشم)	Akshi / Netra (नेत्र)	Chashm is Persian; Aankh is Sanskrit origin
Disease	Marz (مرض) / Bimari	Rog (रोग) / Vyadhi	Marz from M-R-DH; Rog from Sanskrit
Medicine/Cure	Dawa (دوا) / Ilaj (علاج)	Aushadh (औषध) / Chikitsa	Dawa from D-W-Y; both systems in India
Mind/Brain	Zehn (ذهن) / Dimagh (دماغ)	Manas (मन) / Chitta	Zehn from Aramaic/Arabic
Life	Zindagi (زندگی) / Hayat (حيات)	Jeevan (जीवन) / Prana	Hayat from H-Y-Y (to live)

5.5 Social and Familial Life

Social vocabulary in Urdu shows a characteristic pattern where intimate family terms are predominantly Sanskrit-derived, while terms for social institutions, law, and formal relationships more often draw from Arabic-Persian.

Concept	Arabic-Persian (Urdu)	Sanskrit-Origin	Notes
Mother	Walida (والده) / Ammi	Mata (माता) / Maa	Ammi is colloquial Urdu; Mata is formal
Father	Walid (والد) / Abba	Pita (पिता) / Baap	Walid is formal/documentary Arabic
Marriage	Nikah (نكاح) / Shadi	Vivah (विवाह) / Byah	Nikah is Islamic contract term
Family	Khandaan (خاندان) / Aila	Parivar (परिवार) / Kutumb	Khandaan from Persian; Aila is Arabic
Friend	Dost (دوست) / Rafeeq	Mitra (मित्र) / Sakha	Rafeeq from R-F-Q (to accompany)
Teacher	Ustad (استاد) / Muallim	Guru (गुरु) / Acharya	Ustad from Persian; Muallim from '-L-M
Student	Shagird (شاگرد) / Talib	Shishya (शिष्य) / Vidyarthi	Talib from T-L-B (to seek)
King/Ruler	Sultan / Ameer (امير)	Raja (राजा) / Narendra	Sultan is Arabic; Ameer from '-M-R
Prayer/Worship	Namaz / Dua (دعا)	Pooja (पूजा) / Prarthana	Namaz from Persian; Dua is Arabic

Love	Mohabbat (محبت) / Ishq	Prem (प्रेम) / Sneh	Ishq from '-SH-Q (yearning love)
------	---------------------------	---------------------	-------------------------------------

6. SEMANTIC SHIFT AND CONNOTATIVE DIFFERENCES

When words are borrowed across languages, they rarely retain exactly the same meaning. Semantic change occurs along several dimensions: narrowing (the word takes on a more restricted meaning), widening (the word takes on a broader meaning), amelioration (the word acquires more positive connotations), pejoration (the word acquires more negative connotations), or specialization (the word becomes associated with a particular domain or register). In the case of Arabic loanwords in Urdu, several patterns of semantic shift are observable.

6.1 Narrowing and Specialization

Several Arabic words that had broad meanings in their source language acquired more specialized meanings in Urdu. The Arabic word *ilm*, as discussed above, means any kind of knowledge or learning; in Urdu, it often carries a narrower connotation of formal, religious, or scholarly learning, as opposed to practical skill (*hunar*) or everyday knowing. Similarly, the Arabic word *fiqh* (Islamic jurisprudence) entered Urdu primarily in religious contexts, while its Sanskrit near-equivalent *dharmashastra* remained in specifically Hindu legal discourse.

6.2 Register and Formality

Perhaps the most consistent pattern of semantic differentiation between Arabic-origin and Sanskrit-origin synonyms in Urdu is one of register — that is, the level of formality, prestige, or social context associated with each word. Arabic-Persian vocabulary consistently functions as the H (high) register, appropriate to formal writing, official communication, religious discourse, and literary poetry. Sanskrit-Prakrit vocabulary functions as the L (low) register, appropriate to informal conversation, folk literature, and intimate communication. This is illustrated clearly in the domain of address terms for family members: a speaker might refer to their mother as *ammi* (colloquial, intimate) in conversation but as *walida* (formal, Arabic) in documentary or official contexts, while the Sanskrit *mata* (mother) might appear in formal Hindi contexts but sounds stilted in Urdu.

6.3 Religious and Cultural Indexicality

Beyond register, many synonym pairs in Urdu have acquired what linguists call indexical meaning — they point to or index a particular cultural, religious, or ethnic identity. The use of Arabic-origin vocabulary may signal Islamic identity or education, while the use of Sanskrit-origin vocabulary may signal Hindu identity or a more broadly secular, pre-Partition Indian cultural affiliation. This indexicality is most pronounced in specifically religious vocabulary (*nikah* vs. *vivah* for marriage; *namaz* vs. *pooja* for worship) but extends into domains like ethics and philosophy. This does not mean that speakers mechanically choose one register or the other based on religious identity — educated Urdu speakers typically draw freely from both registers — but it does mean that the choice between synonyms can carry subtle social meaning.

6.4 Spiritual Connotations in Sufi Usage

Sufi poetry provides a unique context where Arabic-origin and Sanskrit-origin vocabularies blend most creatively. Poets like Bulleh Shah, Kabir, and Amir Khusrau freely mixed Islamic and Hindu spiritual vocabularies, using the tension between the two lexical registers to create poetic effects of universal spiritual aspiration that transcended religious boundaries. The word *ishq* (Arabic: passionate, yearning love) was used by Sufi poets as a synonym for and complement to the Sanskrit *prem* (love, especially devotional love) — both pointing toward the mystical love of the Divine that transcends doctrinal

boundaries. In this context, the two lexical traditions are not competing alternatives but complementary expressions of a shared spiritual experience.

7. DIGLOSSIA AND SOCIOLINGUISTIC REGISTER IN URDU

The concept of diglossia, originally applied by Charles Ferguson (1959) to situations where two varieties of the same language serve different social functions, has been extended by Joshua Fishman (1967) to encompass broader situations of multilingualism and register variation. Urdu presents what might be called a case of lexical diglossia: within a single language, two parallel vocabulary sets — one Arabic-Persian, one Sanskrit-Prakrit — serve systematically different social functions.

In classical diglossia, the H variety is used in formal, literary, and official contexts, while the L variety is used in informal, domestic, and conversational contexts. This pattern maps closely onto Urdu's dual lexicon. Arabic-Persian vocabulary (H register) appears in formal writing, literary poetry, religious texts, administrative documents, and educated formal speech. Sanskrit-Prakrit vocabulary (L register) appears in everyday conversation, folk songs, bazaar speech, and informal interaction. However, unlike classical diglossia, both registers are available to educated speakers, and the choice between them is not simply determined by context but can also express identity, emphasis, mood, and aesthetic preference.

The Hindi-Urdu continuum offers a particularly instructive perspective. What is called Urdu at the H-register extreme is largely Arabic-Persian vocabulary written in Nastaliq script; what is called Hindi at the H-register extreme is largely Sanskrit-tatsama (directly borrowed from Sanskrit) vocabulary written in Devanagari script. But at the L-register level — in everyday conversation — Urdu and Hindi are mutually intelligible and share the same basic Khari Boli vocabulary. The famous controversy over Hindustani — the name sometimes given to this common vernacular — was precisely a controversy over whether the H register should privilege the Arabic-Persian or the Sanskrit stream. This controversy itself illustrates how deeply the diglossia between Arabic-origin and Sanskrit-origin vocabulary is embedded in the politics of language and identity in South Asia.

8. URDU AS A BRIDGE BETWEEN ARABIC AND SANSKRIT KNOWLEDGE TRADITIONS

The foregoing analysis suggests that Urdu functions not merely as a language of communication but as a bridge between two major knowledge traditions: the Arabic-Islamic and the Sanskrit-Indic. This bridging function operates at multiple levels.

8.1 Conceptual Translation

At the most basic level, the existence of near-synonym pairs in Urdu means that concepts developed in one tradition can be expressed using vocabulary from the other. A scholar trained in Arabic-Islamic philosophy can discuss the concept of *ilm* using the Arabic term, but can also gesture toward the concept of *gyan* using the Sanskrit term, and in doing so potentially perceive a connection between the two traditions that might be invisible to a monolingual scholar. Urdu speakers navigating between these two lexical registers are effectively navigating between two conceptual worlds, and the near-synonymy of so many pairs suggests that these two worlds, for all their differences, addressed many of the same fundamental questions about knowledge, ethics, consciousness, and the good life.

8.2 Sufi-Bhakti Synthesis

The Sufi-Bhakti synthesis of medieval India — the meeting of Islamic mysticism and Hindu devotional religion that produced some of the most beautiful poetry in Indian literary history — was made possible in part by the linguistic bridge of Urdu. Sufi poets using Urdu could draw simultaneously on Arabic-Islamic mystical vocabulary (*ishq*, *fana*, *wajd* — love, annihilation, ecstasy) and Sanskrit-Indic devotional vocabulary (*prem*, *bhakti*, *samadhi* — love, devotion, absorption) to articulate a spiritual experience that

seemed to transcend the boundaries of any single religious tradition. The fact that both vocabularies were available within a single linguistic space allowed for a kind of conceptual comparison and synthesis that would have been more difficult in a purely Arabic or purely Sanskrit context.

8.3 Scientific Translation

The tradition of Arabic-Sanskrit scientific exchange that had begun in the Abbasid period — with Arabic translations of Sanskrit mathematical and astronomical texts — continued in a different form in Mughal India. The Unani medical tradition, based on Arabic-Islamic medicine (itself derived partly from Greek and partly from Indian Ayurvedic sources), coexisted with and influenced Ayurveda. Medical practitioners in Mughal India sometimes trained in both traditions, and the vocabulary of Urdu medical texts reflects this synthesis, with Arabic-origin terms for many body parts and diseases coexisting with Sanskrit-origin equivalents.

9. URDU AND THE INDIAN KNOWLEDGE SYSTEM (IKS)

The National Education Policy of India (2020) has placed renewed emphasis on the Indian Knowledge System (IKS), calling for the integration of India's traditional intellectual heritage into contemporary education. This has sometimes been interpreted narrowly as focusing exclusively on the Sanskrit-Vedic tradition. The present study argues for a broader and more historically accurate understanding of IKS that includes the Arabic-Islamic contribution to India's intellectual heritage.

The evidence of Arabic loanwords in Urdu — their deep integration into the language, their adaptation to Indian phonological and semantic patterns, their synthesis with Sanskrit vocabulary in the crucible of Sufi poetry and Mughal court culture — demonstrates that the Arabic-Islamic tradition is not an external intrusion into Indian intellectual life but a constitutive part of it. The Indian Muslims who developed Urdu literature, who practiced Unani medicine, who translated and commented on philosophical texts, who built the Mughal administrative and legal system, were not foreigners importing a foreign tradition; they were Indians contributing to the ongoing development of a distinctively Indian intellectual synthesis.

Recognizing the Arabic-Islamic tradition as part of IKS has practical implications for linguistics education. The study of Arabic in Indian universities should be framed not merely as a study of a foreign language but as a study of one strand of India's own multilayered intellectual heritage. Conversely, students of Arabic should be introduced to the Sanskrit-Indic tradition not as something alien to Arabic linguistic culture but as a tradition with which Arabic has been in productive dialogue for over a millennium. The lexical pairs analyzed in this paper provide concrete material for such cross-traditional teaching.

10. CONCLUSION

This paper has examined Arabic loanwords in Urdu alongside their Sanskrit and Prakrit counterparts, mapping synonym pairs across five thematic domains and analyzing the phonological, semantic, and sociolinguistic dimensions of this dual lexicon. Several major findings emerge from this analysis.

First, the Arabic-Urdu loanword phenomenon is not a simple case of one language borrowing words from another, but a complex process of phonological adaptation, semantic shift, and register differentiation shaped by centuries of cultural, religious, and political interaction between Arabic-Islamic and Sanskrit-Indic civilizations on Indian soil.

Second, the near-synonym pairs that result from this process are not merely stylistic variants but carriers of subtly different conceptual emphases, reflecting the different epistemological frameworks, ethical systems, and metaphysical worldviews of their source traditions. The difference between *ilm* and *gyan*,

between farz and dharma, between ishq and prem, is not merely lexical but encodes something of the intellectual heritage of the traditions that generated these concepts.

Third, the co-existence of Arabic-origin and Sanskrit-origin vocabulary in Urdu is a manifestation of the broader principle of diglossia, with Arabic-Persian vocabulary serving the H (formal, literary, religious) register and Sanskrit-Prakrit vocabulary serving the L (informal, conversational, folk) register. This diglossia reflects the social history of Urdu's formation in the cosmopolitan, multilingual environment of medieval and early modern India.

Fourth, and most significantly for the Indian Knowledge System framework, Urdu's dual lexicon demonstrates that the Arabic-Islamic and Sanskrit-Indic traditions are not opposed civilizational alternatives but complementary strands of a shared Indian intellectual heritage. The greatest achievements of Urdu literature — the poetry of Ghalib, the prose of Premchand, the philosophy of Iqbal — draw freely from both traditions, weaving them into a synthesis that is neither purely Arabic nor purely Sanskrit but distinctively and irreducibly Indian.

Future research directions suggested by this study include: a corpus-based computational analysis of register distribution of Arabic-origin vs. Sanskrit-origin vocabulary in contemporary Urdu media; a sociolinguistic survey of register preferences among different generational and educational cohorts of Urdu speakers; a comparative study of the Hindi-Urdu lexical split as a case study in language planning and identity politics; and a philosophical analysis of conceptual convergences and divergences between Arabic-Islamic and Sanskrit-Indic epistemological categories as illuminated by their Urdu lexical realizations.

Urdu stands as living proof that languages — and the knowledge traditions they carry — do not simply compete for dominance but can, under the right historical conditions, enter into creative synthesis. In a world where civilizational dialogue is more urgently needed than ever, Urdu's example of linguistic pluralism is both an intellectual resource and an inspiration.

REFERENCES:

1. Platts, J. T. (1884). A dictionary of Urdu, classical Hindi, and English. W. H. Allen & Co. (Reprinted Oxford University Press, 1974).
2. Lane, E. W. (1863-1893). An Arabic-English lexicon (8 vols.). Williams and Norgate.
3. Monier-Williams, M. (1899). A Sanskrit-English dictionary. Clarendon Press. (Reprinted Motilal Banarsidass, 1999).
4. Farhang-e-Aasafiya (1898). (فرہنگ آصفیہ). Syed Ahmad Dehlawi (Ed.). Matba Mujtabai, Delhi.
5. Apte, V. S. (1890). The practical Sanskrit-English dictionary. Prasad Prakashan. (Revised edition, 1965).
6. Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325-340.
7. Fishman, J. A. (1967). Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues*, 23(2), 29-38.
8. Thomason, S. G., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. University of California Press.
9. Myers-Scotton, C. (1993). *Social motivations for codeswitching: Evidence from Africa*. Clarendon Press.
10. Cardona, G., & Jain, D. (Eds.). (2003). *The Indo-Aryan languages*. Routledge.
11. Hock, H. H. (1991). *Principles of historical linguistics* (2nd ed.). Mouton de Gruyter.
12. Masica, C. P. (1991). *The Indo-Aryan languages*. Cambridge University Press.

13. Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. MIT Press.
14. Rahman, T. (2011). *From Hindi to Urdu: A social and political history*. Oxford University Press.
15. Faruqi, S. R. (2001). *Early Urdu literary culture and history*. Oxford University Press.
16. Jain, G. C. (1975). *Urdu ki lughawi tashkeel (The lexical formation of Urdu)*. Taraqqi-e-Urdu Bureau.
17. Rai, A. (1984). *A house divided: The origin and development of Hindi/Hindavi*. Oxford University Press.
18. King, C. R. (1994). *One language, two scripts: The Hindi movement in nineteenth century North India*. Oxford University Press.
19. Naim, C. M. (1999). *Urdu texts and contexts: The selected essays of C. M. Naim*. Permanent Black.
20. Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Georgetown University Press.
21. Wright, W. (1896). *A grammar of the Arabic language (3rd ed., 2 vols.)*. Cambridge University Press.
22. Versteegh, K. (1997). *The Arabic language*. Edinburgh University Press.
23. Al-Biruni, A. R. M. (1030/1964). *Alberuni's India (E. C. Sachau, Trans., 2 vols.)*. S. Chand & Co.
24. Dasgupta, S. (1922-1955). *A history of Indian philosophy (5 vols.)*. Cambridge University Press.
25. Matilal, B. K. (1986). *Perception: An essay on classical Indian theories of knowledge*. Clarendon Press.
26. Nasr, S. H. (1981). *Islamic life and thought*. State University of New York Press.
27. Ernst, C. W. (1997). *The Shambhala guide to Sufism*. Shambhala Publications.
28. Rizvi, S. A. A. (1978). *A history of Sufism in India (2 vols.)*. Munshiram Manoharlal.
29. Ministry of Education, Government of India. (2020). *National Education Policy 2020*. Department of School Education and Literacy.
30. Chatterji, S. K. (1926). *The origin and development of the Bengali language*. Calcutta University Press.
31. Grierson, G. A. (1903-1928). *Linguistic survey of India (11 vols.)*. Superintendent of Government Printing, Calcutta.
32. Saliba, G. (2007). *Islamic science and the making of the European Renaissance*. MIT Press.
33. Pingree, D. (1978). *History of mathematical astronomy in India*. *Dictionary of Scientific Biography*, 15, 533-633.
34. Wujastyk, D. (2003). *The roots of Ayurveda: Selections from Sanskrit medical writings*. Penguin Classics.
35. Metcalf, B. D. (1982). *Islamic revival in British India: Deoband, 1860-1900*. Princeton University Press.