



Writing Malayalam in Roman Script: Manglish, Digital Orthography and Linguistic Identity in Kerala

Josly Joseph

HST, St. Joseph's Anglo-Indian Girls' Higher Secondary School, Calicut, India

Article Information

Received: 4th June 2026

Received in revised form: 2nd July 2026

Accepted: 17th August 2026

Available online: 7th October 2026

Volume: 2

Issue: 4

DOI: <https://doi.org/10.63090/IJLL/3049.3242.0039>

Abstract

Much everyday written Malayalam on phones and social media is no longer written in the Malayalam script but in Roman letters, a practice widely known as "Manglish." Public debate in Kerala often treats Manglish as a sign of declining script literacy and of English eroding the mother tongue. This paper argues instead that Manglish is best understood as an informal, user-created orthography with its own conventions and social meanings. Drawing on research that treats orthography as social practice (Florian Coulmas, Mark Sebba), digital writing as vernacular literacy (David Crystal, Jannis Androutsopoulos, Jan Blommaert) and code-mixing as a normal feature of multilingual communication (Braj Kachru, Carol Myers-Scotton), it analyses the main features of Manglish spelling, including digraphs such as "zh" and "nj," the collapsing of dental and retroflex distinctions, inconsistent and expressive marking of vowel length, and widespread variation. It shows how the Roman script removes the visible boundary between Malayalam and English and so supports fluid code-mixing. The paper also places Manglish within the longer history of Malayalam written in several scripts, including Arabi-Malayalam, Suriyani Malayalam and the reformed script of 1971. It considers why Manglish has spread, among young people and the Gulf diaspora in particular, and concludes with implications for language education, language technology and debates about linguistic identity in Kerala.

Keywords:- Manglish, Malayalam, Orthography, Romanisation, Digital Communication, Code-Mixing, Sociolinguistics, Script, Linguistic Identity, Kerala

Introduction

Anyone who reads the messages Malayalis send one another on their phones will notice something striking. Much of what is written in Malayalam is not written in the Malayalam script at all. A friend confirms a plan with "sheri, njan varam" ("okay, I'll come"); a comment under a film trailer reads "pwooli item machaane"; a family group shares news of a wedding in a mixture of Malayalam words and English phrases, all in Roman letters. This practice, widely known as "Manglish," a blend of "Malayalam" and "English," has become one of the most common ways of writing Malayalam in everyday digital communication, especially among young people and the diaspora.

Manglish is often discussed in public debate as a sign of decline: evidence that young Malayalis are losing touch with their script, or that English is eroding the mother tongue. Such anxieties are understandable in a state that takes pride in its literacy, its literature and its language, which was granted classical language status by the Government of India in 2013. Yet

treating Manglish only as a threat leaves important questions unasked. How does Manglish actually work as a writing system? Why has it spread so widely, and what does it allow its users to do? And what does its history, and the longer history of Malayalam being written in different scripts, tell us about the relationship between a language and the letters used to write it?

This paper addresses these questions through a review of research on orthography, digital writing and code-mixing, combined with an analysis of the main features of Manglish spelling. It argues that Manglish is best understood not as a corruption of Malayalam but as an informal, user-created orthography: a set of conventions developed from below, shaped by the constraints of the Roman keyboard and the sound system of Malayalam, and carrying social meanings of its own. It also argues that Manglish belongs to a long tradition in which Malayalam has been written in several scripts by different communities. The paper concludes by considering the implications of Manglish for language education, for language technology and for debates about linguistic identity in Kerala.

Theoretical Framework

Orthography as Social Practice

Linguists have long distinguished between a language and its writing system. Florian Coulmas emphasises that writing systems are human inventions, adapted and borrowed across languages, and that the relation between sound and symbol in any system is a matter of convention rather than nature (Coulmas). Mark Sebba goes further, arguing that orthography is a social practice as much as a technical one. Choices about how to spell, which script to use and which conventions to follow carry meanings about identity, authority and belonging (Sebba). Spelling reforms, script changes and non-standard spellings are therefore sites where social and political relations are negotiated. From this perspective, the question to ask about Manglish is not only whether its spellings are "correct" but what its users are doing, socially, when they choose them.

Digital Writing and Vernacular Literacy

Research on computer-mediated communication has shown that digital writing often departs from standard norms in systematic ways. David Crystal, writing about text messaging in English, argued against the widespread fear that abbreviated spellings were damaging young people's literacy, pointing out that creative non-standard spelling requires an understanding of the standard it plays against (Crystal). Jannis Androutsopoulos describes digital media as spaces where "vernacular" writing, informal, speech-like and often multilingual, becomes visible on a large scale for the first time, and where new norms emerge from the practices of users rather than from institutions (Androutsopoulos). Jan Blommaert's work on "grassroots literacy" similarly draws attention to forms of writing produced outside the institutions that define standard literacy, and argues that they should be studied on their own terms rather than measured only by their distance from the norm (Blommaert). Manglish fits closely into this picture. It is a vernacular orthography that has grown up in the spaces of digital communication, without official codification.

Code-Mixing in Indian Englishes

Manglish is also closely tied to the mixing of Malayalam and English. Braj Kachru's influential account of English in India described the "nativisation" of English in South Asian contexts and the routine mixing of English with Indian languages in everyday speech and writing (Kachru). Carol Myers-Scotton's Matrix Language Frame model offers a way to analyse such mixing, distinguishing a "matrix language," which supplies the grammatical frame of a clause, from an "embedded language," which supplies inserted words and phrases (Myers-Scotton). In much Manglish writing, Malayalam is the matrix language, supplying verb

endings, case markers and word order, while English supplies nouns, adjectives and set phrases. Because both languages are written in the same Roman script, the boundary between them, which would be visible on the page if Malayalam were written in its own script, is no longer marked visually.

Malayalam in Many Scripts: A Historical Perspective

The idea that Malayalam belongs naturally and exclusively to the Malayalam script is itself relatively recent. Before the modern Malayalam script became standard, the language and its predecessors were written in a number of scripts, including Vatteluttu and Kolezhuthu, and the script now used developed in part from the Grantha script used for Sanskrit in South India. Different communities also developed their own ways of writing Malayalam. Arabi-Malayalam, the Malayalam language written in an adapted Arabic script, was used by the Mappila Muslim community of the Malabar coast for religious texts, songs and literature. Suriyani Malayalam, also called Karshoni, was Malayalam written in Syriac script by the St Thomas Christian community. In each case, the choice of script expressed religious and cultural affiliation as well as practical convenience.

The modern Malayalam script has itself undergone significant change. In 1971 the Kerala government introduced a reformed script intended to simplify printing and typing by replacing many complex conjunct letters and vowel signs with more regular, separable forms. The reform was widely adopted in schools and printing, but it was never universally accepted. With the arrival of Unicode and digital fonts, it became possible to display the traditional forms again, and volunteer groups such as Swathanthra Malayalam Computing developed free fonts and input tools that support both traditional and reformed orthography. Today, readers encounter a mixture of styles, and the "correct" form of the script remains a matter of debate.

Seen against this history, Manglish is not an unprecedented break. It is the latest instance of Malayalam being written in a script designed for another language, by users who adapt that script to their own needs. What distinguishes Manglish from Arabi-Malayalam or Suriyani Malayalam is that it is not tied to a particular religious community, and that it has developed almost entirely without institutional support, through millions of individual acts of typing.

Features of Manglish Spelling

Manglish has no official rules, but its spellings are far from random. They follow broadly shared conventions, many of which can be traced to the earlier romanisation of Malayalam place names and personal names in colonial and post-colonial administration. These conventions solve, in different ways, the problem of representing the sounds of Malayalam with the twenty-six letters of the English alphabet. Table 1 sets out some of the most important correspondences, comparing the Malayalam letter with its usual pronunciation, the scholarly transliteration in the ISO 15919 standard and common Manglish spellings.

Table 1. Some correspondences between Malayalam script, standard transliteration and Manglish

Malayalam letter	Sound (IPA)	ISO 15919	Common Manglish spelling	Example
ഴ	/ɻ/ retroflex approximant	ṛ	zh (also l)	Kozhikode, Alappuzha
ഞ	/ɲ/ palatal nasal	ṇ	nj	njan ("I")
ത	/t̪/ dental stop	t	th	Thiruvananthapuram
ട	/ɖ/ retroflex stop	ṭ	t, d	vettam, veedu ("house")
ള	/ɭ/ retroflex lateral	ḷ	l	Kerala (കേരളം)

റ / റ	/r/ trill and /r/ tap	ṛ / r	r for both	Distinction usually lost
ന്ദ്	alveolar /nd/ cluster	nṛ	nt, nd	ente ("my")
ആ (long a)	/a:/ long vowel	ā	aa or a	njaan / njan

Several patterns stand out. First, Manglish uses digraphs, pairs of letters, to represent sounds that English spelling lacks. The most famous is "zh" for the retroflex approximant ʒ , a sound often described as distinctive of Malayalam and Tamil, and familiar from place names such as Kozhikode and Alappuzha. The combination "nj" for the palatal nasal ɲ , as in "njan" ("I"), is another. These digraphs are well established and widely understood.

Second, Manglish often collapses distinctions that the Malayalam script maintains. Malayalam distinguishes dental and retroflex consonants, such as ṭ and Ṣ , and two kinds of "l" and "n" sounds, but Manglish frequently uses the same Roman letter for both members of each pair. Writers typically use "th" for the dental ṭ , as in "Thiruvananthapuram," and "t" or "d" for the retroflex Ṣ . Many other distinctions are left to context. A reader who knows Malayalam can usually recover the intended word without difficulty, much as readers of English recover the different sounds of "th" in "thin" and "this."

Third, vowel length is marked inconsistently. Malayalam distinguishes short and long vowels, and the difference can change meaning. Manglish writers sometimes double the vowel ("njaan," "sheriii") and sometimes do not ("njan"), and doubling may also be used expressively, to convey emphasis or emotion, in the way that repeated letters are used in English digital writing. This expressive use shows that Manglish spelling carries social and affective meaning as well as phonological information.

Fourth, there is considerable variation between writers and even within a single writer's messages. The popular slang word meaning "excellent" or "awesome," for instance, appears as both "poli" and "pwoli," the second spelling capturing a rounded pronunciation that the first does not. Such variation is typical of vernacular orthographies, and it is precisely the kind of feature that Sebba argues should be studied as social practice rather than dismissed as error (Sebba). Certain spellings become markers of style, region or group membership.

Manglish and Code-Mixing

Manglish is rarely written in "pure" Malayalam. In most messages, Malayalam words and grammatical endings mix with English words and phrases. A sentence such as "meeting cancel aayi" ("the meeting got cancelled") takes the English words "meeting" and "cancel" and builds them into a Malayalam clause with the verb "aayi" ("became"). In Myers-Scotton's terms, Malayalam is the matrix language here, providing the grammatical frame, while English is embedded (Myers-Scotton). This kind of mixing is common in spoken Malayalam, especially among educated urban speakers, and Manglish simply transfers it to writing.

The Roman script makes such mixing easy. A writer does not need to switch between keyboards or scripts in the middle of a sentence, and English words can be written in their familiar spellings. When the same sentence is written in Malayalam script, English words must either be transliterated into Malayalam letters or left in Roman script, producing a visible switch on the page. Manglish removes this visible boundary, and in doing so it may encourage more fluid mixing. This is one reason why researchers in language technology have found Malayalam-English code-mixed text on social media difficult to process automatically. Datasets developed for tasks such as sentiment analysis on YouTube comments have had to deal with Malayalam written in Roman script, in Malayalam script and in mixtures of both, together with English words and inconsistent spellings (Chakravarthi et al.).

Identity, Convenience and Ideology

Why has Manglish spread so widely, given that Malayalam keyboards and input tools are freely available? Several factors are relevant. The first is convenience. Most mobile phones and computers default to a Roman keyboard, and many users learned to type on such keyboards. Typing in Malayalam requires installing or switching to a different input method, and some users find it slower. Transliteration tools that convert Roman input into Malayalam script reduce this barrier, but they still require an extra step.

The second factor is education and literacy. Many young Malayalis, especially those who attended English-medium schools, read and write English more fluently than Malayalam script, even when Malayalam is their home language. For these users, Manglish allows them to write in the language they speak most naturally, using the script they know best. The same is true, to an even greater degree, for the large Malayali diaspora in the Gulf countries and elsewhere, where children may grow up speaking Malayalam at home but with limited instruction in reading and writing it. For such speakers, Manglish can be a means of maintaining written communication in Malayalam that might otherwise be lost.

The third factor is social meaning. Writing in Manglish signals informality, friendliness and a certain youthful, cosmopolitan identity. Writing in Malayalam script, by contrast, may be associated with formality, seriousness or literary style. Users often choose between the two according to context: Manglish for friends and casual exchanges, Malayalam script for formal announcements, literary writing or messages to older relatives. In this sense the two scripts function as different registers within the same linguistic repertoire, much as Androutsopoulos describes for other digital vernaculars (Androutsopoulos).

These practices exist alongside a strong public ideology of linguistic pride. Kerala has taken measures to strengthen the place of Malayalam in education and administration, and organisations such as the Malayalam Mission work to teach the language to children of the diaspora. Critics of Manglish argue that it weakens script literacy and makes the language seem less dignified. Supporters, or simply users, respond that Manglish keeps Malayalam alive in spaces where it might otherwise be replaced by English entirely. Both positions reflect real concerns. The sociolinguistic perspective adopted here suggests that the relationship between Manglish and Malayalam script is not simply competitive. Many users move between the two, and the spread of Manglish may coexist with continued attachment to the Malayalam script as a symbol of cultural identity.

Implications for Education and Technology

Recognising Manglish as a systematic vernacular orthography has practical implications. In language education, teachers can use Manglish as a bridge rather than treating it only as an error to be corrected. Students who habitually write in Manglish already possess an intuitive analysis of Malayalam sounds, since they must decide how to represent them in Roman letters. Comparing Manglish spellings with Malayalam script and with standard transliteration, as in Table 1, can make students aware of features of Malayalam phonology, such as the distinction between dental and retroflex consonants or the role of vowel length, that they might otherwise take for granted. For diaspora learners, Manglish may offer an entry point into reading and writing Malayalam, which can be extended gradually into the Malayalam script.

In language technology, Manglish poses both challenges and opportunities. Systems for search, translation, speech and sentiment analysis must be able to handle Malayalam written in Roman script, with its variable spellings and frequent mixing with English. Transliteration tools, which convert Roman input into Malayalam script, already rely on the conventions of Manglish, and their success depends on recognising the most common correspondences. Research on code-mixed Malayalam-English text has begun to address these challenges, but

there is a need for more work that combines computational methods with sociolinguistic understanding of how and why people write as they do (Chakravarthi et al.).

For researchers in linguistics and literature, Manglish also raises questions that this paper has only begun to address. Systematic corpus studies could measure the frequency of different spelling conventions and how they vary by age, region and platform. Ethnographic research could examine how users decide when to use Manglish and when to use Malayalam script. Literary scholars might explore the growing use of Manglish in creative writing, advertising, film titles and popular culture, where it is increasingly used deliberately for stylistic effect.

Conclusion

This paper has argued that Manglish, the writing of Malayalam in Roman script, is best understood as an informal, user-created orthography rather than as a simple corruption of Malayalam. Drawing on research that treats orthography as social practice, digital writing as vernacular literacy and code-mixing as a normal feature of multilingual communication, it has shown that Manglish follows broadly shared conventions, solves real problems of representation and carries social meanings of its own. It has also placed Manglish within a longer history in which Malayalam has been written in several scripts, including Arabi-Malayalam and Suriyani Malayalam, and in which the Malayalam script itself has been reformed and contested.

None of this means that concerns about script literacy are unfounded. The Malayalam script is a vital part of Kerala's cultural heritage, and the ability to read and write it gives access to a rich literary tradition. But the choice is not necessarily between Manglish and Malayalam script. Many users move between the two according to context, and Manglish may help sustain written Malayalam among speakers who would otherwise write only in English. Understanding Manglish on its own terms is a first step toward policies, teaching practices and technologies that support Malayalam in all the forms in which its speakers actually use it. A language, as the history of Malayalam shows, is larger than any one of its scripts.

Works Cited

- Androutsopoulos, Jannis. "Language Change and Digital Media: A Review of Conceptions and Evidence." *Standard Languages and Language Standards in a Changing Europe*, edited by Tore Kristiansen and Nikolas Coupland, Novus Press, 2011, pp. 145–61.
- Blommaert, Jan. *Grassroots Literacy: Writing, Identity and Voice in Central Africa*. Routledge, 2008.
- Chakravarthi, Bharathi Raja, et al. "A Sentiment Analysis Dataset for Code-Mixed Malayalam-English." *Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL)*, European Language Resources Association, 2020, pp. 177–84.
- Coulmas, Florian. *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge UP, 2003.
- Crystal, David. *Txtng: The Gr8 Db8*. Oxford UP, 2008.
- Kachru, Braj B. *The Alchemy of English: The Spread, Functions, and Models of Non-native Englishes*. Pergamon, 1986.
- Myers-Scotton, Carol. *Duelling Languages: Grammatical Structure in Codeswitching*. Clarendon Press, 1993.
- Sebba, Mark. *Spelling and Society: The Culture and Politics of Orthography around the World*. Cambridge UP, 2007.