



Language Rights in the Digital Age: Expanding Scopes, Emerging Limits

Ms. Monamie Mukherjee

Assistant Professor, Department of English and Literary Studies Brainware University 398, Ramkrishnapur Road, Kazipara, Barasat, Kolkata-700125. West Bengal, India. mm.hu@brainwareuniversity.ac.in
ORCID ID: 0009-0001-8836-4751

ABSTRACT

Language rights have traditionally been framed as essential to cultural identity, equality, and democratic participation. However, recent developments have significantly reshaped their scope and limitations. This paper examines emerging challenges to language rights in the contemporary era, particularly in the context of technological advancement, globalization, and shifting political dynamics. It highlights how digital platforms and artificial intelligence systems risk reinforcing linguistic hierarchies by privileging dominant languages while marginalizing low-resource and indigenous ones. At the same time, the paper explores tensions between language preservation and national integration, as well as the practical difficulties of implementing mother-tongue education policies. Additionally, it addresses growing concerns over data ownership and the use of linguistic resources by private technology actors, raising questions about cultural sovereignty in the digital age. By critically analysing these developments, the study argues that language rights must be reconceptualized beyond formal recognition to ensure meaningful inclusion, equitable access, and protection of linguistic diversity in an increasingly interconnected world.

KEYWORDS

Language rights, Language technology, Linguistic marginalization, Social inclusion, AI governance.

ARTICLE HISTORY

Received: 02 July 2026
Revised: 15 July 2026
Accepted: 04 August 2026
Published: 05 September 2026

CITATION

Mukherjee M., (2026 Language Rights in the Digital Age: Expanding Scopes, Emerging Limits. *Journal of Indian Knowledge Studies (J. Indian Knowl. Stud.)*, 1(1), 26-29. <https://doi.org/10.71371/jiks.1126-29>

INTRODUCTION

To examine language right, we should first focus on the question, what is language right? Well, we can say, language rights are those rights that empowers one to use her/his own language in every sphere of life. Language rights movement typically talks about the discrimination faced by ethnic minorities. Speakers of languages of ethnic minorities often face discrimination on many levels, including linguistic. They are the most visible faces of such discrimination. Sometimes they enjoy some rights; tolerance-oriented and promotion-oriented. Tolerance oriented rights are the rights to use your own language at home and close-knit social functions, etc. This is a way of telling the minority groups that they should remain within a well-defined

usage pattern as far as their language is concerned. Promotion oriented rights however are those, which are more often than not denied to the ethnic minorities. This requires active support for use of the language in various public domains, otherwise the attempts made by the minority groups to maintain their language and culture, in comparison to the dominant majority language, remains hampered.

THE MAJOR MOVEMENTS RELATED TO LANGUAGE RIGHT

1. The language ecology movement

This movement is primarily concerned with preservation and revitalization of languages, relating

language death with the imminent collapse of total ecosystem. Situated within the biological diversity model, which metaphorically places Languages as species, it derives that each language is like a species in an ecosystem and that linguistic diversity is comparable to biodiversity. In that case, Language loss can be equated to species extinction. So, when a language dies, unique knowledge, culture, and worldview disappear, similar to extinction in biology. Thus, multilingualism can be a healthy ecosystem, where societies are with many languages, seen as more culturally resilient, just like ecosystems with many species are more stable. Just as invasive species can disrupt ecosystems, dominant global languages (e.g., English) can overshadow smaller ones, and that is what it had happened in many Indian societies and culture. So, the priority lies primarily with the language and not the speaker.

The main interest is Language, not the people. Languages should contain a storehouse of knowledge systems about the environment, animal species, technologies for cultivation and domestication, medicinal properties of plants, as well as cultural achievements like epic tales, riddles, lullabies etc. Linguistic enquiry into the grammatical system of these languages can contribute substantially to the scientific investigation on human cognition. Work in this domain brings researchers in close contact with the small groups of people speaking the language, very often resulting into language revitalization activities.

Another significant model is the 'Language nest model', where fluent elder speakers of the language take the responsibility of teaching the children how to 'speak and live' the language. For example, the Maori in New Zealand. This is usually feasible with communities with written material available. And also, the number of speakers need to be high. The master-apprentice model however, was followed in California in 1992. Here there are real-life languages facing death and the elder speakers are themselves not very fluent and have not been using the language since long time. In addition, the speaker number is so low in each group that there is practically no way to reverse the process of endangerment overnight. The goal of this model is modest compared to the language nest model. Here the idea is to take help of masters from within the community and put them together with apprentices who are expected to learn the basic skills over a period of time in simple, real-life activities. This model also does not attempt to revitalize the speaker base. Mufwene said, "language endangerment is a more wicked problem than has been acknowledged in the literature. It sometimes boils down to a choice between saving speakers from their economic predicament and saving a language. Seldom can both goals be congruent..."

Thus, the link between the welfare of language and its

speakers is complex. Language revitalization should be initiated only if there is expressed wish of the community for doing so, as sometimes the speakers do not have a choice. For examples, if we consider the cases in El Salvador in 1932, where following a peasant uprising, anyone who was identified as Indian risked being killed by the Salvadorian soldiers. This brings to light that loss of language and violation of people's right to decide for themselves are converged. In another change in the village of Gapun in Papua New Guinea, a large number of people have shifted from the local language Taiap to Tok Pisin as they moved out to work and brought the new language of status back to the village. But this change was so slow and gradual, that many of them did not even notice it, so here, no violation of rights has taken place. But in case of Southeast China again, the Sui clans have what is called clanlects, where language varieties differentiated along clan lines. In this community women are to be married outside their clan and move in to the husband's village. Children are to learn the father's clanlect and not the mothers. As a result, the mother's clanlect does not get transmitted to the next generation.

2. The linguistic human rights movement

This movement is situated within and draws its strength from the human rights movement, equating language with other basic human rights like food, safety and liberty. A distinctive characteristic of this movement is that these are considered necessary and basic, rather than enrichment-oriented ones. Learning a mother tongue is necessary while learning a foreign language is not, as mother tongue is the one which cannot be separated from its speakers. Preservation of language and protection of people go hand in hand. This movement also distinguishes subsets of linguistic human rights into individual and collective Rights

Whereas individual rights include the right to learn his mother tongue and have this identification accepted and respected by others, it also emphasises on use of mother tongues in most official situations, including schools and higher education. Everybody whose mother tongue is not the official language of the country, has the right to become bi/trilingual in the mother tongue/s and the official language. On the other hand, Collective rights include ensuring that indigenous peoples are allowed to 'reproduce themselves as distinct groups, with their own languages and cultures'. The idea of choice is a key element reflected in this movement, but only in case of official languages. The choice of mother tongue is not an issue with the supporters of this movement as even when the shift is voluntary, they argue that the speakers were not informed about the consequences of the shift and hence it does not qualify to be called a choice. The notion of essentialism associated with this movement,

i.e., language and ethnic identity are one and same, has been criticized severely. Hence the dichotomies of ‘Free choice Vs force’ and ‘Mother tongue Vs foreign language’, as well as ‘Necessity Vs enrichment’, are seen as the movement’s tendency towards essentialism clearly.

3. The minority language rights movement

This movement is set within the larger background of democratic rights; it aims at initiating a dialogue between the dominant and dominated, through equality. Language rights as part of a broader concern about the rights of minority groups. Language is given importance as a significant aspect of a group’s identity. Protection of language rights as critical part of a liberal democratic perspective and contributes to an equitable relationship between the speakers of dominant and minority languages. Priority of this movement is welfare of speakers rather than well-being of languages, because it enforces that Language is a group right.

LIMITATIONS AS THREATS TO LANGUAGE RIGHTS IN THE DIGITAL AGE

The rapid expansion of artificial intelligence (AI) and large-scale language technologies has intensified longstanding debates around **linguistic rights**, particularly in relation to **data ownership and control over linguistic resources**. Traditionally grounded in human rights discourse, language rights now intersect with digital infrastructures, where language is increasingly transformed into **data** for computational use. In this context, **linguistic resource in corpora**, speech recordings, and annotated dataset have become critical assets in AI development; their extraction, ownership, and commercialization raise new concerns that extend beyond classical language rights into domains of **data sovereignty** and **digital justice**. Language data is often treated as a commodity, characterized by large-scale scraping of online texts and speech, aggregation of datasets from multiple, often undocumented sources and also, weak attribution and unclear licensing structures.

The soaring dominance of ‘private technology actors’, or the corporations in language technology has led to a **reconfiguration of linguistic authority**. Traditional institutions like language academies, public bodies, etc. are losing influence as private actors control and **reshape notions of language ownership and legitimacy**, often side-lining public institutions and communities. This shift raises critical questions like - who owns language when it is encoded into AI systems? or, who has the authority to define “correct” or “standard” usage in algorithmic contexts? The answers increasingly lie with private entities rather than linguistic communities.

A major concern is the emergence of **data extractivism**, where linguistic data is harvested—frequently without consent from communities, particularly from marginalized or indigenous groups. Documented cases show that AI systems have been trained on indigenous language recordings scraped from the internet without community permission, raising questions about ownership and ethical use. Such practices echo colonial patterns of resource extraction, now applied to **cultural and linguistic inequality**. So, the key issues include lack of informed consent, absence of benefit-sharing mechanisms, and cultural appropriation of linguistic knowledge. The concept of **community data sovereignty** has emerged in response, asserting that communities should control how their linguistic data is collected, used, and shared.

Again, language technologies often reproduce **global inequalities** when most AI systems are trained predominantly on high-resource languages such as English, leading to systematic underrepresentation of others. This imbalance results into poor performance for “low-resource” languages, reduced digital presence of marginalized communities, and loss of culturally-specific knowledge. Scholars describe this as ‘**epistemic injustice**’, where certain knowledge systems and worldviews are excluded or distorted by technological systems, showing strong linguistic marginalisation.

GOVERNANCE GAPS AND ETHICAL CHALLENGES

Current legal and ethical frameworks struggle to address these issues. The unclear provenance of training data complicates questions of authorship and rights. Even when datasets are publicly available, their reuse in commercial AI systems raises unresolved legal questions. Ethical data collection requires meaningful community engagement, yet many AI systems rely on passive data extraction rather than participatory approaches; no consideration for consent and participation is observed ever. Power asymmetries are clearly evident when private corporations possess disproportionate control over data infrastructure, computational resources, and over the global distribution channels. Such concentration of power limits the ability of communities and public institutions to assert linguistic rights. **CARE principles**, i.e., Collective Benefit, Authority to Control, Responsibility, Ethics are violated often and **participatory data governance models** involving communities in decision-making fails often too. Scholars argue that ethical language technology must move from **data extraction to “data care”**, prioritizing long-term relationships with communities over short-term technological gains.

CONCLUSION

The intersection of language rights and AI-driven data economies reveal a fundamental tension where we find a failure in data governance; while language technologies promise increased access and preservation, they also risk reinforcing inequalities through **privileged control of linguistic resources**.

The growing concerns over data ownership highlight the need to re-assert linguistic rights as **collective and**

cultural rights, not merely individual ones to be downplayed with. It fails to develop in the process, a robust framework for ensuring that linguistic communities retain agency over their languages in digital environments. Ultimately, safeguarding linguistic diversity in the age of AI requires aligning technological innovation with principles of **justice, equity, and community sovereignty** on the same footing.

REFERENCES

1. Bird, S., & Simons, G. (2003). Language documentation and description. *International Journal of Language Documentation and Conservation*, 1(1), 1–20.
2. Erdocia, I., Schneider, B., & Migge, B. (2025). *Language in the age of AI technology*. Cambridge University Press
3. Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2024). Diversity and language technology: How language modeling bias causes epistemic injustice. *Ethics and Information Technology*, 26(1), Article 8. <https://doi.org/10.1007/s10676-023-09742-6>
4. Longpre, S., Mahari, R., Chen, A., Obeng-Marnu, N., Sileo, D., Brannon, W., Muennighoff, N., Khazam, N., Kabbara, J., Perisetla, K., Wu, X. A., Shippole, E., Bollacker, K., Wu, T., Villa, L., Pentland, S., & Hooker, S. (2024). *A large-scale audit of dataset licensing and attribution in AI*. *Nature Machine Intelligence*, 6, 975–987. <https://doi.org/10.1038/s42256-024-00878-8>
5. Schulz, M., & Loewen-Colón, J. (2026). *Preventing AI extractivism: The case for braiding Indigenous data justice with ABS for stronger AI data governance*. *AI & Society*. <https://doi.org/10.1007/s00146-026-02931-z>
6. Smart, A., Hutchinson, B., Amugongo, L. M., Dikker, S., Zito, A., Ebinama, A., Wudiri, Z., Wang, D., van Liemt, E., Sedoc, J., Olojo, S., Uwakwe, S., Wornyo, E., Schmer-Galunder, S., & Smith-Loud, J. (2025). *Socially responsible data for large multilingual language models*. arXiv. <https://arxiv.org/abs/2409.05247>
7. Sustainability Directory. (2025, June 25). *Community data sovereignty in AI-augmented TEK applications*. Sustainability Directory. <https://prism.sustainability-directory.com/scenario/community-data-sovereignty-in-ai-augmented-tek-applications/>

COPYRIGHT:

© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>