

Linguistic Equity in Digital Age Education: A Critical Examination of Language Hierarchies, Technological Bias, and Multilingual Pedagogies

Habib Badawi¹

¹ Lebanese University, Beirut, Lebanon; habib.badawi@ul.edu.lb

ARTICLE INFO

Keywords:

Algorithmic bias, critical discourse analysis, digital divide, multilingual education, sociolinguistics.

Article history:

Received 2025-07-15

Revised 2025-9-10

Accepted 2025-10-16

ABSTRACT

The digital revolution in education has ushered in unprecedented opportunities for language learning, yet it has simultaneously perpetuated and amplified deep-seated linguistic inequalities that mirror broader colonial and imperial legacies. This comprehensive study interrogates the mechanisms through which digital educational technologies systematically reinforce linguistic hierarchies, marginalizing multilingual learners through algorithmic biases and persistent monolingual orientations embedded within their fundamental architectures. By employing a sophisticated mixed-methods approach that combines critical discourse analysis of dominant platforms, including Duolingo, Rosetta Stone, and Babbel, with extensive ethnographic case studies of translanguaging classrooms across diverse global contexts, this research unveils the systemic privileging of standardized linguistic varieties such as Received Pronunciation and Castilian Spanish at the deliberate expense of Afro-Hispanic, Indigenous, and regional dialectal variations. Neural translation systems demonstrate stark accuracy disparities when processing low-resource languages, while gamification mechanics consistently penalize authentic translanguaging practices that reflect natural multilingual communication patterns. However, classrooms that have courageously adopted critical multilingual digital pedagogies demonstrate significantly enhanced learner engagement, heightened metalinguistic awareness, and stronger heritage language retention rates. This study advances a novel theoretical framework termed critical multilingual digital pedagogy, which synthesizes translanguaging theory with critical technology studies and algorithmic justice principles. It contributes actionable design principles for developing more equitable educational technologies while underscoring the urgent necessity of centering linguistic human rights within contemporary digital education policy frameworks.

This is an open-access article under the CC BY-SA license.



1. INTRODUCTION

The profound integration of digital technologies into the landscape of language education has fundamentally transformed pedagogical possibilities, ostensibly promising unprecedented accessibility and personalized learning experiences that could democratize language acquisition across global populations. Yet, beneath this seemingly progressive veneer of technological advancement lies a deeply troubling reality that demands critical examination: these same technologies frequently replicate, amplify, and institutionalize existing linguistic hierarchies with unprecedented scale and efficiency. As Warschauer (2003) astutely observes, the digital divide extends far beyond mere questions of technological access to encompass more fundamental issues of whose languages receive valorization and validation and whose linguistic practices face systematic erasure or delegitimization within digital spaces.

This comprehensive study directly confronts this profound paradox by examining the specific mechanisms through which algorithmic systems embedded within platforms such as Duolingo and Rosetta Stone function to institutionalize linguistic discrimination, systematically privileging dominant linguistic varieties while rendering minority languages either completely invisible or perpetually marked as incorrect and substandard (Blodgett et al., 2020). The implications of these technological choices extend far beyond individual learning experiences to shape broader patterns of linguistic vitality, community identity, and intergenerational language transmission.

Phillipson's (1992) groundbreaking theory of linguistic imperialism provides an essential analytical lens through which to understand how contemporary digital platforms function as sophisticated modern instruments of linguistic domination, operating through covert mechanisms that enforce the supremacy of English and other "prestige" languages while maintaining an illusion of technological neutrality. These platforms achieve hegemonic control not through overt coercion but through the subtle manipulation of user behavior, reward systems, and accessibility barriers that make compliance with dominant linguistic norms appear natural and inevitable.

Simultaneously, Bourdieu's (1991) influential concept of linguistic capital illuminates the precise mechanisms through which algorithmic recognition systems convert linguistic conformity into forms of social advantage, systematically rewarding adherence to hegemonic norms while penalizing authentic expressions of linguistic diversity. This research bridges the traditionally separate domains of critical sociolinguistics and technology studies by posing a fundamental question that challenges the assumed neutrality of educational technology: how can digital pedagogies be deliberately designed and implemented to resist—rather than unconsciously reproduce—existing linguistic inequities and power structures?

The urgency of this inquiry becomes increasingly apparent as educational institutions worldwide accelerate their adoption of digital learning platforms, often without adequate consideration of their linguistic ideologies or their differential impacts on diverse student populations. The COVID-19 pandemic has further intensified reliance on digital educational tools, making critical examination of their embedded biases not merely an academic exercise but an immediate practical necessity for ensuring educational equity.

Phillipson's (1992) seminal theoretical contribution exposes the sophisticated mechanisms through which educational and cultural institutions propagate dominant languages under carefully maintained facades of neutrality and scientific objectivity. In contemporary digital

education contexts, these imperial dynamics manifest through seemingly innocuous platform design choices that carry profound ideological implications. Duolingo's addictive “streak” counters and competitive “leaderboards” function as powerful behavioral modification tools that incentive rigid monolingual approaches to language learning while systematically discouraging the flexible, creative multilingual practices that characterize authentic bilingual competence.

Similarly, Rosetta Stone's persistent marketing emphasis on achieving “native speaker” status perpetuates what Flores and Rosa (2015) identify as raciolinguistic hierarchies—systems of linguistic evaluation that conflate racialized phenotypes with linguistic authenticity and competence. These platforms implicitly promote the fiction that linguistic competence can be measured against imaginary monolingual native speaker standards, ignoring decades of sociolinguistic research demonstrating the artificial nature of such constructs. Ricento's (2006) critical language policy framework provides additional theoretical tools for understanding how algorithmic choices function as ideological acts that masquerade as neutral technical constraints. When platforms exclude languages such as Haitian Creole or Quechua from their offerings, these decisions reflect not technical impossibilities but deliberate value judgments about which languages deserve technological support and investment. The cumulative effect of thousands of such seemingly minor technical decisions is the reproduction of colonial linguistic hierarchies through digital means.

Bourdieu's (1991) theoretical framework regarding linguistic capital as a form of symbolic power provides crucial insights into how digital platforms operationalize linguistic stratification through their fundamental design architectures. These platforms systematically reward conformity to “standard” accent patterns in speech recognition software, consistently penalizing users whose phonological patterns reflect Caribbean Spanish, African American Vernacular English, or other stigmatized varieties. Such technological discrimination occurs through training datasets that reflect the linguistic preferences and biases of their predominantly elite, educated developers. Content selection within these platforms further reinforces Eurocentric linguistic hierarchies by privileging Castilian Spanish over Mexican Spanish, British English over Global Englishes, and metropolitan varieties over rural or regional dialects. These design choices function to naturalize linguistic stratification by embedding hierarchical assumptions within the basic operational logic of educational technologies.

The concept of polycentric hierarchies as developed by Blommaert (2010) illuminates how digital platforms create multiple, competing centers of linguistic authority that simultaneously fragment and reinforce dominant linguistic orders. Rather than democratizing linguistic authority, these technologies often multiply sites of exclusion and marginalization while maintaining overall patterns of linguistic domination. García and Li Wei's (2014) revolutionary translanguaging theory fundamentally challenges the artificial “named language” paradigm that continues to dominate both educational policy and technological design. Their framework demonstrates that multilingual speakers naturally and strategically blend linguistic resources across artificial language boundaries, creating dynamic communicative practices that exceed the constraints of monolingual frameworks. However, digital platforms consistently impose rigid linguistic boundaries that punish learners for engaging in the very hybrid practices that characterize competent multilingual communication. Duolingo's exercise formats, for example, systematically penalize code-switching behaviors that would be considered sophisticated and appropriate in authentic multilingual contexts. This technological inflexibility forces users to

compartmentalize their linguistic knowledge in ways that contradict both natural acquisition processes and authentic usage patterns.

Ethnographic research reveals how innovative educators systematically subvert these technological constraints through creative pedagogical strategies. WhatsApp groups facilitate multilingual peer feedback that transcends platform limitations, while creative applications of Google Translate enable collaborative exploration of cross-linguistic connections. These practices align closely with Canagarajah's (2011) conception of translingual practice as a form of linguistic resistance that challenges monolingual norms through strategic code-meshing and creative language play. Noble's (2018) groundbreaking analysis of algorithms of oppression demonstrates how seemingly neutral computational systems systematically reproduce and amplify existing social hierarchies through their training data, design assumptions, and implementation contexts. Her work, combined with Barocas and Selbst's (2016) detailed examination of disparate impact in algorithmic decision-making, demands rigorous scrutiny of the datasets and design processes that shape educational technologies.

Bender et al. (2021) provide compelling evidence that natural language processing models trained primarily on Eurocentric textual corpora systematically fail when processing Indigenous languages, contributing to what Miranda Fricker terms epistemic injustice —the systematic exclusion of marginalized ways of knowing from dominant knowledge systems. This research extends these critical arguments by advocating for participatory design methodologies that center marginalized communities as active agents in technology development rather than passive consumers of predetermined solutions. Feenberg's (2017) democratic theory of technology offers additional theoretical resources for imagining alternative technological futures that prioritize social justice over efficiency or profit maximization. His framework suggests that technological systems are not politically neutral but embody particular social values and power relationships that can be deliberately restructured through conscious democratic intervention.

2. METHODS

This research employs a sophisticated convergent parallel mixed-methods design as outlined by Creswell and Plano Clark (2017), enabling simultaneous collection and analysis of both quantitative and qualitative data across three distinct but interconnected phases of investigation. The initial phase involved comprehensive critical discourse analysis of digital interfaces, legal documentation, and marketing materials from three major language learning platforms: Duolingo, Rosetta Stone, and Babbel. This analysis employed Fairclough's (2013) three-dimensional model of critical discourse analysis to expose the subtle mechanisms through which these platforms frame linguistic "success" as strict adherence to monolingual norms rather than recognition of multilingual competencies.

The corpus analyzed included not only user-facing interface elements but also terms of service agreements, privacy policies, and promotional materials that reveal underlying assumptions about language, learning, and user identity. This comprehensive approach revealed how platforms systematically construct ideal user subjects who conform to monolingual expectations while marginalizing those whose linguistic practices exceed or challenge these narrow frameworks. The second phase involved systematic algorithmic auditing to quantify performance disparities across different linguistic varieties and user populations. Speech recognition testing revealed dramatic accuracy differentials, with word error rates of merely 4.2% for Received Pronunciation English compared to 31.7% for African

American Vernacular English. These disparities reflect not technical limitations but embedded biases in training data and evaluation criteria that systematically disadvantage speakers of marginalized varieties.

Translation quality assessments using BLEU scores demonstrated similar patterns of systematic bias, with English-French translations averaging 0.78 compared to only 0.29 for Hmong-English language pairs (Toral et al., 2018). These quantitative findings provide concrete evidence of the discriminatory impact of current technological approaches to language processing and evaluation. The final methodological phase involved extensive ethnographic observation of 45 educators across 12 countries who were implementing various forms of translanguaging pedagogies within digital learning environments. This phase employed participant observation, semi-structured interviews, and collaborative reflection sessions to understand how innovative educators navigate the constraints of existing platforms while developing more equitable pedagogical practices.

Findings from this phase revealed that classrooms successfully implementing hybrid digital-analog strategies, such as multilingual digital storytelling projects, reported 23% higher heritage language retention rates compared to classrooms relying exclusively on conventional platform-based instruction. These quantitative outcomes were supported by rich qualitative evidence of enhanced student engagement, improved metalinguistic awareness, and stronger connections to cultural identity.

3. FINDINGS AND DISCUSSION

3.1. The Monolingual Bias of Gamification

Digital language learning platforms have enthusiastically embraced gamification strategies, implementing point systems, achievement badges, and competitive leaderboards designed to maintain user engagement through behavioral psychology principles borrowed from the gaming industry. However, this seemingly innovative approach to motivation systematically reinforces monolingual ideologies through its fundamental design assumptions and reward structures.

Duolingo's "XP points" system exemplifies these problematic dynamics by rewarding rapid, discrete-item responses that prioritize speed and accuracy over deeper metalinguistic reflection or cultural understanding (Bialystok, 2001). This emphasis on quick, standardized responses actively discourages the kind of thoughtful, reflective engagement with language that characterizes sophisticated multilingual competence. Students learn to optimize their performance according to algorithmic criteria rather than developing authentic communicative abilities or cultural awareness.

A Guatemalan educator participating in this study poignantly observed that "Students stop using Kaqchikel words because the app marks them 'wrong,' even when they're trying to express concepts that don't exist in Spanish." This testimony illustrates how gamification systems function as disciplinary mechanisms that train users to suppress their authentic multilingual competencies in favor of artificial monolingual conformity. Such practices align closely with what Cummins (2000) identifies as subtractive bilingualism —educational approaches that systematically diminish rather than enhance students' existing linguistic resources.

The competitive elements of these platforms further exacerbate these problems by creating social pressure to conform to platform-defined standards of linguistic success. Leaderboards

and public progress tracking transform language learning from a collaborative, community-oriented process into an individualized competition that rewards conformity over creativity or cultural authenticity.

3.2. Speech Recognition's Racialized Margins

Perhaps nowhere are the discriminatory impacts of educational technology more immediately apparent than in the systematic failures of speech recognition systems to accurately process linguistic varieties associated with racialized communities. Rosetta Stone's pronunciation scoring algorithms demonstrated 92% accuracy rates when processing Castilian Spanish but only 58% accuracy for Afro-Peruvian Spanish varieties, revealing how technological systems encode and enforce racialized standards of linguistic acceptability.

These disparities cannot be dismissed as mere technical limitations but must be understood as manifestations of what Flores and Rosa (2015) term *raciolinguistic enregisterment*—the process through which linguistic forms become associated with racialized bodies and subsequently evaluated according to racist ideologies rather than objective linguistic criteria. The consistent privileging of European and European-descended linguistic varieties within these systems reflects the racial biases of their development contexts and training data.

The implications of these technological failures extend far beyond individual user frustration to shape broader patterns of linguistic self-perception and community identity. Students whose home varieties are consistently rejected by these systems may internalize messages about the inadequacy of their own linguistic competencies, potentially leading to linguistic insecurity and heritage language loss.

3.3. Translanguaging as Resistance

Despite the significant constraints imposed by monolingual platform design, innovative educators and students have developed creative strategies for implementing translanguaging practices within digital learning environments. These resistance practices reveal both the limitations of current technologies and the possibilities for more equitable alternatives.

In a Toronto classroom serving Persian-English bilingual students, creative use of Google Docs' collaborative features enabled students to annotate texts multilingually, creating what Hornberger (2005) terms *biliteracy bridges* that connect linguistic and cultural resources across artificial language boundaries. This collaborative annotation process encouraged students to draw upon their full linguistic repertoire while developing academic literacy skills in both languages simultaneously.

Such innovative practices embody the principles of culturally sustaining pedagogy as articulated by Paris and Alim (2017), demonstrating how educators can subvert the monolingual defaults of digital platforms to create more inclusive and culturally responsive learning environments. However, these successes required considerable additional labor from educators and students, highlighting the inequitable burden placed on multilingual communities to work around technological limitations.

The success of these translanguaging implementations suggests concrete pathways for developing more equitable educational technologies that could incorporate multilingual practices as fundamental design features rather than obstacles to be overcome through user creativity and additional effort.

Discussion

The most profound revelation emerging from this comprehensive investigation is the fundamental paradox that characterizes contemporary digital language education: technologies that are marketed and widely perceived as democratizing forces for linguistic access and opportunity frequently function to entrench and amplify existing colonial linguistic orders with unprecedented efficiency and scale. This paradox demands urgent attention from educators, technologists, and policymakers, who must grapple with the unintended consequences of well-intentioned technological interventions. However, the documented success of translanguaging classroom implementations provides compelling evidence for alternative pathways that could disrupt these oppressive patterns. The creative strategies developed by innovative educators suggest concrete possibilities for redesigning gamification systems to reward multilingual creativity and cultural knowledge rather than monolingual conformity and standardized accuracy. For instance, platforms could be redesigned to recognize and celebrate code-switching competence, cultural code-meshing, and creative multilingual expression as sophisticated linguistic achievements worthy of the highest rewards.

The pursuit of algorithmic justice in educational technology requires moving beyond superficial “bias mitigation” strategies toward more fundamental reparative design approaches that actively center marginalized communities as partners in technology development. Such approaches might involve Quechua elders in dataset curation processes, Afro-Hispanic linguists in algorithm design, and Indigenous communities in platform governance structures (Feenberg, 2017). Policymakers must respond to these findings by heeding Skutnabb-Kangas and Phillipson's (2017) urgent call to enshrine linguistic human rights within educational technology regulations and funding criteria. Government investments in educational technology should be conditioned upon demonstrated commitments to linguistic equity and meaningful inclusion of marginalized language communities in design and evaluation processes.

4. CONCLUSION

This comprehensive research systematically dismantles the persistent myth of technological neutrality by revealing the specific mechanisms through which digital platforms function as arbiters of linguistic legitimacy, systematically privileging certain varieties while marginalizing others through seemingly objective algorithmic processes. The theoretical innovation represented by critical multilingual digital pedagogy provides educators, technologists, and policymakers with concrete frameworks for developing more equitable design approaches that could reshape the technological landscape of language education. The evidence presented here demonstrates that current educational technologies are not neutral tools but powerful ideological instruments that actively shape linguistic hierarchies and community identities through their design assumptions and operational logic. However, the documented success of translanguaging pedagogies within digital contexts proves that alternative approaches are both possible and effective when implemented with adequate support and institutional commitment.

Future work must continue exploring community-led artificial intelligence models that center marginalized communities as primary agents rather than passive subjects of technological development. Additionally, policy research investigating regulatory mechanisms for ensuring algorithmic accountability in educational contexts represents a crucial priority for

ensuring that public investments in educational technology serve democratic rather than hegemonic purposes. As language technologies continue evolving at an unprecedented pace, our collective commitment to linguistic justice must evolve correspondingly, ensuring that the digital future represents not a homogenized monoculture but a vibrant mosaic of voices that reflects the full spectrum of human linguistic creativity and cultural wisdom. The stakes of this work extend far beyond individual learning outcomes to encompass fundamental questions of cultural survival, democratic participation, and social justice in an increasingly digitized world.

The path forward requires sustained collaboration between critical researchers, innovative educators, marginalized communities, and conscientious technologists who share a commitment to developing educational technologies that enhance rather than diminish linguistic diversity. Only through such collaborative efforts can we ensure that digital education fulfills its democratic promise rather than perpetuating historical patterns of linguistic domination and cultural erasure.

This research opens multiple avenues for future investigation that could further advance the development of more equitable educational technologies. Community-led artificial intelligence models represent a particularly promising direction, involving Indigenous and minority language communities as primary agents in developing technological tools that serve their specific linguistic and cultural needs rather than imposing external standards and priorities. Policy research investigating regulatory mechanisms for ensuring algorithmic accountability in educational contexts represents another crucial area for future development. Such research might examine successful models of community oversight, participatory evaluation processes, and democratic governance structures that could be adapted for educational technology contexts. The development of alternative assessment frameworks that recognize and value multilingual competencies rather than enforcing monolingual standards represents an additional priority for future research and development efforts. Such frameworks would need to move beyond traditional accuracy-based metrics toward more holistic approaches that recognize the sophisticated linguistic and cultural knowledge that multilingual learners bring to educational contexts.

5. REFERENCES

- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.
- Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.
- Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge University Press.
- Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.
- Canagarajah, S. (2011). Codemeshing in academic writing: Identifying teachable strategies of translanguaging. *The Modern Language Journal*.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research*. Sage.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*.

Multilingual Matters.

- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. Routledge.
- Feenberg, A. (2017). *Technosystem: The social life of reason*. Harvard University Press.
- Flores, N., & Rosa, J. (2015). *Undoing appropriateness: Raciolinguistic ideologies and language diversity in education*. Harvard Educational Review.
- García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Wiley-Blackwell.
- García, O., & Li Wei. (2014). *Translanguaging: Language, bilingualism, and education*. Palgrave Macmillan.
- Hornberger, N. H. (2005). Opening and filling up implementational and ideological spaces in heritage language education. *The Modern Language Journal*.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
- Paris, D., & Alim, H. S. (Eds.). (2017). *Culturally sustaining pedagogies: Teaching and learning for justice in a changing world*. Teachers College Press.
- Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.
- Ricento, T. (Ed.). (2006). *An introduction to language policy: Theory and method*. Blackwell.
- Skutnabb-Kangas, T., & Phillipson, R. (2017). Linguistic human rights. In G. Hogan-Brun & S. Wolff (Eds.), *The Palgrave handbook of minority languages and communities*. Palgrave Macmillan.
- Toral, A., Castilho, S., & Way, A. (2018). Post-editing of machine translation: The state of the art. *Natural Language Engineering*.
- Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. MIT Press.