OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue - iv, Published on October issue 2026, Page No. 747 - 752
Website: https://tirj.org.in/tirj, Mail ID: editor@tirj.org.in
(SJIF) Impact Factor 8.111, e ISSN : 2583 - 0848

Between Global Vitality & Local Vulnerability : A Multidimensional Reassessment of Language Endangerment in Bengali

Dr. Md Siddique Hossain
Assistant Professor, Department of Bengali
Bangabasi Morning College, 19, Rajkumar Chakraborty Sarani, Kolkata – 700009
Email ID: mdsh803@gmail.com

Received Date *30. 09. 2026*
Selection Date *15. 10. 2026*

Keyword
Bengali language, language endangerment, language vitality, domain loss, low-resource languages, dialect continuum, diaspora heritage language, UNESCO Atlas of Languages in Danger, reversing language shift.

Abstract
Bengali is, by every conventional demo-linguistic measure, one of the world's major languages, ranked among the top ten to fifteen most-spoken languages globally with an estimated 230–300 million speakers across Bangladesh, India, and a substantial diaspora. On this basis, standard sociolinguistic typologies classify it as 'safe.' Yet a purely numerical reading of vitality obscures a more complex reality: several closely related varieties within the greater Bengali-Assamese dialect continuum are formally classified as vulnerable or endangered by UNESCO; standard Bengali itself is experiencing measurable domain loss in higher education, judiciary, science, and digital infrastructure; heritage transmission is visibly weakening across second and third-generation diaspora communities; and the language remains classified as 'low-resource' in natural language processing despite its speaker base. This paper argues for a domain-specific, multidimensional model of endangerment—drawing on Fishman's Graded Intergenerational Disruption Scale and UNESCO's nine vitality factors—that can hold Bengali's demographic robustness and its structural vulnerabilities in the same analytical frame. Using the 1952 Bengali Language Movement as a historical counter-case of successful reversal of language shift, the paper concludes with policy recommendations for language planning in education, digital infrastructure, and diaspora heritage programs.

Discussion

1. Introduction : Discussions of language endangerment have historically centred on languages with small, dwindling speaker populations - languages with a few hundred or a few thousand remaining speakers, often undocumented, and facing imminent extinction within a generation (Krauss, 1992; Moseley, 2010). Bengali sits uneasily, and productively, outside this paradigm. With an estimated 230 to 300 million native and second-language speakers, official status in Bangladesh and the Indian states of West Bengal, Tripura, and Assam, and a rich, continuous literary tradition stretching back over a millennium, Bengali would appear to be the antithesis of an endangered language.

OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

This paper argues that the binary of 'safe' versus 'endangered', while analytically convenient, does not capture the layered reality of a language as large and internally diverse as Bengali. Endangerment, properly understood through frameworks such as UNESCO's nine factors of language vitality (UNESCO Ad Hoc Expert Group on Endangered Languages, 2003) or Fishman's Graded Intergenerational Disruption Scale (Fishman, 1991), is not a single property of a language as a whole but a set of conditions that can apply unevenly across a language's internal varieties, domains of use, and speaker communities. A language can be numerically 'safe' at the level of the standard, codified variety while simultaneously losing ground in specific domains, among specific populations, or in specific communicative functions. This paper develops that argument for Bengali across four empirical dimensions - (i) the endangerment status of related and subordinate varieties within the greater Bengali-Assamese continuum; (ii) domain loss in education, administration, and science; (iii) the erosion of intergenerational transmission in diaspora settings; and (iv) Bengali's persistent classification as a 'low-resource' language in digital and computational infrastructure. It closes by revisiting the 1952 Bengali Language Movement as a historically documented case of successful language-status reversal, and by drawing out its implications for contemporary language planning.

2. Theoretical Framework: Defining Endangerment Beyond Speaker Counts—

2.1. Fishman's Graded Intergenerational Disruption Scale : Joshua Fishman's (1991) Graded Intergenerational Disruption Scale (GIDS) reframed language endangerment as fundamentally a question of intergenerational transmission rather than aggregate speaker numbers. Fishman's eight-stage scale distinguishes languages that are transmitted naturally within the home and community (the most robust stages) from languages that survive only in institutional, ceremonial, or archival contexts (the most endangered stages), irrespective of how many total speakers a language may claim across all age groups. A language can therefore have millions of adult speakers and still be in incipient shift if children in a given community are no longer acquiring it as a first language - a condition directly relevant to Bengali diaspora communities discussed in Section 6.

2.2 UNESCO's Nine Factors of Language Vitality : The UNESCO Ad Hoc Expert Group on Endangered Languages (2003) proposed nine factors for assessing language vitality, including intergenerational transmission, absolute number of speakers, proportion of speakers within the total population, trends in existing language domains, response to new domains and media, availability of materials for language education and literacy, governmental and institutional language attitudes and policies, community members' own attitudes toward their language, and the type and quality of documentation. Crucially, this framework treats 'domains of use' and 'response to new domains and media' as independent axes of vitality — meaning a language can retain strong transmission and identity value in the home while simultaneously ceding ground in law, higher education, science, and digital technology. This factor-based, rather than binary, approach is the analytical lens adopted throughout this paper.

3. Bengali : A Demo-linguistic Profile— Bengali (endonym Bangla) is an Indo-Aryan language of the Bengali-Assamese branch, native to the historical region of Bengal, now divided between the sovereign state of Bangladesh and the Indian state of West Bengal. Estimates of its global speaker population vary by source and methodology - reflecting well-documented difficulties in defining language versus dialect boundaries and reconciling census, native-speaker, and total-speaker counts - but converge in the range of roughly 230 to 300 million speakers, placing Bengali among the world's five to fifteen most-spoken languages depending on the ranking criterion used (Eberhard, Simons, & Fennig, 2024). It is the sole official language of Bangladesh, one of the official languages of India, and a recognized state language in West Bengal, Tripura, and Assam. Bengali also carries UNESCO-designated global symbolic significance: 21 February, the anniversary of the killing of student protesters demanding recognition of Bengali as a state language of then-East Pakistan in 1952, was declared International Mother Language Day by UNESCO in 1999 and is now observed worldwide (Ghosh & Chatterjee, 2024). By any conventional demo-linguistic or institutional measure, therefore, Bengali is unambiguously a vital, 'safe' language at the level of the standard variety.

OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

4. The Paradox of a 'Safe' Macro-Language : Endangerment Within the Bengali-Speaking Continuum— The classification of Bengali as safe becomes considerably less straightforward once attention shifts from the standardised, literary variety to the wider dialect continuum from which it draws its speaker base. Several varieties that are popularly treated as 'dialects of Bengali' are classified by linguists - and in several cases by the UNESCO Atlas of the World's Languages in Danger itself - as structurally distinct languages with their own, more precarious, vitality profiles. Sylheti, spoken by an estimated ten million people primarily in the Sylhet region of Bangladesh and in the Barak Valley of Assam, is mutually unintelligible with standard Bengali by most linguistic assessments and is formally classified as Vulnerable by the UNESCO Atlas (Moseley, 2010). Chittagonian, spoken by an estimated thirteen million people in south-eastern Bangladesh, is similarly non-mutually-intelligible with standard Bengali despite frequently being treated administratively as a nonstandard dialect. Rajbangsi, Kharia Thar, and Mal Paharia, though closely related to western Bengali dialects, are typically classified as separate languages in their own right (New World Encyclopedia, n.d.), and Chakma, spoken in the Chittagong Hill Tracts and written in its own script, likewise retains contrastive tonal features distinguishing it from standard Bengali phonology.

This creates what might be termed a nested-endangerment structure: the "macro" language, standard Bengali, is demographically and institutionally robust, but several of the vernacular varieties that feed into the wider Bengali-speaking population are independently vulnerable, and their speakers face social and educational pressure to shift toward the standardised variety for purposes of literacy, employment, and administrative recognition — the very dynamic that, at smaller scale, has historically driven language shift and loss elsewhere in South Asia (List of Endangered Languages in Bangladesh, n.d.). Treating 'Bengali' as a single undifferentiated entity in vitality assessments therefore risks masking active shift and loss processes occurring within its own linguistic family.

5. Domain Loss in Standard Bengali—

5.1. Higher Education, Administration, and the Judiciary : Even where Bengali retains strong home and community transmission, UNESCO's vitality framework identifies domain loss - the retreat of a language from formal, prestige, or specialized functions - as an independent risk factor. In West Bengal, Bengali competes with English and, increasingly, Hindi for domains such as higher education, the higher judiciary, corporate employment, and civil-service recruitment, a pattern that recurs across many Indian regional languages under the pressure of a hierarchically organized, English-and-Hindi-dominant national linguistic economy. In Bangladesh, where Bengali enjoys unambiguous constitutional and administrative primacy following the 1952 Language Movement, English nonetheless dominates higher technical and scientific education, international business, and much of the upper judiciary, producing a diglossic pattern in which Bengali retains near-universal vernacular and literary vitality while ceding specific high-prestige technical and administrative domains.

5.2 The Digital Domain : Bengali as a 'Low-Resource' Language : Perhaps the most counter-intuitive dimension of Bengali's vulnerability lies in digital and computational infrastructure. Despite its enormous speaker base, Bengali is consistently classified in computational linguistics as a 'low-resource' language - a term denoting the scarcity of large-scale, high-quality, machine-readable text and annotated data available for training natural language processing (NLP) systems, translation engines, and speech technologies (Ghosh & Caliskan, 2023). This classification is not a function of speaker numbers but of colonial and postcolonial digital infrastructure history: the earliest large multilingual corpora were assembled from European-language sources such as the Bible and the European Parliament proceedings, and internet penetration, digitisation investment, and corpus-building efforts have historically privileged languages of the Global North (Ghosh & Chatterjee, 2024). The practical consequence is that machine translation and generative-AI systems handling Bengali exhibit measurably higher error rates than they do for high-resource languages, including systematic gender-assignment errors in English–Bengali translation documented in empirical evaluations of large language models (Ghosh & Caliskan, 2023). Domain retreat in this sense is prospective as much as it is historical: as science, commerce, and administration increasingly mediate themselves through AI-driven digital

OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

tools, a language under-represented in the training data underlying those tools risks a renewed and accelerating form of domain loss, even while its number of human speakers remains large and stable.

6. Diasporic Bengali : Intergenerational Transmission Under Threat— Fishman's (1991) intergenerational-transmission criterion is most acutely tested not in Bangladesh or West Bengal but among Bengali-speaking diaspora communities in the United Kingdom, North America, and the Gulf states. Across immigrant heritage-language contexts generally, the well-documented three-generation shift pattern - in which the immigrant generation retains the heritage language, the second generation becomes bilingual with growing dominance in the host-country language, and the third generation frequently loses productive fluency altogether - places diaspora Bengali at meaningfully greater risk of attrition than Bengali in its home territories, notwithstanding the aggregate demographic robustness of the language worldwide. Community-run supplementary schools, heritage-language weekend classes, and cultural organization built around festivals such as Language Movement Day and Durga Puja function, in Fishman's terms, as compensatory institutional transmission mechanisms substituting for weakening home transmission - a pattern typical of GIDS stages 6 through 4, where a language survives through deliberate institutional effort rather than unreflective intergenerational continuity. This diasporic vulnerability is structurally distinct from, and in important respects more acute than, the domain loss discussed in Section 5, because it implicates the foundational transmission mechanism on which all other dimensions of vitality ultimately depend.

7. Script Erosion and Digital Romanization— A further, comparatively under-examined dimension of domain vulnerability concerns the Bengali script itself. In digital messaging, social media, and informal online communication, both within Bengal and in the diaspora, Bengali is frequently rendered in Romanized transliteration - colloquially termed ‘Benglish’ - rather than in the Bengali (Eastern Nagari) script, driven historically by early mobile keyboards and typing interfaces that lacked native Bengali input support and, more recently, by habit and platform convention even where Unicode Bengali keyboards are now freely available. While this practice does not threaten spoken transmission of the language, it constitutes a form of domain loss specific to literacy and written culture, and raises longer-term questions about script literacy among younger, digitally socialized speakers whose primary written exposure to their own language may increasingly occur outside its native orthography.

8. Historical Counter-Narrative : The 1952 Language Movement as a Model of Linguistic Resilience— Bengali's contemporary vitality cannot be understood without reference to the Bhasha Andolan, or Language Movement, of 1952 — arguably the twentieth century’s most consequential instance of successful reversal of language shift through organized political action. When the state of Pakistan sought to impose Urdu as the sole national language over the Bengali-majority population of East Pakistan (present-day Bangladesh), student-led protests in Dhaka culminated on 21 February 1952 in police firing on demonstrators, producing a set of deaths that became a foundational national memory (Umar, 1970, 1976, 1981). The movement secured constitutional recognition of Bengali as a state language of Pakistan in 1956 and subsequently became a central ideological current in the Bangladesh Liberation War of 1971 (Umar, 2004, 2006). Its long-term institutional legacy — UNESCO’s 1999 declaration of 21 February as International Mother Language Day — makes the 1952 movement a documented historical case in which deliberate collective action reversed an active trajectory of language subordination, offering a instructive counterpoint within Fishman's (1991) reversing-language-shift framework: it demonstrates that domain loss and institutional marginalization, even when backed by state power, are not irreversible, provided sufficient community mobilization and, eventually, institutional and constitutional support follow.

9. Discussion : Toward a Domain-Specific Model of Endangerment— The evidence assembled above resists a single-line verdict on Bengali’s endangerment status, and this paper argues that resisting such a verdict is itself the correct scholarly conclusion. A domain-specific model - distinguishing the vitality of the standard, codified variety from that of subordinate dialects; distinguishing home and community transmission from institutional and digital domain presence; and distinguishing the situation

OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

of home-territory speakers from that of diaspora speakers - accommodates the empirical reality that Bengali is simultaneously one of the world's most secure languages by demographic and constitutional measures and a language experiencing genuine, documentable vulnerability along specific, non-overlapping axes: dialectal (Sylheti, Chittagonian, and related varieties), institutional-domain (education, judiciary, science), technological (low-resource status in NLP), and intergenerational-diasporic (heritage transmission decline). This reframing has direct implications for language-endangerment scholarship more broadly: it suggests that large, demographically "safe" languages should not be exempted by default from vitality monitoring, and that UNESCO-style factor-based assessment should, where feasible, be applied at the level of specific domains and specific speaker populations rather than solely at the level of the named macro-language.

10. Policy Recommendations— Three areas of intervention follow from the analysis above. First, in education and documentation policy, Bangladesh and West Bengal's education authorities, in coordination with linguistic researchers, could extend formal documentation and, where communities request it, literacy support to vulnerable and endangered varieties within the Bengali-Assamese continuum, consistent with UNESCO's documentation-adequacy factor, without displacing the standard variety's role in formal education. Second, in digital-language planning, investment in large-scale, ethically sourced Bengali corpora and community-partnered NLP development, following models such as human-centred data-collection frameworks proposed in recent computational-linguistics scholarship (Ghosh & Caliskan, 2023), would directly address Bengali's low-resource classification and its attendant risks of future domain loss. Third, in diaspora heritage-language policy, formal recognition and funding support for community-run Bengali supplementary schools, alongside research-informed heritage-language pedagogy attentive to the specific bilingual profiles of second- and third-generation learners, would strengthen the intergenerational-transmission dimension identified as most acutely at risk in Section 6.

11. Conclusion— Bengali stands as an instructive case for the theory of language endangerment precisely because it does not fit comfortably within either the 'safe' or 'endangered' category as conventionally applied. Its scale, constitutional status, and unbroken literary continuity place it among the world's most secure languages; yet dialectal varieties within its own linguistic family are formally endangered, its standard form is losing ground in specific high-prestige and technological domains, and its diaspora communities face the same intergenerational attrition risks documented across heritage-language contexts worldwide. A multidimensional, domain-specific model of endangerment - rather than a single categorical label - offers both a more accurate description of Bengali's present condition and a more actionable basis for language planning. The 1952 Language Movement remains a powerful historical reminder that such vulnerabilities, however serious, are not fixed: where domain loss has been successfully reversed before through documented collective and institutional action, comparable interventions in education, digital infrastructure, and diaspora transmission remain both plausible and, on the evidence presented here, urgently warranted.

Bibliography:

Bengali Language. New World Encyclopedia, New World Encyclopedia contributors, www.newworldencyclopedia.org/entry/ Bengali_language. Accessed 1 Sept. 2026.

Eberhard, David M., et al., editors. Ethnologue : Languages of the World. 27th ed., SIL International, 2024.

Fishman, Joshua A. Reversing Language Shift: Theoretical and Empirical Foundations of Assistance to Threatened Languages. Multilingual Matters, 1991.

Ghosh, Sourojit, and Aylin Caliskan. "ChatGPT Perpetuates Gender Bias in Machine Translation and Ignores Non-Gendered Pronouns: Findings across Bengali and Five Other Low-Resource Languages." Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, Association for Computing Machinery, 2023, pp. 901–12, doi.org/10.1145/3600211.3604672.

OPEN ACCESS
IF 8.111
Trisangam International Refereed Journal (TIRJ)
A Double-Blind Peer Reviewed Research Journal on Language, Literature & Culture
Volume - vi, Issue – iv, October 2026, TIRJ/ October 2026/article - 84
Website: https://tirj.org.in/tirj, Page No. 747 - 752
Published issue link: https://tirj.org.in/tirj/issue/archive

Ghosh, Sourojit, and Srishti Chatterjee. "Misgendering and Assuming Gender in Machine Translation When Working with Low-Resource Languages." Gendered Technology in Translation and Interpreting: Centering Rights in the Development of Language Technology, edited by Esther Monzó-Nebot and Vicenta Tasa-Fuster, Routledge, 2024.
Krauss, Michael. "The World's Languages in Crisis." Language, vol. 68, no. 1, 1992, pp. 4–10.
"List of Endangered Languages in Bangladesh." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/List_of_endangered_languages_in_Bangladesh. Accessed 1 Sept. 2026.
Moseley, Christopher, editor. Atlas of the World's Languages in Danger. 3rd ed., UNESCO Publishing, 2010.
"Sylheti Language." Wikipedia, Wikimedia Foundation, en.wikipedia.org/wiki/Sylheti_language. Accessed 1
Sept. 2026.
Umar, Badruddin. The Emergence of Bangladesh: Class Struggles in East Pakistan. 2 vols., Oxford UP, 2004–06.
Umar, Badruddin. Purba Banglar Bhasha Andolan o Tatkalin Rajniti [The Language Movement in East Bengal and Contemporaneous Politics]. 3 vols., 1970–81.
UNESCO Ad Hoc Expert Group on Endangered Languages. Language Vitality and Endangerment. UNESCO, 2003.