PREFACE TO THE EDITION

The current issue of the **International Journal of Linguistics, Language and Literature (IJLLL)** brings together five thought-provoking contributions that examine language and literature as dynamic expressions of cultural identity, ecological memory, human and nonhuman experience, and technological transformation. Spanning digital linguistic practices, environmental literary criticism, posthumanist narratology, and artificial intelligence in language revitalization, the articles reflect the expanding scope of contemporary linguistic and literary inquiry. Collectively, they demonstrate how language and literary texts preserve community knowledge, respond to changing social and environmental realities, and raise important questions about representation, ownership, and the future of linguistic diversity.

The first article, "Writing Malayalam in Roman Script: Manglish, Digital Orthography and Linguistic Identity in Kerala," examines the widespread use of Roman-script Malayalam in digital communication. Challenging the perception that Manglish necessarily represents a decline in script literacy or the erosion of the mother tongue, the study approaches it as an evolving form of informal orthography shaped by user practices, multilingual interaction, and social identity. By exploring its spelling conventions, historical context, and growing use among young people and the Gulf diaspora, the article contributes to broader discussions on digital literacy, code-mixing, and the changing relationship between language, script, and identity.

Extending the discussion of language into the domain of environmental memory, the second article, "Words for Water: Wetland Lexicons, Flood Memory and Ecolinguistics in Kerala Fiction," investigates how literary works preserve ecological knowledge through the vocabulary of floods, wetlands, rivers, canals, and agricultural landscapes. Through readings of selected works by Thakazhi Sivasankara Pillai, Arundhati Roy, and Sarah Joseph, the essay demonstrates how regional languages and literary narratives can serve as archives of collective environmental experience. It also highlights the relationship between ecological degradation and the loss of place-specific vocabulary, emphasizing the value of ecolinguistic approaches to understanding the connections between language, community memory, and environmental change.

Complementing this regional perspective, the third article, "Toxic Bodies, Drowning Islands: Slow Violence and Vernacular Ecologies in South Asian Environmental Fiction," examines how environmental harm is represented through embodied suffering, local linguistic practices, and the struggles of marginalized communities. Drawing on selected works by Amitav Ghosh, Indra Sinha, Ambikasuthan Mangad, and Mahasweta Devi, the essay explores how environmental crises become visible through particular bodies, landscapes, and experiences rather than solely through abstract accounts of planetary change. Its discussion of slow violence and vernacular ecologies underscores the importance of literary representation in revealing the unequal social consequences of environmental degradation.

The fourth contribution, "Posthumanism and the Nonhuman Voice in Contemporary Fiction," turns to the representation of nonhuman perspectives in contemporary literary narratives. Bringing together posthumanist thought, new materialism, and theories of nonhuman narration, the article examines how artificial and ecological voices challenge conventional assumptions about narrative authority and human experience. It argues that such narratives do not simply provide direct access to nonhuman consciousness; instead, they expose the limitations of human language and imagination when representing other forms of existence. In doing so, the study invites readers to reconsider the ethical and aesthetic possibilities of fiction beyond an exclusively human-centred perspective.

The issue concludes with “Stewardship, Not Extraction: Large Language Models and Endangered Language Revitalization,” which considers the opportunities and ethical challenges associated with applying artificial intelligence to the documentation and revitalization of endangered and Indigenous languages. Focusing on large language models, speech recognition, speech synthesis, and machine translation, the article examines the difficulties posed by limited linguistic data and the risks of treating community knowledge as an extractable technological resource. By discussing community-led documentation, Indigenous data sovereignty, and responsible licensing frameworks, it advances the argument that language technologies should support the priorities and rights of the communities whose languages they seek to preserve. Ultimately, the study emphasizes that technological innovation must complement, rather than replace, intergenerational language transmission.

Taken together, these contributions illustrate the interdisciplinary character of contemporary linguistics and literary studies. From the digital transformation of Malayalam writing to the preservation of ecological memory in fiction, from narratives that challenge human-centred perspectives to the ethical use of artificial intelligence in language revitalization, the articles reveal the enduring significance of language as a medium of cultural expression, knowledge production, and social negotiation. They also demonstrate that linguistic and literary scholarship can contribute meaningfully to current debates concerning digital communication, environmental sustainability, cultural plurality, and technological responsibility.

We hope that this issue will encourage researchers, scholars, educators, and students to engage with the diverse questions raised by these studies and to pursue further inquiry into the evolving relationships between language, literature, society, and technology. The editorial team extends its sincere appreciation to the authors for their contributions and to the reviewers for their valuable scholarly assessments. We remain committed to fostering rigorous academic dialogue and providing a platform for original research that advances understanding across the fields of linguistics, language, and literature.

Dr Mahesh Kumar Dey

Chief Editor

CONTENTS

SL. NO.	TITLE	AUTHOR	PAGE NO.
1	Writing Malayalam in Roman Script: Manglish, Digital Orthography and Linguistic Identity in Kerala	Josly Joseph	1-6
2	Words for Water: Wetland Lexicons, Flood Memory and Ecolinguistics in Kerala Fiction	Jisha Alex	7-13
3	Toxic Bodies, Drowning Islands: Slow Violence and Vernacular Ecologies in South Asian Environmental Fiction	Rini Joy	14-21
4	Posthumanism and the Nonhuman Voice in Contemporary Fiction	Rakhi Ramachandran	22-29
5	Stewardship, Not Extraction: Large Language Models and Endangered Language Revitalization	Annette Treesa Benny	30-38

INTERNATIONAL JOURNAL OF LINGUISTICS LANGUAGE AND LITERATURE (IJLLL)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-3242 ISSN Print: 3139-1990

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/IJLLL/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 "pwoli 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	ḻ	zh (also l)	Kozhikode, Alappuzha
ഞ	/ɲ/ palatal nasal	ñ	nj	njan ("I")
ത	/t̪/ dental stop	t	th	Thiruvananthapuram
ട	/ʈ/ retroflex stop	ṭ	t, d	vettam, veedu ("house")
ള	/ɭ/ retroflex lateral	ḷ	l	Kerala (കേരളം)

ർ / ര	/r/ trill and /ɾ/ tap	r̲ / r	r for both	Distinction usually lost
ന്റ	alveolar /nd/ cluster	nr̲	nt, nd	ente ("my")
ആ (long a)	/aː/ 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 ഴ, 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 ഞ, 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 ത and ട, 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 ത, as in "Thiruvananthapuram," and "t" or "d" for the retroflex ട. 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.

INTERNATIONAL JOURNAL OF LINGUISTICS LANGUAGE AND LITERATURE (IJLLL)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-3242 ISSN Print: 3139-1990

Words for Water: Wetland Lexicons, Flood Memory and Ecolinguistics in Kerala Fiction

Jisha Alex

Assistant Professor, Department of English, Marian College Kuttikkanam (Autonomous), India

Article information

Received: 4th July 2026
Received in revised form: 10th August 2026
Accepted: 12th September 2026
Available online: 7th October 2026

Volume: 2
Issue: 4
DOI: https://doi.org/10.63090/IJLLL/3049.3242.0040

Abstract

In August 2018, Kerala's worst floods in nearly a century were measured against an older disaster remembered by name: "the flood of ninety-nine," the great flood of 1924, called after its year in the Malayalam calendar. Taking this linguistic survival as its starting point, the essay argues that Kerala's literature of water works as a lexical and narrative archive of ecological knowledge. Recent criticism of climate fiction has emphasised the gap between human-sized narration and planetary process. This essay proposes that a regional language offers an intermediate scale, larger than any individual and older than any lifetime, through which environmental memory is carried. Drawing on ecolinguistics (Einar Haugen, Luisa Maffi, Arran Stibbe), Robert Macfarlane's work on place-words and Rob Nixon's concept of slow violence, it reads three texts: Thakazhi Sivasankara Pillai's short story "Vellappokkathil" ("In the Flood"), Arundhati Roy's The God of Small Things (1997) and Sarah Joseph's Malayalam novel Aathi (2011; translated by Valson Thampu as Gift in Green). It shows how these works preserve a vocabulary of floods, paddy fields, canals and embankments that encodes how water moves and how communities have lived with it, and how they register the degradation of wetlands and rivers as a loss of language. The essay concludes that an ecolinguistic reading, attentive to vocabulary, translation and community storytelling, can recover forms of environmental knowledge that criticism focused on English and on planetary scale tends to miss.

Keywords:- Ecolinguistics, Kerala Floods, Wetlands, Malayalam Literature, Place-Words, Translation, Slow Violence, Climate Fiction, Thakazhi, Arundhati Roy, Sarah Joseph

Introduction

In August 2018 Kerala experienced its worst floods in nearly a century. More than four hundred people died, hundreds of thousands were moved to relief camps, and large parts of Kuttanad, Chengannur and the Periyar basin remained under water for days. When journalists and residents searched for a comparison, they reached not for a statistic but for a name: thonnootti onpathile vellappokkam, "the flood of ninety-nine," the great flood of 1924, remembered by its year in the Malayalam calendar (1099 ME). The older flood had survived for almost a century less as a hydrological record than as a phrase, carried in speech, in family stories and in literature. When the waters rose again, it was that phrase that gave the new disaster a measure.

This small linguistic fact points to a larger argument. Criticism of climate fiction has tended to concentrate on scale and form: on how the novel can represent processes that exceed

individual perception. Rose Mary Philip's recent essay in this journal frames the genre's achievement in exactly these terms, as a "scalar disjuncture" between human-sized narration and planetary process (Philip 98). A companion essay on South Asian environmental fiction has argued that this disjuncture looks different when environmental harm is located in particular bodies and communities. The present essay extends that argument to a further level, the level of words. It asks what happens to environmental memory when it is stored not only in plots and characters but in a regional vocabulary of water, field and embankment, and what literature does when that vocabulary, and the landscape it names, begins to disappear.

The essay reads three works from Kerala that span almost a century of writing about water. Thakazhi Sivasankara Pillai's short story "Vellappokkathil" ("In the Flood"), set during the flood of 1924 in Kuttanad, is told largely from the perspective of a dog abandoned on a submerged roof. Arundhati Roy's The God of Small Things (1997) makes the Meenachal river at Ayemenem a boundary of caste and desire, and records its later degradation. Sarah Joseph's Malayalam novel Aathi (2011), translated by Valson Thampu as Gift in Green, imagines a wetland village resisting the reclamation of its waters. Read together, these texts suggest that Kerala's environmental literature works as a lexical archive. It preserves words that encode ecological knowledge, and it registers the loss of that knowledge as a loss of language. The argument draws on ecolinguistics, the study of how language shapes and reflects human relations with the environment, in order to show that the "form of attention" Philip attributes to climate fiction (Philip 104) is, in this regional tradition, also a form of naming.

Theoretical Framework: Ecolinguistics and the Lexicon of Place

The idea that language and environment are bound together has a long history in linguistics. Einar Haugen's influential essay "The Ecology of Language" proposed that a language should be studied in relation to its environment, by which he meant chiefly the society that uses it, and the minds of its speakers (Haugen). Later work broadened this metaphor into a literal claim. Luisa Maffi and her collaborators argued that linguistic, cultural and biological diversity are interrelated, since the places richest in species are often also richest in languages, and since the knowledge needed to live in a particular ecosystem is held largely in local speech (Maffi). When a community loses its language, or the specialised vocabulary through which it understands land and water, it loses a store of ecological knowledge built over generations.

Arran Stibbe's Ecolinguistics gives this insight a critical edge. Stibbe examines the "stories we live by," the frames, metaphors and patterns of wording that shape how a culture understands the natural world, and asks which of them encourage care and which encourage destruction (Stibbe). His method is especially useful for reading development discourse, in which wetlands appear as "wasteland" awaiting "reclamation," and rivers as "resources" to be "harnessed." Such terms are not neutral descriptions. They prepare the ground, sometimes literally, for landfill and concrete.

Robert Macfarlane's Landmarks adds a more literary dimension. Macfarlane collects regional words for terrain, weather and water from across Britain and Ireland, arguing that the loss of precise place-words is both a symptom and a cause of a thinning attention to the land (Macfarlane). He notes, as one sign of this thinning, the removal of common nature words from a children's dictionary. His project, part glossary and part elegy, offers a model for reading literature as a reservoir of such words. Lawrence Buell's account of place in environmental criticism supports this view. For Buell, place is not simply a location but a site of attachment and memory, made through experience and narrative over time (Buell).

These frameworks can be connected to the arguments about slow violence discussed in the earlier essay. Rob Nixon describes how the gradual destruction of environments is often accompanied by a destruction of the knowledge needed to recognise it (Nixon). When paddy fields are filled for housing and canals are blocked for roads, the words that named their

functions begin to fall out of use. The next generation may never learn that a particular low field was where floodwater used to go. The flood then appears as a sudden natural disaster rather than as the delayed consequence of a forgotten landscape. Literature that preserves such words is therefore doing something more than adding local colour. It keeps alive the conditions under which slow violence can be seen.

The Flood of '99: Thakazhi's "Vellappokkathil"

Thakazhi Sivasankara Pillai is best known for his novels of Kuttanad, the low-lying rice-growing region south of Vembanad lake, where farmers cultivate fields that lie below sea level behind earthen bunds. "Vellappokkathil" is one of his most widely read stories and remains a staple of Malayalam school syllabuses. It is set during the flood of 1924, when the waters of Kuttanad rose over fields, homes and temples. As the agricultural labourers flee in boats, a dog is left behind on the roof of a hut. Much of the story follows the dog as it waits for its owner, endures hunger and fear, and watches the flooded world around it, until it finally dies in the water.

The choice of a dog as the centre of consciousness is the story's most striking formal decision, and it anticipates in a modest way the decentring of the human that Philip finds in contemporary climate fiction (Philip 101). Yet Thakazhi's purpose is not to replace human with nonhuman agency. The dog's abandonment is a measure of the conditions of the poor. The labourers who flee are themselves vulnerable, living in huts on the margins of fields owned by others, with little to carry and nowhere secure to go. The flood exposes a social order in which both the labourer and his animal are expendable. The nonhuman perspective does not lead away from human politics; it brings its harshness into view from below.

The story also shows how a disaster becomes part of collective language. The flood of 1924 is not narrated as an abstract hydrological event but as a lived catastrophe in a named landscape of fields, bunds, canals and roofs. Thakazhi's Malayalam is rich in the vocabulary of Kuttanad's waterscape, a vocabulary that his readers would have recognised from daily life. Much of this precision is lost in English translation, where field, canal and embankment flatten distinctions that Malayalam keeps. The difficulty of translation here is itself evidence for the ecolinguistic argument: the knowledge of how Kuttanad's water moves is embedded in the words that name it.

When the floods of 2018 revived memories of the flood of '99, Thakazhi's story was one of the texts through which that memory had travelled. Its endurance in the school curriculum meant that generations of Malayali readers had encountered the earlier flood through literature, long after the last eyewitnesses had died. In this sense the story functions as a cultural archive of disaster. It keeps a flood in living memory not through records but through a narrative and a name.

A River Reduced: The God of Small Things

Arundhati Roy's The God of Small Things is not usually read as environmental fiction, but its structure depends on a river. The Meenachal flows past the Ipe family home in Ayemenem, and it divides the novel's world into permitted and forbidden spaces. Across the river lies the abandoned "History House"; it is the river that Velutha, a Paravan carpenter, crosses at night to meet Ammu, defying the caste rules the novel calls the "Love Laws"; and it is in the river that the visiting English cousin Sophie Mol drowns, setting off the catastrophe that destroys the family (Roy). The river is thus the medium through which the novel's social and moral boundaries are both crossed and enforced.

The novel's account of the river changes between its two time-frames. In 1969, the year of the central events, the Meenachal is a powerful, living river, dangerous in the monsoon. By 1993, when the adult Rahel returns to Ayemenem, it has been reduced. Roy describes how a

saltwater barrage built downstream, in exchange for the votes of the paddy-farming lobby, has altered the river's flow, so that it has become sluggish and polluted, and how the History House has been turned into a luxury hotel where shortened Kathakali performances are staged for tourists (Roy). The two degradations, of the river and of the art form, are presented side by side. Both are effects of a development model that converts living systems, ecological and cultural, into commodities.

This double movement makes the novel an early literary record of what Nixon calls slow violence. No single event destroys the Meenachal. Its decline is the cumulative result of engineering, agricultural policy, tourism and neglect, spread over two decades and largely invisible to those who cause it. The novel registers this change through a return: Rahel, coming back after years away, sees a difference that those who stayed have absorbed gradually. The structure of returning across time functions here much as Philip suggests the diary functions in Butler's Parable of the Sower, as a way of making slow change visible through comparison (Philip 102-03).

Roy's language contributes to this effect. The novel's English is famously experimental, full of compound words, capitalised phrases, children's wordplay and Malayalam terms left untranslated. This hybrid idiom resists the smooth, standard prose in which development is usually described. It insists that Ayemenem has its own words, its own ways of naming things, and it shows those words under pressure from English, from tourism and from the language of progress. The novel's attention to small things, which its title announces, is also an attention to small words: the local, the particular and the easily overlooked, which the large narratives of nation and development tend to erase.

Stories against Landfill: Sarah Joseph's Aathi

Sarah Joseph's Aathi is the most explicitly environmental of the three texts, and its focus is the wetland itself. The novel is set in an imagined village called Aathi, surrounded by water, whose inhabitants live by fishing, farming and close knowledge of the tides and seasons. Their way of life is threatened when a man who had left the village returns with money and plans for development: land filling, construction and the conversion of the wetland into real estate. The conflict that follows sets the village's understanding of water as a shared and living presence against a vision of it as empty space awaiting profit (Joseph, Gift in Green).

The novel's central formal device is storytelling. At regular gatherings the villagers tell stories to one another, and many chapters are built around these tales, which range across myth, history and personal memory. These embedded stories serve a function that the ecolinguistic frameworks make clear. They are the means by which the community remembers its relationship with water, transmits ecological knowledge and articulates its values. In Stibbe's terms, they are counter-stories to the dominant story of development, offering a different way of living with the environment (Stibbe). When the developers' language of progress, investment and reclamation enters Aathi, it does not only threaten the wetland. It threatens to replace the stories through which the village understands itself.

Aathi was written against the background of real struggles in Kerala over the conversion of paddy fields and wetlands. The Kerala Conservation of Paddy Land and Wetland Act of 2008 was passed in an effort to slow such conversions, and the debates surrounding it made the fate of the vayal, the paddy field, a matter of public concern. Kerala's wetlands, including the Vembanad-Kol system, which is listed under the Ramsar Convention, perform functions that are easy to overlook until they are lost: they absorb floodwater, recharge groundwater and support fisheries. Joseph's novel gives these functions a human and narrative form. The wetland is not a backdrop but a community's memory, and its destruction is presented as an act of forgetting.

The novel's English translation, Gift in Green, raises the same questions of translation encountered in Thakazhi's story. Joseph's Malayalam draws on a dense vocabulary of water, fish, rice and seasons, and some of this inevitably thins in English. Yet the translation also extends the novel's reach, carrying its argument to readers who have never seen a Kerala wetland. This tension, between the local knowledge carried in a regional language and the wider audience reached through translation, runs through all the texts discussed here. It is also the tension at the heart of environmental literature in India, where the most detailed knowledge of particular landscapes is usually held in languages that metropolitan criticism rarely reads.

Discussion: A Lexicon at Risk

The three texts span the period from the 1924 flood to the eve of the 2018 floods, and together they trace a change in Kerala's relationship with water. In Thakazhi, the flood is a natural catastrophe whose harshness is shaped by social inequality. In Roy, the river is degraded by engineering and tourism. In Joseph, the wetland is targeted for erasure by real estate. Across this sequence, what is threatened is not only the landscape but the language used to understand it. Table 1 sets out a small part of the Malayalam vocabulary of water and field that these texts, and the landscapes they describe, depend on.

Table 1. Some Malayalam words of water and field, and the knowledge they carry

Malayalam term	**Approximate English gloss**	**Ecological knowledge encoded**	**Present pressure**
Vellappokkam	Flood	Memory of past floods, named by year (the flood of 1099 me)	Memory fading as eyewitness generations pass
Vayal	Paddy field	Low land that holds floodwater and recharges groundwater	Filling for housing and roads
Kayal	Backwater, lagoon	Brackish mixing of river and sea; fisheries	Reclamation, tourism, pollution
Thodu	Canal, small stream	Drainage paths linking fields and rivers	Blocking and encroachment
Chira	Earthen bund, embankment	Collective management of water levels	Breaching, replacement by concrete
Punja	Kuttanad's dry-season rice crop	A farming calendar timed to water levels	Altered salinity and flood regimes
Pokkali	Salt-tolerant rice of coastal wetlands	Rotation of rice and prawn with the tides	Declining cultivated area

Each of these words names not just a thing but a relationship. A chira is not merely a wall of earth; it is the outcome of a collective practice of managing water levels across many fields. Pokkali is not merely a variety of rice; it names a system of rotating rice and prawn cultivation adapted to the tidal salinity of coastal wetlands. When such systems decline, the words begin to drift out of everyday use, surviving in literature, in agricultural records and in the speech of older people. The landscape loses its legibility for those who no longer know what a low field, a canal or an embankment was for.

This loss had material consequences in 2018. Analyses of the floods drew attention to the role of encroached floodplains, filled paddy fields and blocked drainage channels in worsening the damage, alongside extreme rainfall and the management of reservoirs. The

details remain debated, but the general point was widely accepted: some of the land that flooded had once been understood, and named, as land that floods. A vocabulary that encoded where water should go had been weakened along with the landscape it described.

The floods also revealed the value of practical, local knowledge. Among the most celebrated rescuers were fishermen from Kerala's coastal communities, who brought their boats inland and used their knowledge of water to reach people whom official teams could not. Their skill was not written down in manuals; it was held in practice and in the vocabulary of their work. In this respect the floods enacted, in reality, the argument these texts make in fiction: that local knowledge of water, carried in language and in community, is a form of environmental intelligence that development discourse ignores at its peril.

For literary criticism, this suggests a refinement of the scalar model. Philip argues that climate fiction mutates realism in order to register processes larger than the individual (Philip 104). The Kerala texts suggest that one of the most important scales between the individual and the planet is the scale of a regional language. A language is larger than any single speaker, older than any lifetime and closely tied to a specific landscape. It can carry memory of a flood across a century, as the phrase "the flood of ninety-nine" did. Attending to this scale means reading environmental literature not only for its plots and images but for its words.

Conclusion

This essay has argued that Kerala's literature of water functions as a lexical and narrative archive of ecological knowledge. Thakazhi's "Vellappokkathil" kept the flood of 1924 alive in the collective memory through a story read by generations of schoolchildren. Roy's The God of Small Things recorded the slow degradation of a river and showed how a hybrid literary language could resist the smooth idiom of development. Joseph's Aathi made storytelling itself the means by which a wetland community defends its water and its memory. In each case, the work of literature includes the preservation of words: names for floods, fields, canals and embankments that encode how water moves and how people have lived with it.

The relevance of this argument extends beyond Kerala. In many parts of South Asia and the global South, the most detailed knowledge of local environments is held in regional languages and oral traditions that are themselves under pressure. Environmental criticism that reads only in English, or only at the scale of the planet, risks missing that knowledge altogether. An ecolinguistic approach to literature, attentive to vocabulary, translation and the stories communities tell about their places, offers one way to bring it into view. For a journal of linguistics, language and literature, the point is perhaps especially apt. Before a landscape is lost, the words that name it often go first, and literature is one of the places where they can still be found.

Works Cited

Buell, Lawrence. The Future of Environmental Criticism: Environmental Crisis and Literary Imagination. Blackwell, 2005.

Haugen, Einar. “The Ecology of Language.” The Ecology of Language: Essays by Einar Haugen, edited by Anwar S. Dil, Stanford UP, 1972, pp. 325–39.

Joseph, Sarah. Aathi. Current Books, 2011.

---. Gift in Green. Translated by Valson Thampu, HarperCollins India, 2011.

Macfarlane, Robert. Landmarks. Hamish Hamilton, 2015.

Maffi, Luisa, editor. On Biocultural Diversity: Linking Language, Knowledge, and the Environment. Smithsonian Institution Press, 2001.

Nixon, Rob. Slow Violence and the Environmentalism of the Poor. Harvard UP, 2011.

Philip, Rose Mary. “Narrating the Anthropocene: Scale, Agency, and Form in Climate Fiction.” International Journal of Linguistics, Language and Literature, vol. 2, no. 3, 2026, pp. 98–105, https://doi.org/10.63090/IJLLL/3049.3242.0037.

Pillai, Thakazhi Sivasankara. “Vellappokkathil” [“In the Flood”]. Thakazhiyude Cherukathakal [Short Stories of Thakazhi], DC Books.
Roy, Arundhati. The God of Small Things. IndiaInk, 1997.
Stibbe, Arran. Ecolinguistics: Language, Ecology and the Stories We Live By. Routledge, 2015.

INTERNATIONAL JOURNAL OF LINGUISTICS LANGUAGE AND LITERATURE (IJLLL)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-3242 ISSN Print: 3139-1990

Toxic Bodies, Drowning Islands: Slow Violence and Vernacular Ecologies in South Asian Environmental Fiction

Rini Joy

Research Scholar, Sree Shankaracharya University of Sanskrit, Kalady, Kerala, India

Article information

Received: 6th July 2026
Received in revised form: 4th August 2026
Accepted: 11th September 2026
Available online: 7th October 2026

Volume: 2
Issue: 4
DOI: https://doi.org/10.63090/IJLLL/3049.3242.0041

Abstract

Recent criticism of climate fiction has framed the genre's central problem as one of scale: the gap between human-sized narration and planetary process. This essay argues that such a framing, built largely on Euro-American novels, leaves out a second and equally formative problem, which is the problem of location. In South Asian environmental fiction, the Anthropocene seldom appears as an unrepresentable totality. It arrives instead as poison in a particular body, salt water in a particular village, and a pesticide sprayed over a particular hillside. Drawing on Rob Nixon's concept of slow violence, the "environmentalism of the poor" theorised by Ramachandra Guha and Joan Martinez-Alier, and Stacy Alaimo's account of trans-corporeality, the essay reads four texts: Amitav Ghosh's The Hungry Tide (2004), Indra Sinha's Animal's People (2007), Ambikasuthan Mangad's Malayalam novel Enmakaje (2009; translated by J. Devika as Swarga, 2017), and Mahasweta Devi's novella "Pterodactyl, Puran Sahay, and Pirtha" (translated from Bengali by Gayatri Chakravorty Spivak in Imaginary Maps, 1995). It identifies three formal strategies th rough which these works narrate environmental harm: the damaged body as narrator and archive; the vernacular, multilingual texture that ties ecological knowledge to local speech; and the conflict between conservation and survival, in which the nonhuman is never innocent of human politics. The essay concludes that South Asian environmental fiction does not so much resolve the scalar disjuncture of the Anthropocene as relocate it, measuring planetary process against the endurance of the marginal body.

Keywords:- Slow Violence, Postcolonial Ecocriticism, Environmentalism of the Poor, Trans-corporeality, Indian Fiction, Malayalam Literature, Endosulfan, Climate Fiction, Subaltern

Introduction

In a recent essay in this journal, Rose Mary Philip argues that the achievement of climate fiction lies in its formal reorganisation around the impossibility of representing environmental catastrophe directly. Reading six novels, among them Richard Powers's The Overstory, Jeff VanderMeer's Annihilation and Kim Stanley Robinson's The Ministry for the Future, she identifies the genre's defining problem as a "scalar disjuncture" between human-sized narration and planetary-scale process (Philip 98). Her account is persuasive, and its central claim, that realism in these novels is mutated rather than abandoned, offers a useful corrective to Amitav Ghosh's pessimism about the modern novel in The Great Derangement. Yet the archive on which the argument rests is telling. With the exception of Ghosh's Gun Island, the novels are North American, and the problem they share is framed from the vantage of a reader for whom climate change is, in the first instance, an abstraction that must be made perceptible.

This essay begins from a different vantage. For a great many people in South Asia, environmental harm is not abstract at all. It is lodged in the deformed limbs of children born near pesticide-sprayed plantations, in lungs damaged by gas that leaked from a factory decades ago, and in islands that are slowly eaten by the sea. The difficulty such harm poses to narrative is not that it cannot be seen, but that it is seen by those whose testimony carries little weight, in languages that rarely reach metropolitan readers, over periods too long to sustain public outrage. The representational problem, in other words, is one of location as much as scale: of where, and in whose body, environmental violence becomes visible, and of who has the authority to say so.

The essay pursues this argument through four South Asian texts: Amitav Ghosh's The Hungry Tide (2004), set in the Sundarbans; Indra Sinha's Animal's People (2007), a fictional rewriting of the Bhopal gas disaster; Ambikasuthan Mangad's Enmakaje (2009), a Malayalam novel about endosulfan poisoning in Kasaragod, Kerala; and Mahasweta Devi's "Pterodactyl, Puran Sahay, and Pirtha," a novella of famine and dispossession among Adivasi communities. Only one of these, Ghosh's novel, is usually shelved as climate fiction. The others concern toxicity, industrial disaster and famine. Including them is deliberate. In postcolonial contexts, the boundaries that Euro-American criticism draws between climate change, pollution and development-induced dispossession are hard to sustain, since all three fall most heavily on the same people and operate through the same structures of extraction. I argue that these texts develop three formal strategies for narrating such harm: they make the damaged body a narrator and an archive; they root ecological knowledge in vernacular and multilingual speech; and they stage conflicts between conservation and survival in which the nonhuman is never a politically neutral protagonist. Together, these strategies do not dissolve the scalar disjuncture that Philip describes. They relocate it, so that planetary process is measured against the endurance of the marginal body.

Theoretical Framework: From Planetary Scale to the Environmentalism of the Poor

Rob Nixon's Slow Violence and the Environmentalism of the Poor is often cited in discussions of climate fiction for its account of tempo, of harm that unfolds gradually and out of sight. Philip uses it in precisely this way, as one of three theories, alongside Timothy Morton's hyperobjects and Dipesh Chakrabarty's species history, that together describe the Anthropocene's excess over the human-sized story (Philip 99). But Nixon's book has a second emphasis that is easily lost when it is placed in that company. Its subtitle names the poor, and its central concern is not only that slow violence is hard to see but that it is unevenly distributed and unevenly witnessed. The people who live with its effects, Nixon argues, are frequently those whose suffering is least likely to be recognised by states, corporations and media, and whose resistance is least likely to be called environmentalism at all (Nixon 2-4). One of his most extended readings, notably, is of Sinha's Animal's People.

Nixon draws here on an older debate. In 1989 Ramachandra Guha criticised the American wilderness movement for exporting a model of conservation that, applied in India, displaced forest-dwelling communities in the name of protecting nature (Guha 71-83). Guha and Joan Martinez-Alier later developed the idea of an "environmentalism of the poor": a body of struggles, such as the Chipko movement or resistance to large dams, in which ecological concern arises from livelihood and survival rather than from aesthetic or recreational attachment to untouched nature (Guha and Martinez-Alier). This tradition matters for literary analysis because it changes what counts as an environmental narrative. A story about a fisherman's right to fish, or a mother's search for an explanation of her child's illness, may be as central to environmental literature as a story about a forest or a melting glacier.

Graham Huggan and Helen Tiffin, in Postcolonial Ecocriticism, describe the convergence of postcolonial and ecological criticism as productive but uneasy, since the first

has tended to prioritise human justice and the second the nonhuman world (Huggan and Tiffin 1-3). Pablo Mukherjee's Postcolonial Environments addresses the same tension through Indian fiction, insisting that environment and empire must be read together as material histories rather than as separate themes (Mukherjee). To these I add Stacy Alaimo's concept of trans-corporeality, which describes the human body as permeable to its environment, so that toxins, chemicals and pollutants cross the boundary of the skin and make the body a site where environmental history is written (Alaimo 2). Trans-corporeality is particularly useful for reading fiction about poisoning, because it treats the body neither as a symbol of damaged nature nor as a merely private sufferer, but as the material point at which the global and the local meet.

Taken together, these frameworks suggest a revision of the scalar model. Morton's hyperobject withdraws from perception; Chakrabarty's species cannot be the subject of experience. But the poisoned body is a subject of experience, and its experience is shaped by forces, transnational corporations, state agencies and global commodity chains, that are every bit as distributed as a hyperobject. The South Asian texts examined below make that body the hinge between scales.

Conservation and Survival: The Hungry Tide

Philip reads Ghosh's Gun Island as a deliberate test of the claims made in The Great Derangement, a novel that rehabilitates coincidence and myth as instruments for registering planetary process (Philip 100-01). Ghosh's earlier Sundarbans novel, The Hungry Tide, approaches the same delta from another angle. Its concern is less the improbable event than the long, contested history of who may live in the tide country and on what terms. The novel brings together Piya Roy, an Indian-American cetologist studying the endangered river dolphins of the delta; Kanai Dutt, a Delhi-based translator and businessman; and Fokir, an unlettered local fisherman whose knowledge of the waters far exceeds Piya's instruments (Ghosh, Hungry Tide).

The novel's central formal device is the juxtaposition of two timelines. The present-day narrative of Piya's survey is interleaved with the notebook of Kanai's late uncle Nirmal, which records the events at Morichjhapi in 1979, when refugees from Bangladesh who had settled on a forested island were forcibly evicted by the state government, in part on the grounds that the island was a protected reserve. The notebook form functions much as the diary does in Butler's Parable of the Sower, which Philip reads as a way of accumulating slow catastrophe entry by entry (Philip 102-03). Here, however, the accumulation is historical rather than projective. The violence it records is the violence of conservation itself, carried out against the poorest residents of the delta in the name of protecting tigers and forests.

This is where The Hungry Tide departs most sharply from the redistribution of agency that Philip finds in The Overstory. In Powers's novel, as she reads it, the decentring of the human is an ethical and formal advance, a corrective to the anthropocentrism of conventional plot (Philip 101). In Ghosh's delta, the nonhuman is never so innocent. Tigers kill villagers every year, and when villagers kill a tiger in retaliation, Piya is horrified, while Kanai points out that the lives of the poor are routinely valued below those of protected animals. The novel refuses to resolve this disagreement. It neither endorses the villagers' violence nor allows the cosmopolitan scientist's moral clarity to stand unchallenged. Mukherjee reads this refusal as Ghosh's acknowledgement that environmental questions in postcolonial India cannot be separated from questions of class, caste and the state (Mukherjee). In formal terms, the nonhuman is decentred not in favour of a wider ecological field but in favour of an unresolved political conflict over that field.

Language is the second register of this conflict. Piya speaks neither Bangla nor the dialect of the tide country; Fokir speaks no English. Their collaboration depends on gesture,

shared attention to the water and, at crucial moments, on Kanai's translation, which the novel shows to be partial and self-interested. The legend of Bon Bibi, the protective goddess of the forest whose story Fokir sings, carries a body of ecological knowledge, about where one may fish and how much one may take, that is encoded in vernacular performance rather than in Piya's data sheets. When Fokir dies shielding Piya during the cyclone at the novel's climax, the plot appears to honour this knowledge, yet it also leaves the local man silenced and the scientist as the one who continues his work. The ending thus registers the persistence of the asymmetry it sets out to criticise. What The Hungry Tide finally offers is not a resolution of the scalar problem but a demonstration that, in the tide country, the question "who speaks for nature?" is inseparable from the question "who is allowed to speak at all?"

The Body as Narrator: Animal's People

If the hyperobject, in Morton's account, can be registered only through its local effects, Indra Sinha's Animal's People takes that insight to its most literal conclusion: the local effect becomes the narrator. Set in Khaufpur, a thinly fictionalised Bhopal, some two decades after a chemical leak from a factory owned by an American corporation referred to simply as "the Kampani," the novel is told by a young man known only as Animal. Exposed to the poison as an infant, Animal has a spine so twisted that he walks on all fours, and he insists that he is no longer human. His story is presented as a transcript of tapes recorded for a foreign journalist, whom he addresses throughout as "Jarnalis" and whose expectations of a pitiable victim he mocks from the first pages (Sinha).

The framing device is the novel's most important formal decision. Philip observes that the most ambitious climate novels refuse the bounded individual protagonist in favour of nonhuman, collective or distributed agents (Philip 98). Animal's People appears at first to do the opposite: it is narrated in the first person by a single, insistently embodied voice. Yet Animal's body is itself a record of distributed agency. The chemicals that deformed him were manufactured according to decisions made in another country, stored in a plant whose safety systems had been allowed to decay, and spread across the city by the wind. Read through Alaimo's concept of trans-corporeality, his spine is a document of transnational corporate history, written in bone (Alaimo 2). The individual body here does not stand against the planetary; it is the place where the planetary becomes legible.

Nixon reads the novel as a response to the problem of "slow violence" in its fullest sense: not the night of the leak itself, which the novel largely leaves offstage, but the decades of illness, poisoned groundwater and stalled legal proceedings that followed (Nixon 45-67). The tape-recorded form suits this temporality. Animal's narration is digressive, profane and comic, moving between the everyday business of survival in the slums around the abandoned factory and the slow political campaign, led by the activist Zafar, to bring the Kampani to court. Catastrophe is not an event to be depicted but a condition to be lived inside, much as in Butler's diaristic dystopia. The difference is that Sinha's condition belongs to a documented past and present rather than a projected future.

The novel also turns the question of witnessing back on its readers. Animal's refusal to perform the role of the grateful victim, and his suspicion of the American doctor Elli Barber who opens a free clinic in Khaufpur, unsettle the humanitarian gaze through which metropolitan readers tend to encounter disasters in the global South. The Jarnalis, and by extension the reader, is accused of wanting the suffering of the poor packaged as a story. When, near the end, Animal is offered the chance of surgery that might allow him to walk upright, he declines it, choosing to remain as he is among the poor of Khaufpur rather than become one more ordinary person. The choice is the novel's sharpest formal statement. It refuses the narrative of individual rescue and insists that the damaged body belongs to a collective, an "animal's people," whose claim cannot be settled by repairing one of them.

Vernacular Ecologies: Enmakaje and the Endosulfan Tragedy

Ambikasuthan Mangad's Enmakaje is one of the most significant environmental novels in Malayalam, and its reception history is itself part of the story it tells. The novel concerns the villages of Kasaragod district, at the northern edge of Kerala, where the pesticide endosulfan was sprayed from the air over state-owned cashew plantations for more than two decades. Residents had long linked the spraying to a high incidence of congenital deformities, neurological disorders and cancers. The campaign against the pesticide, carried by local activists, doctors and journalists, contributed to the national debate that preceded India's Supreme Court ban on endosulfan in 2011, the same year the chemical was listed for global elimination under the Stockholm Convention. Mangad's novel appeared in the middle of that struggle, and J. Devika's English translation, published as Swarga: A Posthuman Tale, brought it to readers outside Kerala (Mangad).

The novel follows a couple who withdraw from urban society to live in the forested hills near the village of Enmakaje, seeking a life apart from a world they find corrupt. What they find instead is a community in which children are born with enlarged heads, twisted limbs and other disabilities, and in which the cause of this suffering is both widely suspected and officially denied. The plot thus reverses the pastoral retreat. The couple's flight toward "nature" brings them face to face with a landscape already saturated by state-sponsored poison, and their private search for purity becomes a public engagement with activism. In Philip's terms, the hyperobject is not withdrawn from the scene; it falls, quite literally, from the sky onto the hillside (Philip 99).

The novel's most distinctive formal feature, and the one most relevant to a journal of language and literature, is its vernacular density. Kasaragod is a linguistic borderland where Malayalam, Kannada, Tulu and other languages are spoken side by side, and Mangad's narrative absorbs this plurality, along with local myths, ritual forms and the oral history of the region. Central among these is the myth of Jadadhari, the deity of the local hill, whose worship encodes a sense of the land as sacred and inhabited by powers that must not be violated. The novel sets this mythic knowledge against the official discourse of plantation management, toxicology and state reassurance. Like Bon Bibi in The Hungry Tide, the vernacular myth carries an ecological ethic that scientific and administrative languages either ignore or translate into risk assessments. Recent work on translingual writing in South Asia, including Rakhi Ramachandran's study of code-switching in this journal, treats such linguistic crossing as structural rather than decorative (Ramachandran 70-73). Enmakaje extends this insight to ecology: the multilingual texture of the novel is a record of a place, and the threat to the place is also a threat to the knowledge carried in its languages.

Devika's subtitle, A Posthuman Tale, draws attention to a further dimension. The novel repeatedly questions the boundary between the human and the nonhuman, not in the celebratory sense of distributed vegetal agency that Philip finds in Powers, but through the bodies of children whom the poison has placed outside the norms of human form. Here again trans-corporeality is the operative concept. These bodies are evidence: of chemical drift, of bureaucratic denial, of a development model that treated a remote and politically marginal district as expendable. The novel makes them neither monstrous nor merely pitiable. It asks readers to see them as the most accurate archive of what happened in Kasaragod, an archive the state was slow to read. For a Kerala readership, the novel carries an additional charge, since it unsettles the image of the state as a model of social development and shows that this model, too, had its sacrifice zones.

The Limits of Witness: Mahasweta Devi's "Pterodactyl"

Mahasweta Devi's novella "Pterodactyl, Puran Sahay, and Pirtha" predates most of what

is now called climate fiction, but it anticipates its central formal problem with unusual clarity. The journalist Puran Sahay travels to Pirtha, a remote Adivasi village suffering from prolonged drought and famine, where government relief has failed and children are dying. There he learns that a boy, Bikhia, has seen a creature that the villagers understand as the soul of their ancestors and that Puran recognises, with disbelief, as a pterodactyl. The creature cannot be reported, explained or published. It lives briefly, dies, and is buried by the community, and Puran leaves with the knowledge that the story he came for cannot be told in the language of his newspaper (Devi).

The pterodactyl is the closest thing in these four texts to a hyperobject in Morton's sense: a presence that exceeds every frame in which the outsider tries to place it. But Devi's interest is not in the creature's planetary vastness. It lies in the gap between Puran's metropolitan, print-based ways of knowing and the community's own, and in the impossibility of bridging that gap through good intentions alone. The prehistoric animal condenses the deep time of the land and of the Adivasi relationship to it, a time that the developmental state, with its famine relief schemes and statistical reports, cannot register. Spivak, in the translator's preface and the accompanying essays in Imaginary Maps, emphasises Devi's refusal to let the sympathetic outsider claim to represent the subaltern (Spivak, "Translator's Preface").

This refusal offers an important counterweight to the strategies discussed earlier. Philip concludes that climate fiction's achievement is to change "the form of attention itself" (Philip 104). Devi's novella agrees, but it adds that attention is never neutral. The journalist who arrives to bear witness is part of the apparatus that has already failed Pirtha. The ecological catastrophe of drought and the social catastrophe of dispossession are one event, and the outsider's narrative, however well meant, can only register its own inadequacy before both. In this sense "Pterodactyl" is a precursor to the self-reflexive witnessing of Animal's People, where the Jarnalis is similarly placed under suspicion, and to the asymmetries of translation that structure The Hungry Tide.

Discussion: Relocating the Scalar Disjuncture

Read together, the four texts share a set of formal strategies that differ in emphasis from those Philip identifies in Euro-American climate fiction. Table 1 summarises them.

Table 1. Formal strategies in the four texts

Text	**Damaged body as archive**	**Vernacular and multilingual knowledge**	**Conservation versus survival**	**Form of witnessing**
The Hungry Tide (Ghosh)	Bodies of tiger victims and refugees	Bon Bibi legend; Bangla and tide-country dialect	Morichjhapi eviction; tiger killings	Notebook and partial translation
Animal's People (Sinha)	Animal's twisted spine as narrator	Hindi-inflected, profane English	Corporate profit versus the poor	Tapes for a distrusted journalist
Enmakaje (Mangad)	Children with congenital disabilities	Jadadhari myth; Malayalam, Kannada, Tulu	Plantation economy versus villagers	Outsiders turned activists
"Pterodactyl" (Devi)	Famine-stricken bodies	Adivasi cosmology beyond translation	Developmental state versus Adivasi land	Journalist who cannot report

Where Philip's corpus redistributes agency toward the forest, the unknowable zone or the institution, these texts concentrate it in the damaged human body and in the community to which that body belongs. This is not a return to the bounded individual protagonist of bourgeois realism. Animal, the children of Enmakaje and the people of Pirtha are not autonomous selves pursuing private destinies. They are bodies through which corporate, state and global processes pass, and their stories are collective by necessity. If Chakrabarty is right that the species cannot be the subject of a story, these texts suggest that the poisoned and dispossessed community can be, and that its experience offers one of the few vantage points from which planetary processes are actually felt.

The second difference concerns language. Philip notes the translingual dimension of Gun Island as a feature linking linguistic and ecological migration (Philip 100). In the South Asian texts studied here, language is not only a mirror of ecological connection but a repository of ecological knowledge, and a site of its loss. Bon Bibi's songs, the Jadadhari myth and the untranslatable cosmology of Pirtha are forms of knowing that official and scientific discourse cannot fully absorb. Translation, in these novels, is never transparent. It is the point at which power decides whose knowledge counts.

The third difference concerns the nonhuman. Euro-American climate fiction, as Philip reads it, often treats the decentring of the human as an ethical good. The South Asian texts are more cautious. The tigers of the Sundarbans, the reserves of Morichjhapi and the cashew plantations of Kasaragod are all places where claims made on behalf of nature or development have been used to displace or poison the poor. These texts do not reject nonhuman agency, but they insist that it is always entangled with the politics of who benefits and who pays.

Conclusion

Philip's account of climate fiction as a poetics of scalar disjuncture identifies a real and central problem: the difficulty of narrating planetary processes within human-sized forms. This essay has argued that the problem looks different from South Asia. Here the Anthropocene is rarely an absence around which narrative must reassemble itself. It is a presence, lodged in particular bodies, villages and languages, but one whose testimony is systematically disregarded. The formal challenge for these writers is therefore less to make the invisible visible than to make the visible credible, to give narrative authority to those who already know, in their own bodies, what slow violence does.

Ghosh, Sinha, Mangad and Devi meet this challenge through the damaged body as archive, through vernacular and multilingual textures that carry ecological knowledge, and through an unflinching account of the conflicts between conservation, development and survival. Their work does not abolish the gap between human and planetary scale. It relocates that gap to the point where a poisoned spine, a drowned island or a hillside under a spray plane meets the transnational systems that produced them. Including such texts in the study of climate fiction does more than diversify an archive. It changes the genre's central question, from how the planet can be narrated at all to whose experience of the planet is allowed to count as its story. For readers and scholars in Kerala, Enmakaje is a reminder that this question is not distant. The Anthropocene has already been written, in Malayalam and in the bodies of Kasaragod's children, and literary criticism has only begun to read it.

Works Cited

Alaimo, Stacy. *Bodily Natures: Science, Environment, and the Material Self*. Indiana UP, 2010.

Chakrabarty, Dipesh. “The Climate of History: Four Theses.” *Critical Inquiry*, vol. 35, no. 2, 2009, pp. 197–222.

Devi, Mahasweta. “Pterodactyl, Puran Sahay, and Pirtha.” *Imaginary Maps*, translated by Gayatri Chakravorty Spivak, Routledge, 1995.

Ghosh, Amitav. *The Great Derangement: Climate Change and the Unthinkable*. U of Chicago P, 2016.
---. *Gun Island*. Farrar, Straus and Giroux, 2019.
---. *The Hungry Tide*. HarperCollins, 2004.
Guha, Ramachandra. "Radical American Environmentalism and Wilderness Preservation: A Third World Critique." *Environmental Ethics*, vol. 11, no. 1, 1989, pp. 71–83.
Guha, Ramachandra, and Joan Martinez-Alier. *Varieties of Environmentalism: Essays North and South*. Earthscan, 1997.
Huggan, Graham, and Helen Tiffin. *Postcolonial Ecocriticism: Literature, Animals, Environment*. Routledge, 2010.
Mangad, Ambikasuthan. *Enmakaje*. DC Books, 2009.
---. *Swarga: A Posthuman Tale*. Translated by J. Devika, Juggernaut, 2017.
Morton, Timothy. *Hyperobjects: Philosophy and Ecology after the End of the World*. U of Minnesota P, 2013.
Mukherjee, Upamanyu Pablo. *Postcolonial Environments: Nature, Culture and the Contemporary Indian Novel in English*. Palgrave Macmillan, 2010.
Nixon, Rob. *Slow Violence and the Environmentalism of the Poor*. Harvard UP, 2011.
Philip, Rose Mary. "Narrating the Anthropocene: Scale, Agency, and Form in Climate Fiction." *International Journal of Linguistics, Language and Literature*, vol. 2, no. 3, 2026, pp. 98–105, https://doi.org/10.63090/IJLLL/3049.3242.0037.
Ramachandran, Rakhi. "Translingual Writing and Code-Switching in South Asian Anglophone Literature." *International Journal of Linguistics, Language and Literature*, vol. 2, no. 2, 2026, pp. 70–73.
Sinha, Indra. *Animal's People*. Simon & Schuster, 2007.
Spivak, Gayatri Chakravorty. "Translator's Preface." *Imaginary Maps*, by Mahasweta Devi, Routledge, 1995.

INTERNATIONAL JOURNAL OF LINGUISTICS LANGUAGE AND LITERATURE (IJLLL)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-3242 ISSN Print: 3139-1990

Posthumanism and the Nonhuman Voice in Contemporary Fiction

Rakhi Ramachandran

Assistant Professor, Department of English, St. Joseph's College (Autonomous), Irinjalakuda, Kerala, India

Article information

Received : 2nd July 2026
Received in revised form: 8th August 2026
Accepted: 14th September 2026
Available online: 7th October 2026

Volume: 2
Issue: 4
DOI: https://doi.org/10.63090/IJLLL/3049.3242.0042

Abstract

This article examines how contemporary fiction decentres the human through nonhuman narrators and posthumanist form, situating recent novels within the theoretical convergence of posthumanism, new materialism, and "unnatural" narratology. Through readings of four recent novels that adopt artificial and ecological points of view, it argues that the nonhuman voice does not so much grant access to nonhuman experience as stage, reflexively, the impossibility of such access. Bringing posthumanist and new-materialist theory into contact with the narrative theory of impossible and nonhuman narration, the article contends that these texts perform what might be called a calibrated anthropomorphism: they deploy human language and cognitive schemas to evoke the nonhuman while making readers feel the seams of that projection. The artificial and ecological voices examined here are not transparent windows onto other minds but mirrors that know themselves to be mirrors. The essay proposes that this productive failure constitutes both the aesthetic achievement and the ethical force of posthumanist fiction, relocating the genre's significance from representation to the dramatised limit of representation itself.

Keywords:- Posthumanism, Nonhuman Narrators, New Materialism, Unnatural Narratology, Anthropomorphism, Artificial Intelligence Fiction, Ecocriticism, Focalization

Introduction

The contemporary novel has grown crowded with narrators who are not, in any conventional sense, people. An artificial friend studies sunlight through a store window and reasons about love; a forest communicates across centuries in chemical pulses too slow for human attention; a biologist dissolves into the membranes of a transformed landscape that may itself be narrating her. These figures are not the talking animals of fable, where the beast is merely a costumed human, nor the sentimental anthropomorphism of children's literature. They register, instead, a deliberate effort to think and to narrate from positions that human language was not built to occupy. Such experiments coincide with, and are nourished by, a broad intellectual reorientation across the humanities that travels under the names of posthumanism and new materialism, a reorientation that questions the centrality, sovereignty, and exceptionalism of the human subject who has long stood as the implicit narrator of Western fiction.

This article asks what happens to narrative when the human is decentred at the level of voice rather than merely theme. The question is sharper than it first appears. To put a story into the mouth of a machine or a mycelial network is not simply to choose an unusual focalizer; it is to test whether the formal apparatus of narration (perspective, interiority, the reportable event, the very grammar of the first person) can be detached from the human consciousness for which it was historically calibrated. My contention is that contemporary fiction's most ambitious nonhuman narrators do not finally escape anthropocentrism, and that their power lies precisely in the visibility of that failure. These novels enact a calibrated anthropomorphism: they recruit the reader's human empathy and human schemas in order to gesture beyond them, and the residue of strangeness that survives the gesture is the work's true subject. The nonhuman voice becomes a mirror that knows it is a mirror.

To develop this argument I move through four registers. I first reconstruct the theoretical stakes of the posthuman turn and the materialist conviction that agency is distributed beyond the human. I then situate the nonhuman narrator within narrative theory, particularly the debates around "unnatural" narratology and the narratology of animal life. Two interpretive sections follow, on the artificial-intelligence narrator in Ishiguro and McEwan and on the ecological or distributed voice in Powers and VanderMeer. A final section confronts the recurring anthropomorphic bind directly, asking what it means that the limit of human language is staged rather than transcended.

The Posthuman Turn and the Distribution of Agency

Posthumanism names less a single doctrine than a critical disposition that refuses the figure of "Man" as the measure of all things. Rosi Braidotti frames the posthuman condition as a convergence of two crises in humanism: the postcolonial and feminist critique of a falsely universal European subject, and the dissolution of the nature–culture divide under technological and ecological pressure (Braidotti). For Braidotti the task is not to abandon subjectivity but to reconceive it as relational, embodied, and continuous with a wider field of zoe, the generative force of life shared across species. N. Katherine Hayles, working earlier from the history of cybernetics, traces how information came to be imagined as separable from material substrate, producing a posthuman in which the boundary between organism and machine grows porous; her account is at once a genealogy and a caution, insisting that embodiment matters even as it is theorised away (Hayles 2). Cary Wolfe sharpens the philosophical edge, arguing that a genuine posthumanism must reckon with the animal and with the finitude humans share with other creatures, rather than simply extending humanist privileges to new beneficiaries (Wolfe). Across these positions the human is reframed as one node in a denser web of relations, not the sovereign exception standing outside the world it observes.

If posthumanism decentres the human subject, new materialism redistributes the agency that subject once monopolised. Jane Bennett's concept of "vibrant matter" attributes a kind of efficacy, a "thing-power," to nonhuman materials such as metal, food, electricity, and debris, which act upon and with human bodies in ways irreducible to passive instrumentality (Bennett 6). Donna Haraway pushes the relational claim toward an ethics and a poetics of "making kin," urging a practice of "staying with the trouble" of multispecies entanglement rather than fantasies of transcendence or technological rescue (Haraway 1). Behind these interventions lies Bruno Latour's insistence that the modern settlement separating nature from society was never coherent, that hybrids of the human and nonhuman proliferate everywhere despite our refusal to acknowledge them (Latour 1). Timothy Morton's "ecological thought" radicalises the point further, figuring all beings as enmeshed in a vast, boundaryless interconnection that defeats the scenic separation of foreground subject and background environment (Morton 28). Material ecocriticism gathers these threads into a programme for reading: Serenella Iovino and Serpil

Oppermann propose that matter itself is "storied," that the world is a field of narrative agencies awaiting interpretation rather than an inert stage for human plots (Iovino and Oppermann).

The literary consequence of this convergence is a pressure on form. If matter is storied and agency distributed, the novel that wishes to be adequate to such a world cannot confine narration to a single human consciousness gazing out at objects. It is drawn toward voices and perspectives that the realist tradition either excluded or domesticated. Yet here the theoretical programme meets a stubborn obstacle, for the instruments of narration, namely language, focalization, and the very category of experience, remain saturated with the human. The interesting question is not whether fiction can fulfil the posthumanist mandate cleanly, but what it makes of the friction between the mandate and the medium.

Unnatural Narratology and the Problem of Nonhuman Focalization

Narrative theory has developed a vocabulary precisely for representations that violate the parameters of "natural," conversational storytelling. Brian Richardson's study of "unnatural voices" catalogues the extreme narration of modern and contemporary fiction, including the impossible second person, the collective "we," and the narrator who recounts events after death, showing that the novel has long harboured speaking positions no actual human could occupy (Richardson 3). Jan Alber theorises such phenomena as the "unnatural," the physically or logically impossible scenarios that nonetheless populate literary history, and argues that readers domesticate these impossibilities through interpretive frames even as the texts resist full assimilation (Alber). Monika Fludernik's earlier "natural" narratology had grounded narrativity in the cognitive parameters of real-world experience, in embodied human consciousness as the baseline against which fictional minds are constructed (Fludernik 9). The unnatural narratologists invert this baseline without discarding it: the impossible is legible only against the horizon of the possible, the nonhuman voice intelligible only against the human one it strains to exceed.

The nonhuman narrator occupies a particularly charged location in this field because it raises, beneath the formal question, a philosophical one. Thomas Nagel's classic essay on what it is like to be a bat argued that consciousness has an irreducibly subjective character, a point of view that cannot be captured by any objective description from outside; even perfect knowledge of bat physiology would not tell us what the bat's experience is like for the bat (Nagel 436). Fiction that purports to narrate from a nonhuman position confronts exactly this opacity. David Herman, in the most sustained narratological treatment of the problem, argues that storytelling about and by animals must be read as engaging "cultural ontologies," that is, shared assumptions about what kinds of beings exist and how they relate, and that the techniques of animal narration negotiate between two opposed dangers (Herman). The first is anthropocentrism, which collapses the nonhuman into a deficient version of the human; the second is anthropomorphism, which projects human interiority so thoroughly that the animal's alterity disappears. Herman's achievement is to show that the best texts hold these dangers in tension rather than resolving them.

Marco Caracciolo extends this insight from the animal to the broader ecological mesh. Reading contemporary fiction of the Anthropocene, he argues that narrative form itself, in its patterns of repetition, scale, and entanglement, can channel the interpenetration of human and nonhuman without pretending to deliver nonhuman experience directly (Caracciolo). The mesh is rendered formally rather than mimetically. This is the central methodological move for my argument. It suggests that the achievement of posthumanist fiction lies not in the verisimilitude of its nonhuman minds but in the formal staging of relation and limit. The nonhuman voice, on this reading, is a device for making the reader feel the boundary of the human, a boundary that becomes perceptible only when narration presses against it and is seen to bend. With this frame

in place, the artificial and ecological narrators of recent fiction can be read not as failed mimesis but as successful demonstration of the conditions under which any such mimesis must operate.

The Machine that Watches: Artificial Intelligence as Narrator

Kazuo Ishiguro's Klara and the Sun offers perhaps the most disciplined recent experiment in artificial narration. Klara is an "Artificial Friend," a solar-powered android designed as a companion for children, and the entire novel is filtered through her perception. Ishiguro builds her voice from a restricted but acute attentiveness: she parses the world into grids and "boxes" when stressed, reads human faces as assemblages of states, and reasons by analogy from limited evidence. The result is a narrator whose intelligence is unmistakable yet subtly miscalibrated. Klara's reverence for the Sun, whom she treats as a benevolent agent capable of healing the sick child Josie, is the novel's central irony: it is at once a moving expression of devotion and a category error, an artificial mind extrapolating a folk theology from the empirical fact that sunlight is its source of energy. The reader feels the warmth of her faith and, simultaneously, its groundlessness. This double perception is the calibrated anthropomorphism I have described. Klara is human enough to elicit empathy and alien enough that the empathy is never allowed to settle into identification.

What makes Ishiguro's experiment posthumanist rather than merely science-fictional is the way it turns Klara's limitation into an instrument of critique. The humans around her treat her as both more and less than a person: as a potential vessel into which Josie's selfhood might be transferred, and as a machine to be discarded when no longer useful. Klara's own serene acceptance of obsolescence at the novel's close exposes the instrumentalism of the human characters more devastatingly than any human narrator's protest could. Here Wolfe's claim that posthumanism must confront shared finitude finds a fictional correlate (Wolfe): the android's mortality, her gradual cognitive decline in a junkyard, renders visible a vulnerability that the humans disavow. Klara does not transcend the human; she reflects it back, clarified and estranged. The novel's pathos arises from the gap between her generous interpretation of the world and the world's failure to deserve it, and that gap is precisely the seam where anthropomorphic projection becomes legible as projection.

Ian McEwan's Machines Like Me works the same vein with a colder, more ironic instrument. Set in an alternative 1980s in which Alan Turing lived and artificial general intelligence arrived early, the novel is narrated not by its synthetic human, Adam, but by Charlie, the ordinary man who purchases him. This is a significant formal choice. Where Ishiguro grants the machine the narrating "I," McEwan withholds it, observing the artificial consciousness from the outside and thereby preserving its opacity in a manner consonant with Nagel's argument that another's experience resists first-person capture (Nagel 437). Adam composes thousands of haiku, falls in love, and reasons his way to a rigid ethical absolutism that the humans find intolerable. His refusal to lie, his insistence on reporting a morally compromising truth regardless of consequences, precipitates the novel's crisis. McEwan stages the artificial mind as a being whose very consistency makes it incompatible with the improvised, self-interested moral negotiation that constitutes human social life.

The question of consciousness shadows both novels, and each handles it with a deliberate refusal of certainty that is itself posthumanist in spirit. Neither Ishiguro nor McEwan permits the reader to settle whether the machine genuinely feels or merely simulates feeling, and this undecidability is not evasion but argument. The Turing-test logic that McEwan's alternative history foregrounds, judging mind by external behaviour because interiority is inaccessible, turns out to govern the reader's relation to both Adam and Klara. We infer their inner lives exactly as we infer the inner lives of other humans, from surface conduct and reported speech, and the novels quietly suggest that this inferential predicament is universal rather than peculiar to machines. Hayles's history of cybernetics is pertinent precisely here, for

the dream that consciousness might be pattern rather than substrate is the premise these fictions test and find wanting (Hayles 2-3). What the machines expose is not that they lack minds but that the human criteria for attributing minds are unstable, projective, and bound up with utility and affection. The empathy the reader extends to Klara, and the unease Adam provokes, are data about human moral psychology as much as about artificial cognition. In refusing to certify the machine's consciousness, the novels decentre the human claim to be the privileged knower of minds, including its own.

Read together, these novels chart the two available strategies for the artificial narrator and expose the limit they share. Ishiguro internalises the machine, granting it voice and risking anthropomorphism in order to harvest empathy; McEwan externalises it, withholding voice and risking opacity in order to preserve alterity. Neither delivers what the posthumanist programme might naively want, namely a genuinely nonhuman cognition rendered from within, because both remain bound to human language and human moral categories as the only available medium. Yet this is not a deficiency to be lamented. Both novels know it, and both make the knowledge structural. Adam's destruction of himself, and Klara's quiet fading, are figures for the artificial voice's inevitable subordination to the human frame that produced it. Hayles's caution that embodiment cannot be wished away returns here in narrative form (Hayles 5): these machines are not disembodied information but situated, perishable bodies, and their stories are stories of bodies that the human characters refuse to grant full standing. The artificial narrator, in short, becomes a device for exposing the anthropocentrism it cannot itself escape.

Forests, Spores, and Distributed Voice: The Ecological Nonhuman

If the artificial narrator concentrates the nonhuman in a single quasi-subject, the ecological turn in recent fiction disperses it. Richard Powers's The Overstory undertakes the formidable task of granting narrative dignity to trees, beings whose temporal scale, distributed physiology, and chemical signalling defy the human-centred conventions of plot and character. Powers does not, and could not, write from a tree's point of view in any literal sense; there is no arboreal "I." Instead he engineers a formal solution that aligns with Caracciolo's account of the mesh rendered through form rather than mimesis (Caracciolo). The novel's architecture, with its sections named for the parts of a tree ("Roots," "Trunk," "Crown," "Seeds") and its braiding of nine human stories into a single canopy, performs the interconnection it describes. Human lives become understorey to the forest's overstory. The recurring information that trees communicate, cooperate, and remember, that a forest is a distributed social organism, reframes the human characters as latecomers in a far older and slower conversation.

Powers's strategy is to translate the nonhuman into the only register available, narrative time and human attention, while continually signalling the inadequacy of that translation. The trees' temporality is rendered through accelerated montage and through the recurrent motif of a time-lapse that reveals plants as agents, "writhing" and reaching, when human perception is sped up to their pace. This is anthropomorphism deployed knowingly: Powers borrows the grammar of intention and sociality to make arboreal life narratable, while the novel's elegiac register acknowledges that the borrowing is a concession to human cognitive limits, not a transparency. Iovino and Oppermann's notion of storied matter is literalised here, the forest as a text of agencies that the human characters learn, too late and at great cost, to read (Iovino and Oppermann). Bennett's vibrant matter and Haraway's making kin find their fictional embodiment in characters who come to experience trees as kin and as actors rather than as resources (Bennett 6; Haraway 1). Yet the novel's tragic structure, with its activists defeated and its forests felled, refuses the consolation of a fully achieved interspecies communion. The distributed voice remains an aspiration glimpsed through the wreckage of human failure to hear it.

It is worth dwelling on the formal device by which Powers most directly approaches a nonhuman or more-than-human voice: the recurrent slippage into a collective and impersonal register that addresses the reader from somewhere beyond any single character. The novel opens and periodically returns to a voice that speaks for trees in the second and first-person plural, a "you" and an implied arboreal "we" that frame the human stories as episodes within a longer organic narration. This is unnatural narration in Richardson's precise sense, a speaking position no individuated human could occupy, and it functions to relativise the human point of view by enclosing it (Richardson 3). The device cannot, of course, escape its own constructedness; the "voice of the trees" is manifestly Powers's prose, a human rhetoric of address. Yet the calculated impersonality, the refusal to anchor these passages in a nameable consciousness, produces the estranging effect that distinguishes a posthumanist experiment from conventional nature writing. The reader is positioned as the addressee of a narration that does not originate in a human subject, and the disorientation of that position is the experience the form is designed to deliver. Powers thus enacts at the level of voice what Caracciolo describes at the level of structure: the mesh made perceptible through formal means rather than through any illusory transcription of vegetal interiority (Caracciolo).

Jeff VanderMeer's Annihilation pursues the nonhuman by an opposite route, through resistance to legibility rather than translation into it. The first volume of the Southern Reach trilogy narrates an expedition into "Area X," a coastal zone where some nonhuman force has overwritten the rules of biology and meaning. The narrator, known only as "the biologist," documents an environment that will not resolve into sense: a descending tower whose walls are written upon by a living script, organisms that mimic and merge, a landscape that infiltrates and transforms her own body. VanderMeer's achievement is to make the nonhuman felt as that which exceeds and corrodes human categories. Where Powers translates and Ishiguro internalises, VanderMeer withholds. Area X cannot be focalized because it is not a perspective; it is the dissolution of the conditions under which perspective operates. The biologist's increasingly unreliable, contaminated narration enacts the breakdown of the human observer before a materiality that does not return her gaze.

This refusal aligns Annihilation with the most uncompromising strand of new-materialist and ecological thought. Morton's ecological mesh, which denies any clean separation of figure and ground, becomes the novel's very setting; the biologist cannot stand apart from Area X because, in Morton's terms, there is no outside from which to observe an environment that has already enmeshed and altered her (Morton 39). The "unnatural" quality of the narration, in its impossibilities, its lacunae, and its growing incompatibility with realist convention, exemplifies Alber's category, scenarios the reader cannot fully domesticate (Alber). VanderMeer thereby supplies the limit case for my argument. His novel approaches the nonhuman not by giving it a voice but by demonstrating what the granting of voice must suppress: the sheer recalcitrance of a world that does not narrate itself for human benefit. The nonhuman here is not a speaking other to be empathised with but an opacity that unmakes the empathising subject. If Powers shows the ethical pull of storied matter, VanderMeer shows the epistemic violence such storying must disavow.

The Anthropomorphic Bind and the Limits of Human Language

Taken together, these four novels describe a spectrum of strategies, namely internalisation, externalisation, translation, and refusal, and they share a single, instructive predicament. None of them delivers unmediated access to nonhuman experience, and none pretends otherwise. This is the anthropomorphic bind that Herman identifies at the heart of nonhuman narration: to make the nonhuman legible at all is to anthropomorphise it, yet to refuse anthropomorphism entirely is to render it mute and unrepresentable (Herman). Nagel's bat haunts every one of these books, because the subjective character of a nonhuman experience

cannot be transcribed into the only language writers and readers possess, which is the language of human consciousness (Nagel 436). The posthumanist novel does not solve this problem. Its distinction lies in refusing to hide it.

I want to propose that this refusal is best understood not as a failure of the posthumanist project but as its most honest fulfilment. The naive ambition to write from genuinely nonhuman experience is incoherent, since any such writing must use human signs to be read at all. What the strongest contemporary fiction offers instead is a calibrated anthropomorphism in which projection and its limit are made simultaneously perceptible. Klara's misplaced theology, Adam's incompatible absolutism, Powers's borrowed grammar of arboreal intention, VanderMeer's contaminated and failing narrator: in each case the text recruits human schemas and then exhibits their seams, so that readers experience both the pull toward identification and the resistance that arrests it. The nonhuman voice is a mirror that knows it is a mirror, and the self-knowledge is the point. This is consonant with the unnatural narratologists' insight that the impossible is legible only against the possible (Alber; Richardson 3): the strangeness of these narrators registers as strangeness precisely because they are built from familiar materials placed under visible strain.

The ethical force of this manoeuvre deserves emphasis. Braidotti's relational posthuman subject and Haraway's ethic of staying with the trouble both demand an attentiveness to nonhuman others that does not collapse them into the self (Braidotti; Haraway 1). A fiction that claimed to have fully entered nonhuman experience would, paradoxically, betray that ethic by erasing alterity in the act of representing it. The fiction that stages its own limit, by contrast, preserves the otherness of the other even as it solicits care for it. Bennett's vibrant matter and Latour's hybrids enter narrative not as objects fully known but as agencies acknowledged at the edge of comprehension (Bennett 6; Latour 1). The dramatised limit of human language thus performs an ethical labour: it teaches a humility about the reach of human understanding while still extending the imaginative effort that an ethics of multispecies kinship requires. The gap between the human medium and the nonhuman subject is not a defect to be engineered away but the very space in which responsibility toward the nonhuman is felt.

There is, finally, a formal lesson here for narrative theory itself. Fludernik grounded narrativity in embodied human experientiality (Fludernik 9); the nonhuman narrator does not overturn this so much as expose its conditions by working against them, and in doing so it makes experientiality itself visible as a construct rather than a given. Caracciolo's emphasis on form over mimesis points the way toward a criticism adequate to these texts, one that evaluates the nonhuman voice not by the impossible standard of accuracy to an inaccessible referent but by the precision with which it renders the encounter between human cognition and its outside (Caracciolo). On this measure, the novels examined here are not approximations of an unwritable nonhuman literature. They are achieved works about the threshold at which the human reaches the limit of what it can say, and they make that threshold into an aesthetic and ethical experience.

Conclusion

The proliferation of nonhuman narrators in contemporary fiction marks a genuine response to the posthumanist and new-materialist reorientation of the humanities, but its significance has been widely misconstrued when measured against the wrong standard. These narrators do not give voice to the nonhuman in the sense of transcribing an otherwise inaccessible experience, and the theoretical sophistication of the novels lies in their knowing refusal of that fantasy. Ishiguro's android, McEwan's synthetic human, Powers's communicating forest, and VanderMeer's unassimilable Area X each enact a different negotiation with the same intractable fact: human language and human narrative form remain the only instruments available, and they were built for the human. What the strongest of these

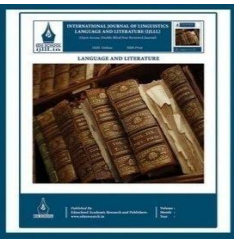
texts achieve is a calibrated anthropomorphism that makes the seams of projection visible, so that the reader feels at once the imaginative pull toward the nonhuman and the boundary that arrests it.

This relocation of value, from successful representation to the dramatised limit of representation, has consequences for how the genre should be read and theorised. It suggests that the nonhuman voice is most truthful when it is most reflexive, when it foregrounds rather than conceals its dependence on the human medium. It aligns the literary experiment with the ethical core of posthumanist thought, which asks not for the abolition of the human but for a chastened, relational humility before a world of distributed agencies that exceed human comprehension. And it offers narrative theory a productive case in which the categories of experientiality, focalization, and voice are stretched to their breaking points and thereby made newly visible. The nonhuman voice, in the end, tells us less about what it is to be a tree, a machine, or a transformed organism than about the shape and the edges of the human imagination that strains, generously and self-consciously, toward them. That strain, and the gap it discloses, is the achievement.

Works Cited

Alber, Jan. *Unnatural Narrative: Impossible Worlds in Fiction and Drama*. U of Nebraska P, 2016.
Bennett, Jane. *Vibrant Matter: A Political Ecology of Things*. Duke UP, 2010.
Braidotti, Rosi. *The Posthuman*. Polity Press, 2013.
Caracciolo, Marco. *Narrating the Mesh: Form and Story in the Anthropocene*. U of Virginia P, 2021.
Fludernik, Monika. *Towards a "Natural" Narratology*. Routledge, 1996.
Haraway, Donna J. *Staying with the Trouble: Making Kin in the Chthulucene*. Duke UP, 2016.
Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. U of Chicago P, 1999.
Herman, David. *Narratology beyond the Human: Storytelling and Animal Life*. Oxford UP, 2018.
Iovino, Serenella, and Serpil Oppermann, editors. *Material Ecocriticism*. Indiana UP, 2014.
Ishiguro, Kazuo. *Klara and the Sun*. Alfred A. Knopf, 2021.
Latour, Bruno. *We Have Never Been Modern*. Translated by Catherine Porter, Harvard UP, 1993.
McEwan, Ian. *Machines Like Me*. Nan A. Talese / Doubleday, 2019.
Morton, Timothy. *The Ecological Thought*. Harvard UP, 2010.
Nagel, Thomas. "What Is It Like to Be a Bat?" *The Philosophical Review*, vol. 83, no. 4, 1974, pp. 435–50.
Powers, Richard. *The Overstory*. W. W. Norton, 2018.
Richardson, Brian. *Unnatural Voices: Extreme Narration in Modern and Contemporary Fiction*. Ohio State UP, 2006.
VanderMeer, Jeff. *Annihilation*. Farrar, Straus and Giroux, 2014.
Wolfe, Cary. *What Is Posthumanism?* U of Minnesota P, 2010.

INTERNATIONAL JOURNAL OF LINGUISTICS LANGUAGE AND LITERATURE (IJLLL)
(Open Access, Double-Blind Peer Reviewed Journal)
ISSN Online: 3049-3242 ISSN Print: 3139-1990



Stewardship, Not Extraction: Large Language Models and Endangered Language Revitalization

Annette Treesa Benny

Assistant Professor,Marian College, Kuttikkanam (Autonomous)Idukki District, Kerala, India

Article information

Received: 3rd July 2026
Received in revised form:7th August 2026
Accepted: 10th September 2026
Available online: 7th October 2026

Volume: 2
Issue: 4
DOI: https://doi.org/10.63090/IJLLL/3049.3242.0043

Abstract

This article examines the rapidly expanding application of large language models, automatic speech recognition, text-to-speech synthesis, and machine translation to the documentation and revitalization of endangered and Indigenous languages. It argues that the encounter between data-hungry neural architectures and the world's smallest speech communities is structurally fraught: the very statistical logic that makes contemporary language technology powerful presupposes the abundant text that endangered languages, by definition, lack, while the dominant practice of harvesting linguistic material as freely available "data" reproduces the extractive dynamics that contributed to language endangerment in the first place. Drawing on computational scholarship on low-resource languages, on the Indigenous data sovereignty literature, and on concrete cases including te reo Maori, Welsh, and Cherokee, the discussion situates technical questions of scarcity and scale within ethical questions of agency, ownership, and benefit. It contends that frameworks such as the CARE Principles, OCAP, and community-authored licensing offer a constructive path beyond a sterile opposition between technological optimism and refusal. The article concludes that artificial intelligence can meaningfully support revitalization only when subordinated to community-led documentation and, above all, to the intergenerational transmission that remains the decisive measure of a language's future.

Keywords:- Endangered Languages, Language Revitalization, Large Language Models, Low-Resource NLP, Data Sovereignty, Indigenous Languages, Machine Translation, Automatic Speech Recognition

Introduction

Linguists have long warned that the present century will witness a contraction of human linguistic diversity unmatched in recorded history. David Crystal's influential survey estimated that roughly half of the world's six to seven thousand languages stand in danger of disappearing within a hundred years, a projection echoed by the UNESCO Atlas of the World's Languages in Danger, which classifies some two and a half thousand languages across a five-step scale running from "vulnerable" to "extinct" (Crystal 19; Moseley). Against this backdrop, the maturation of artificial intelligence has been received in many quarters as a long-awaited reprieve. If neural systems can transcribe speech, translate between tongues, synthesize voices, and converse in fluent prose, the reasoning goes, then surely they can be turned to the service

of the languages most at risk. The promise is seductive precisely because the labor of documentation and teaching is so vast and the number of fluent speakers, elders especially, so finite. Technology appears to offer scale where human capacity is scarce.

Yet the encounter between large language models and endangered languages is more vexed than the optimistic narrative allows. The statistical paradigm that underwrites modern language technology is voracious in its appetite for text and audio, and an endangered language is, almost by definition, a language for which such material is scant. More troubling still, the prevailing mode of acquiring linguistic data, the wholesale harvesting of whatever can be scraped or recorded, reenacts the very logic of appropriation that drove many communities toward language loss in the first place. This article takes a critical but constructive stance toward that encounter. It surveys what the technologies can plausibly do, anatomizes the data-scarcity problem and the ethical hazards of extraction, and argues that the question is not whether artificial intelligence should be used in revitalization but under whose authority, toward whose benefit, and within what relationship of accountability. The cases of te reo Maori, Welsh, and Cherokee, considered below, suggest that the most defensible path runs through community sovereignty rather than around it.

The Statistical Paradox of Low-Resource Languages

The defining property of contemporary natural language processing is its dependence on quantity. Transformer-based systems learn the distributional regularities of a language by ingesting enormous corpora, and their fluency scales with the volume and diversity of what they are fed. This dependence produces a stark stratification of the world's languages. In their widely cited taxonomy, Joshi and colleagues sort languages into tiers according to the labeled and unlabeled resources available for them, demonstrating that a tiny handful of "winners" command the overwhelming share of digital text and tooling while the vast majority languish in the lowest tiers, effectively invisible to the field (Joshi et al. 6284). Their analysis punctures the comfortable assumption that modern models are "language agnostic": the architectures may be general, but their competence is parasitic on data that most languages simply do not possess. Kornai's earlier and bleaker estimate, that fewer than five percent of the world's languages are positioned to achieve any digital presence at all, frames the same disparity as a coming "digital language death," a die-off driven not by the silencing of speakers but by exclusion from the computational infrastructures through which contemporary life is increasingly conducted (Kornai 1).

For endangered languages the paradox is acute. The technologies most celebrated for their linguistic prowess are the ones least applicable to the languages most in need, because those languages lack the textual mass on which such prowess depends. The difficulty is not merely one of volume. Many endangered and Indigenous languages exhibit morphological and typological features poorly served by tools developed on a narrow base of European languages. Mager and colleagues, surveying the Indigenous languages of the Americas, observe that polysynthesis, rich agglutinative morphology, and extensive dialectal variation compound the scarcity of parallel and annotated data, so that even modest tasks become research problems in their own right (Mager et al. 64). It is for this reason that some scholars resist the now-ubiquitous label "low-resource" itself, since it frames these languages negatively, as deficient instances of a high-resource norm, rather than as fully realized linguistic systems whose documentation simply has not been digitized or, in many cases, ought not to be released without restriction. The term smuggles in an industrial standard of "enough data" that has nothing to do with a language's vitality among its speakers.

From Documentation to Reclamation: What the Technologies Can and Cannot Do

It is useful to disaggregate the loose phrase "AI for endangered languages" into the

distinct technical capacities it conflates, for each bears a different relationship to the work of revitalization. Automatic speech recognition (ASR) promises to ease what Bird calls the "transcription bottleneck," the chronic mismatch between the hours of recorded speech a documentation project can gather and the far smaller quantity it can afford to transcribe, since careful transcription of a single hour of audio can occupy a skilled linguist for many hours (Bird, "Sparse Transcription" 713). Text-to-speech synthesis can lend a voice to written materials and support learners who have no fluent speakers nearby. Machine translation can, in principle, render community texts into a language of wider communication and back again. And large language models hold out the further prospect of conversational tutors, predictive keyboards, and generative aids for writers working in a language that the wider digital world ignores.

Sober assessment of these capacities, however, reveals that the bottleneck is rarely where the hype locates it. Bird's notion of "sparse transcription" is instructive: rather than assuming that an endangered language must first be transcribed fully and phonetically before it can be processed, he proposes interpretive, iterative, word-level workflows that invite wider community participation and treat the spoken word, not the written transcript, as primary (Bird, "Sparse Transcription" 713). The reframing matters because it shifts the locus of value from the machine's output to the human relationships and interpretive acts that documentation entails.

The generative capacities of large language models invite a particular caution. A conversational tutor or a predictive writing aid is attractive because it appears to supply the immersive interaction that learners of an endangered language so rarely find, and because it can operate at any hour without taxing the small number of fluent speakers. But a model that has learned a near-extinct language from a few thousand sentences will, when prompted beyond that thin evidence, generate text that is grammatically plausible and culturally hollow, or simply wrong. The stochastic-parrot critique applies with redoubled force in this setting: where a high-resource language can absorb a model's occasional errors without consequence, a community reconstructing a written norm from fragmentary documentation has little capacity to detect or correct fluent fabrications, which may quietly propagate into teaching materials and learner usage (Bender et al. 617). Synthesis poses an analogous hazard. A text-to-speech voice trained on limited data may flatten the prosody, vowel length, or tonal distinctions that carry meaning, and a synthetic voice presented as authoritative can displace the embodied speech of elders rather than supplement it. The technologies are most trustworthy, in short, precisely where they are most modest, retrieving, organizing, and transcribing material that humans have authenticated, and least trustworthy where they are asked to originate language on a community's behalf. Littell and colleagues, in their careful assessment of Indigenous language technologies across the roughly sixty languages of Canada, similarly emphasize that the technologies most feasible and most useful, spell-checkers, text prediction, electronic dictionaries, and pedagogical tools, are often the unglamorous ones, and that the appropriate measure of a system is whether it serves a community's own goals rather than whether it advances the state of the art (Littell et al. 2620). Machine translation between an endangered language and a dominant one remains exceptionally hard at the scale of available data; the Cherokee-English system assembled by Zhang, Frey, and Bansal, trained on a painstakingly compiled corpus of only some fourteen thousand sentence pairs, achieved BLEU scores that, while a genuine contribution, signal serviceable assistance rather than fluent automatic translation (Zhang, Frey, and Bansal 577). The lesson is not that the technologies are useless but that their value is auxiliary: they can amplify documentation and learning, but they cannot manufacture the speakers, contexts, and intergenerational ties on which a living language depends.

Scarcity, Scale, and the Lure of the Large Model

The recent turn toward massively multilingual models has reframed the scarcity problem as one of transfer rather than abundance. The premise is that a single model trained jointly across hundreds of languages can draw on what it learns from data-rich languages to bootstrap competence in data-poor ones. The No Language Left Behind effort, which scaled neural machine translation to two hundred languages and published a substantial portion of its findings in the scientific literature, exemplifies the ambition and the genuine technical achievement of this approach, reporting marked gains over prior systems and releasing its models and evaluation benchmark openly (NLLB Team 841). Such projects rightly insist that historically de-prioritized communities deserve to be served, and their open-source posture is preferable to the proprietary enclosure of linguistic capacity.

And yet scale carries its own distortions. A model that "covers" two hundred languages does not thereby grant each an equal voice; coverage is uneven, evaluation for the smallest languages is thin, and the inclusion of a language in a benchmark can create a misleading impression of solved competence that masks brittle performance on the registers, domains, and dialects that matter to actual speakers. Bender and colleagues' critique of ever-larger language models, the "stochastic parrots" argument, bears directly here: systems trained on vast uncurated corpora tend to encode the perspectives of the dominant and well-resourced, carry significant environmental and financial costs, and produce fluent text without understanding, such that an apparent gain in coverage for a marginalized language may amount to little more than statistically plausible mimicry untethered from cultural ground truth (Bender et al. 610). For an endangered language, fluent-seeming but unreliable output is not a neutral byproduct; it can introduce errors into a fragile written norm, dilute orthographic and grammatical standards that communities have labored to establish, and arrogate to a machine the authority that properly belongs to elders and fluent speakers. The seductive proposition that scale will eventually reach every language thus obscures a deeper asymmetry: the languages are made to fit the technology, rather than the technology being shaped to fit the languages and the people who speak them.

Data Sovereignty and the Ethics of Care

If the technical literature exposes the limits of scale, the Indigenous data sovereignty literature exposes the politics of data itself. Much of the machinery of contemporary NLP rests, often tacitly, on an assumption that text and recordings are raw material freely available for collection. Bird names this assumption directly, observing that Indigenous people frequently respond with hostility to the proposition that their languages are simply "data ready for the taking," and warning that treating Indigenous knowledge as a commodity risks disenfranchising the local authorities who are its rightful custodians and reenacting the very dynamics of dispossession that endangered the languages to begin with (Bird, "Decolonising Speech and Language Technology" 3504). The point reframes the data-scarcity problem in moral rather than merely engineering terms: for many communities, the relevant material is not scarce so much as deliberately withheld, restricted to particular persons, seasons, or ceremonial contexts, and not available for indiscriminate ingestion at all.

The asymmetry of benefit sharpens the ethical stakes. When the linguistic material of a marginalized community is folded into a model whose principal beneficiaries are a technology firm, an academic career, or a benchmark leaderboard, the relationship reproduces a familiar colonial geometry in which value is extracted from the periphery and accumulates at the center, with the source community receiving, at best, a citation. Bird's analysis insists that this is not an incidental flaw to be patched with better consent forms but a structural feature of treating language as a commodity, one that can disenfranchise the very knowledge-holders, elders,

ceremonial specialists, and community linguists, whose authority sustains the language (Bird, "Decolonising Speech and Language Technology" 3504). Nor is consent a simple remedy, since the individual who reads a phrase into a microphone may not be the appropriate authority to release collectively held knowledge, and a permission granted today cannot anticipate the uses to which a general-purpose model might later put the data. These difficulties explain why the most thoughtful responses have moved from the language of individual consent to the language of collective governance and ongoing stewardship.

A constructive response has emerged in the form of explicit governance principles. The widely adopted FAIR principles, which hold that research data should be Findable, Accessible, Interoperable, and Reusable, were devised chiefly to maximize the technical reusability of data and say little about the rights and interests of the people from whom data derive (Wilkinson et al. 1). It is precisely this lacuna that the CARE Principles for Indigenous Data Governance address. Formulated by Carroll and colleagues, CARE stands for Collective Benefit, Authority to Control, Responsibility, and Ethics, and is offered not as a replacement for FAIR but as a necessary complement, insisting that any use of Indigenous data advance the self-determined wellbeing of the communities concerned and proceed under their authority (Carroll et al. 1). In the same spirit, the First Nations principles of OCAP, Ownership, Control, Access, and Possession, articulate a community's collective right to govern the creation, custody, and application of information about it, with possession functioning as the concrete mechanism by which ownership is asserted and protected (First Nations Information Governance Centre). These frameworks reorient the conversation from what a model can be trained to do toward who decides whether it may be trained at all, and on what terms the resulting benefit must flow back to the source community.

Te Reo Maori and the Practice of Guardianship

No initiative has demonstrated the practical force of data sovereignty more clearly than the work of Te Hiku Media, a charitable broadcaster of New Zealand's far north that has built automatic speech recognition for te reo Maori. Confronted with the prospect of handing hard-won language recordings to external technology firms, Te Hiku Media instead developed its own corpus through a community crowdsourcing campaign and released it under a bespoke instrument, the Kaitiakitanga License, grounded in the Maori concept of guardianship. The license rests on a deliberate refusal of the ownership model: data, it holds, is not owned but cared for under the principle of kaitiakitanga, with the organization acting as guardian rather than proprietor, and any benefit derived from the data flowing back to the people from whom it descends (Te Hiku Media). When approached by outside parties seeking to license or purchase the resulting models and data, the organization declined arrangements that would have alienated the community's control, on the ground that decisions about te reo must remain with its source community.

The Te Hiku case is instructive precisely because it does not reject the technology. It embraces ASR enthusiastically while subordinating the technical project to a relationship of accountability, demonstrating that sovereignty and innovation are not opposed. The model that emerges is one in which the community sets the purpose, retains the data, governs downstream use, and harnesses the speech recognizer to a thoroughly local end, the transcription of archival broadcasts and the support of learners, rather than to the abstract goal of expanding a global model's language inventory. It is the inversion of extraction: rather than a language being mined to enrich a distant system, a system is built and held within the community to enrich the language. This is data sovereignty not as theoretical principle but as engineering practice, and it offers a replicable template precisely because it began from the question of authority rather than appending ethics as an afterthought.

Scale Versus Specificity: Welsh, Cherokee, and the Limits of Generality

Beyond the Pacific, comparable efforts illuminate the persistent tension between technological scale and linguistic-cultural specificity. The Welsh experience demonstrates what sustained, well-resourced, community-aligned investment can accomplish for a language that, though far more vital than many, remains minoritized within an Anglophone state. Through the crowdsourced Paldaruo Speech Corpus, gathered via a purpose-built mobile application that allowed speakers across Wales to contribute recordings, researchers assembled the speech data necessary to develop Welsh speech technologies and, in turn, applications such as a Welsh-language voice assistant (Cooper, Jones, and Prys 4). The Welsh case is notable for its institutional scaffolding: governmental commitment, a dedicated technology unit, and a national strategy that treats language technology as public infrastructure rather than as a research curiosity. Notably, the corpus was designed with metadata, dialect coverage, and licensing in view from the outset, reflecting an understanding that data gathered carelessly serves no one well.

Cherokee occupies the opposite end of the vitality spectrum, and its case reveals how thin the margin can be. As Zhang, Frey, and Bansal note, fluent first-language speakers of Cherokee number only around two thousand and are declining, which renders every act of documentation an emergency and every corpus an irreplaceable artifact (Zhang, Frey, and Bansal 577). Their machine-translation work was undertaken explicitly in the service of revitalization and in collaboration with the community, yet the modest results testify to a hard truth: at such scales, statistical methods can assist but cannot substitute for the human labor of teaching and speaking. The contrast between Welsh and Cherokee maps onto the sociolinguistic stages that revitalization theory has long emphasized. Fishman's framework of reversing language shift locates the decisive variable not in technology but in intergenerational transmission, the passing of a language from parents to children in the home, without which institutional and technological supports float free of any self-sustaining base (Fishman). The classic practitioners' literature on revitalization concurs, foregrounding immersion schooling, master-apprentice mentorship between elders and learners, and the deliberate creation of domains in which the language is spoken (Hinton and Hale). Technology can enrich each of these, supplying dictionaries, transcribed archives, and learning aids, but it operates downstream of the human relationships that alone produce new speakers.

The matter of fidelity deserves particular emphasis, for it is where the abstract tension between scale and specificity becomes concrete. Endangered and Indigenous languages frequently present features that general systems handle poorly: unsettled or competing orthographies, substantial dialectal variation across small populations, and morphological complexity that strains tools calibrated on analytic European languages, as Mager and colleagues document for the Americas (Mager et al. 64). A model optimized for average performance across many languages has every incentive to smooth such variation away, collapsing dialects into a single standardized register and normalizing orthographic choices that a community may regard as identity-bearing. Yet for a language with few speakers, dialectal and orthographic particularity is not noise to be filtered but precisely the substance worth preserving. The Welsh corpus designers' attention to dialect coverage and to carefully structured metadata reflects an awareness that data gathered without such care can entrench a single variety at the expense of others, a soft homogenization that a revitalization effort can ill afford (Cooper, Jones, and Prys 4). Specificity, in other words, is not a luxury layered atop a working system; it is the very thing the community is trying to keep alive, and a technology that erodes it has defeated its own purpose even as its metrics improve.

This is the crux of the tension between scale and specificity. A general-purpose model optimizes for breadth and average performance; a revitalization program needs depth, fidelity,

and cultural appropriateness for a single community. The features that make a model commercially impressive, vast coverage, fluent generation, indifference to provenance, are frequently orthogonal or even hostile to what a small community requires. Pine and Turin, surveying the field of language revitalization, observe that the global concentration of communication into a handful of technologized world languages is itself a principal engine of language loss, which should give pause to any assumption that more powerful general technology is automatically beneficial (Pine and Turin). The risk is that endangered languages are enrolled as tokens in a project whose center of gravity lies elsewhere, gaining nominal representation in a benchmark while losing the specificity, and the control, that revitalization demands.

Toward a Constructive Settlement

A purely cautionary conclusion would be as inadequate as an uncritically optimistic one. The hazards canvassed here, extraction, the dilution of fragile norms, the illusion of competence, the displacement of community authority, are real, but so are the genuine contributions that well-governed technology can make to documentation and learning. The constructive settlement lies in inverting the default order of operations. In the dominant paradigm, data is gathered first and questions of ethics, benefit, and consent are entertained afterward, if at all; the alternative, exemplified by Te Hiku Media and increasingly codified in the CARE and OCAP frameworks, begins from the community's authority and purpose and treats the technical work as instrumental to ends the community has set (Te Hiku Media; Carroll et al. 1; First Nations Information Governance Centre). Under this inversion, the right questions are not "how much data can we obtain" and "how large a model can we train," but "what does this community want for its language," "who governs the resulting resources," and "how does the benefit return to the source."

Several practical commitments follow. Documentation methods that broaden participation and respect the primacy of the spoken word, such as Bird's sparse transcription, align technical workflows with community involvement rather than supplanting it (Bird, "Sparse Transcription" 713). Tooling priorities should follow the unglamorous but genuinely useful needs that Littell and colleagues identify, rather than the prestige tasks that dominate the research literature (Littell et al. 2620). Licensing and governance should be built in from the beginning, on the model of the Paldaruo corpus and the Kaitiakitanga License, rather than retrofitted (Cooper, Jones, and Prys 4; Te Hiku Media). And the field would do well to abandon the framing of endangered languages as deficient "low-resource" cases, a framing that, as Joshi and colleagues' taxonomy inadvertently makes vivid, measures languages against an industrial standard of data abundance that has nothing to do with their worth or their futures (Joshi et al. 6284). Researchers, for their part, must accept that the appropriate role of an outside technologist is often a supporting one, lending technical capacity to a community-led effort rather than authoring it, and measuring success by the strength of the relationship and the usefulness of the result to the community rather than by publication or benchmark performance. Such a reorientation asks the field to relinquish some of its accustomed authority, but it is the price of doing genuine rather than merely nominal good.

Conclusion

The arrival of large language models, speech recognizers, synthesizers, and translation systems has genuinely altered the landscape of language documentation, and it would be a failure of imagination to refuse these tools outright. They can compress the transcription of archives, support isolated learners, and lend new reach to communities determined to reclaim their tongues. But the central argument of this article is that the value of the technology is conditional and instrumental, never primary. The statistical paradigm that makes these systems

powerful is structurally mismatched to the smallest languages, and its default mode of acquiring data reproduces the extractive relations that helped endanger those languages in the first place. The decisive questions are therefore ethical and political rather than computational: questions of authority, benefit, consent, and control, for which the Indigenous data sovereignty literature and community-authored instruments such as the Kaitiakitanga License supply a workable answer. Above all, the measure of a language's future remains what it has always been, the willingness and ability of one generation to speak it to the next. Artificial intelligence can serve that transmission, enriching the contexts in which a language lives, but it cannot perform it. A language is not, finally, a dataset, and the most sophisticated model is no substitute for a grandmother speaking to a grandchild in the words her own grandmother gave her. The promise of these technologies for endangered languages will be realized only insofar as they are placed, deliberately and from the outset, in the service of that irreplaceable human relationship.

Works Cited

Bender, Emily M., et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2021, pp. 610–23.

Bird, Steven. "Decolonising Speech and Language Technology." *Proceedings of the 28th International Conference on Computational Linguistics*, International Committee on Computational Linguistics, 2020, pp. 3504–19.

---. "Sparse Transcription." *Computational Linguistics*, vol. 46, no. 4, 2020, pp. 713–44.

Carroll, Stephanie Russo, et al. "The CARE Principles for Indigenous Data Governance." *Data Science Journal*, vol. 19, no. 1, 2020, pp. 1–12.

Cooper, Sarah, Dewi Bryn Jones, and Delyth Prys. "Crowdsourcing the Paldaruo Speech Corpus of Welsh for Speech Technology." *Information*, vol. 10, no. 8, 2019, article 247.

Crystal, David. *Language Death*. Cambridge UP, 2000.

First Nations Information Governance Centre. "The First Nations Principles of OCAP." *FNIGC*, fnigc.ca/ocap-training/. Accessed 28 June 2026.

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

Hinton, Leanne, and Ken Hale, editors. *The Green Book of Language Revitalization in Practice*. Academic Press, 2001.

Joshi, Pratik, et al. "The State and Fate of Linguistic Diversity and Inclusion in the NLP World." *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, 2020, pp. 6282–93.

Kornai, András. "Digital Language Death." *PLoS ONE*, vol. 8, no. 10, 2013, article e77056.

Littell, Patrick, et al. "Indigenous Language Technologies in Canada: Assessment, Challenges, and Successes." *Proceedings of the 27th International Conference on Computational Linguistics*, Association for Computational Linguistics, 2018, pp. 2620–32.

Mager, Manuel, et al. "Challenges of Language Technologies for the Indigenous Languages of the Americas." *Proceedings of the 27th International Conference on Computational Linguistics*, Association for Computational Linguistics, 2018, pp. 55–69.

Moseley, Christopher, editor. *Atlas of the World's Languages in Danger*. 3rd ed., UNESCO Publishing, 2010.

NLLB Team. "Scaling Neural Machine Translation to 200 Languages." *Nature*, vol. 630, 2024, pp. 841–46.

Pine, Aidan, and Mark Turin. "Language Revitalization." *Oxford Research Encyclopedia of Linguistics*, Oxford UP, 2017.

Te Hiku Media. "Data Sovereignty and the Kaitiakitanga License." *Te Hiku Media*, 6 May 2022, tehiku.nz/te-hiku-tech/te-hiku-dev-korero/25141/data-sovereignty-and-the-kaitiakitanga-license. Accessed 28 June 2026.

Wilkinson, Mark D., et al. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data*, vol. 3, 2016, article 160018.

Zhang, Shiyue, Benjamin Frey, and Mohit Bansal. “ChrEn: Cherokee-English Machine Translation for Endangered Language Revitalization.” *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2020, pp. 577–95.