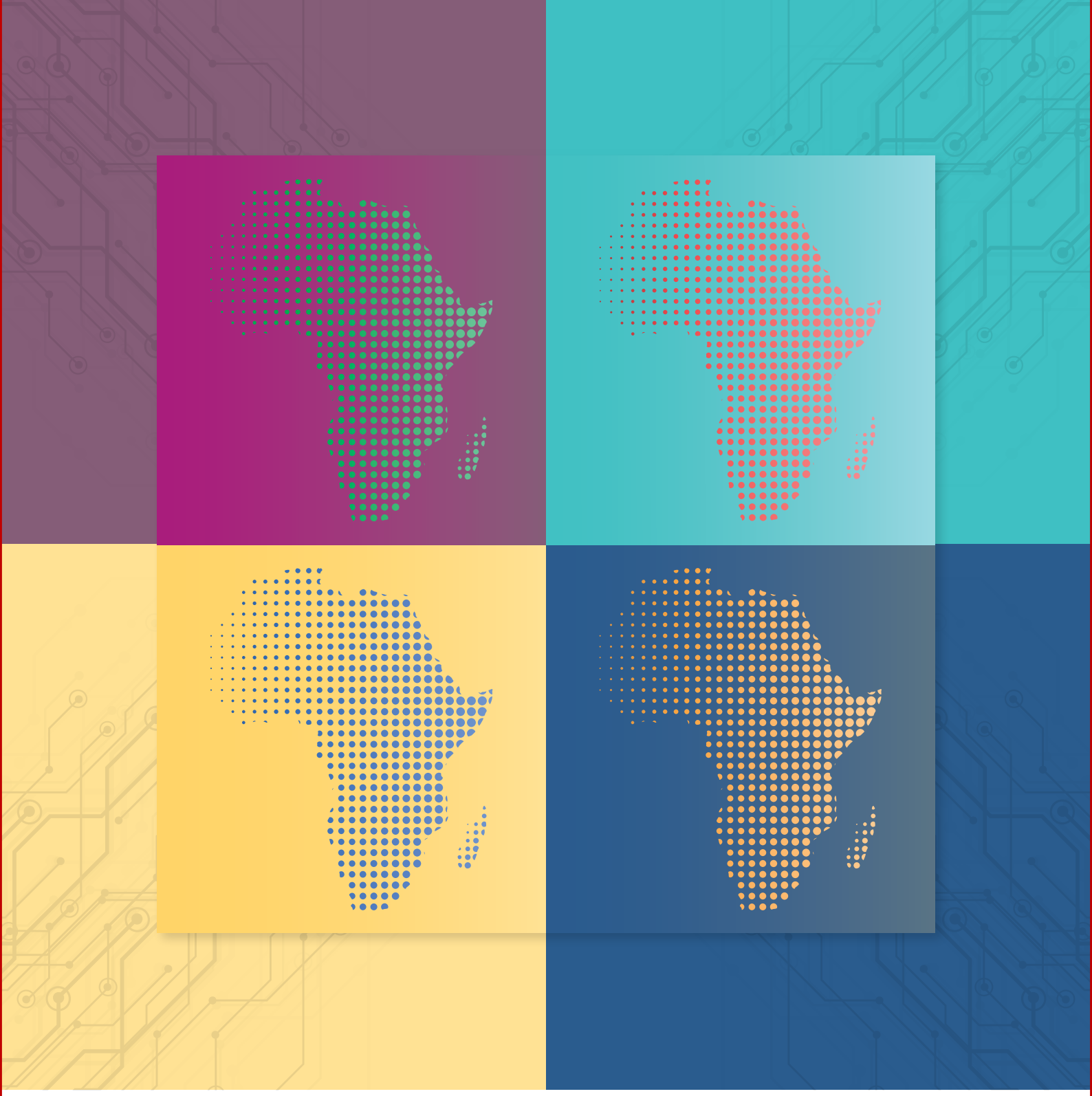

Digital Policy Studies

Digital Policy Studies is an open-access, peer-reviewed interdisciplinary academic journal focused on the empirical, critical and intersectoral study of subjects related to digital policy and the fourth industrial revolution, cybersecurity, the digitalisation of politics, the digital economy, information and communication technology, the convergence of technology and society, new media and related topics. The Journal publishes research articles, review articles and policy commentary in designated sections.

Published by the Department of Politics and International Relations at the University of Johannesburg.

Volume 4 ⊙ Number 1: 2025

UNIVERSITY OF JOHANNESBURG

Contents

EDITORIAL SECTION

Introduction to the Special Issue 'African Languages in the Digital Age'

Sizwe Zwelakhe Dlamini
Guest Editor
Department of African Languages
University of Johannesburg
Johannesburg, South Africa
sizwed@uj.ac.za

Introduction

The rapid pace of digital expansion across Africa has sparked important questions about the future of African indigenous languages. These questions serve as an entry point for deeper reflection, critical inquiry, and strategic thinking about the role these languages should play in public life, education, knowledge production, and technological innovation. Despite being central to the cultural identities and everyday practices of millions of speakers across the continent, African indigenous languages remain at a margin in the digital epoch. Nonetheless, viewed positively, this epoch has a strong potential of gradually opening new spaces in which these languages can be used, developed, refined, revitilised, and preserved. This special issue titled, "African Languages in the Digital Age", is an attempt to discuss how digital technologies are transforming indigenous African languages ecologies, and how these languages provide platform for the digital landscape to be challenged and reshaped. It brings together interdisciplinary scholarship that critically investigates the link between language, technology, and issues of participation, transformation, use, and access. The issue covers peer-reviewed articles that touch on the following themes:

- The incorporation of African indigenous languages to improve ICT comprehension, access and uptake by African populations;
- The incorporation of African indigenous languages in ICT development and use in human activity such as in education, e-government, e-commerce, and banking system;
- The possibilities, benefits, and challenges of Artificial intelligence (AI) in teaching and learning.
- The linguistic complexities and challenges between African indigenous languages and ICTs, and possible solutions;
- The relationship between African indigenous languages, ICS access, and equitable access to digital economic opportunities and access to basic services.

Collectively, the contributions in this issue highlight several important factors. Firstly, they raise questions about the state of language technologies in enabling digital participation, equitable access, and linguistic equity, while revealing inherent political and methodological challenges in the development of tools for under-resourced languages. Secondly, they highlight the socio-linguistic aspect of digital spaces, demonstrating how online platforms

can become the centre of activism and multilingualism. They evaluate the value of artificial intelligence and online platforms in reimagining practices like teaching and learning and service delivery, for instance. Thirdly, the contributions in this issue raise policy and planning questions, scrutinising how states, institutions, and communities can aid and contribute to the development and digital integration of African indigenous languages, in ways that promote inclusivity and sustainability.

By foregrounding multi-disciplinary perspectives, such as sociology, political science, anthropology, public management and governance, education, translation studies, and computational linguistics, this special issue attempts to demonstrate that overall, the digital age poses both challenging and transformative possibilities for African indigenous languages. More importantly, it demonstrates the significance of African scholars and communities in shaping the multilingual digital future. Through this issue, the aim is not only to document current research trends and developments, but also to inspire more research, potential collaborations, and policy intervention strategies that position and uphold African indigenous languages at the centre of the continent's digital transformation. There is still a lot that needs to be done to address the issue of the underrepresentation of African indigenous language in digital spaces, but there is hope because of the work that has already begun in the continent. As the common isiZulu idiomatic expression avers, "Inja iyawaqeda amanzi ngolimi" (The dog finishes water by licking).

Exploring the Integration of South African Indigenous Languages in ICT for Economic Growth

Busisiwe Thabisa Sibizo
Department of Politics and International Relations
University of Johannesburg
Johannesburg, South Africa
btsibizo@gmail.com

Abstract

Information and Communication Technology (ICT) is vital for promoting economic participation and growth. However, the prevalence of English as the main language of ICT presents a significant hurdle for many individuals, especially in Africa, hindering their full involvement in the digital economy. Language is essential for understanding and interaction, and as industries increasingly move to digital platforms, integrating South African indigenous languages into ICT has become necessary for wider participation. Despite extensive research emphasising the need to incorporate African languages into ICT, highlighting key advantages and challenges, a considerable gap remains in examining practical approaches for implementation and a focus on a country perspective with South Africa as one of the examples. This qualitative study aims to explore the reasons behind the slow uptake of South African indigenous languages in ICT and suggest potential strategies for their integration. Data was gathered through a desktop review.

The findings reveal a strong demand for incorporating African languages into ICT; however, there is limited investment to support this initiative. The study also highlights issues with language translation, as many African languages lack support from commonly used translation tools, underscoring the need for greater involvement from South African indigenous language speakers in the development and application of these tools. Additionally, the motivation to embrace South African indigenous languages is often insufficient, as many South African language speakers may not fully appreciate the value of their languages due to the longstanding prevalence of English in global communication. The research suggests that highlighting the significance of these languages, enhancing their perceived worth, employing South African indigenous language speakers for ICT development, and securing increased investment are crucial steps for successful integration.

Keywords: ICT, Economic Development, South African Languages, Digital Inclusion, Language Adoption, and Digital Divide.

Introduction

As digital technology becomes increasingly integrated into Africa, particularly within economic, educational, and social sectors, access to these digital tools remains inconsistent. The digital divide stems not only from infrastructural and affordability issues but also from linguistic exclusion. While this issue affects Africa as a whole, the focus of this study is South Africa. According to Siziba and Maseko (2024: 4), a new language policy was introduced along with the 1996 Constitution in South Africa to direct and influence language use in the post-apartheid era. This policy recognises twelve (12) official languages; namely, Sepedi, Sesotho, Setswana, Siswati, Tshivenda, Xitsonga, Afrikaans, English, isiNdebele, isiXhosa, and isiZulu. This also includes sign language. However, the policy seems to remain undermined as only English and Afrikaans continue to be the dominant languages in instruction, government, and commerce in most contexts.

It is crucial to highlight that many South Africans struggle with reading, writing, understanding, and expressing themselves in English which significantly impacts their interactions in different sectors, especially the ones with digital tools. As noted by Onamadike (2025) in an opinion piece for "Mail & Guardian", language inclusion is frequently overlooked in the development of civic technology; design decisions are often made by teams who do not consider the linguistic realities of the users they aim to serve. Onamadike (Op cit.) highlights that most African civic tech platforms are designed in English or French, neglecting the indigenous languages that users are comfortable with such as Swahili, Yoruba, or isiZulu. Also, he states that "a key reason for this dominance is that English is the primary language of the internet where most training data for language technologies used in digital tools comes from." Furthermore, he acknowledges that South Africa has multiple languages and a high number of civic tech initiatives on the continent; however, government websites, mobile apps, and AI-driven chatbots are English-only. It is to be noted that government websites and mobile apps are significant tools to engage in economic activities, especially the process of searching for jobs or being updated about recent business or job opportunities by government institutions. Therefore, the failure to interact with these, due to language barriers, affects overall economic development, especially with this significant shift to a digital world.

The existing divide is exacerbated by factors such as infrastructure, data and device affordability, and the language barriers discussed in this study. Key obstacles to bridging the digital divide include insufficient infrastructure such as internet connectivity, fixed broadband, and mobile telecommunications alongside affordability concerns, as devices and data remain costly. This is coupled with a lack of knowledge and skills necessary to utilise digital technologies effectively (FoodBev SETA, 2022: 5-6). Consequently, when individuals encounter these devices through educational institutions, workplaces, or government services, they often struggle to engage, especially if they are new to the technology. The added complexity of language barriers further impedes their ability to adapt.

This difficulty and prolonged adjustment, due to language barriers, hinders engagement with economic activities, development, and overall growth. This affects ICT businesses given that these businesses fail to reach a high target of diverse individuals. Non-English speaking individuals who are seeking employment and small businesses that are owned by non-English speaking individuals are mainly affected by this. This is because language barriers reinforce exclusion from digital markets, financial tools, e-governance systems, and general everyday duties that require the use of digital tools. Mahlaule et al. (2024: 691) conducted a study on the constraints, challenges, and opportunities faced by entrepreneurs

and digital businesses in South African townships and rural areas. In their study, they highlight that businesses fail at effectively advertising their products and services to customers with diverse language and cultural backgrounds. They further highlight that the digital sphere has great potential for economic growth and inclusion; however, language and cultural diversity should be considered to ensure effective implementation. It is therefore important to note that overcoming key issues such as language barriers will assist rural and township businesses to "tap into the global markets, fostering economic diversification, and poverty alleviation" (Mahlaule et al., 2024: 694).

Information and Communication Technology (ICT) has emerged as a vital tool for fostering economic participation and growth. However, the lack of language inclusion within ICT creates barriers that impede this growth, as comprehension, knowledge, and skills are essential for successful engagement. The predominance of English in ICT tools restricts full participation, particularly for non-English speakers. Thus, this study is critical in demonstrating how the incorporation of all twelve (12) official South African languages into ICT tools can enhance economic participation, leading to greater economic growth. The study outlines the issues associated with language translation, as commonly used translation tools inadequately support many South African languages. These tools dominantly use English in South Africa. As highlighted by Onamadike (2025) earlier, "a key reason for this dominance is that English is the primary language of the internet where most training data for language technologies used in digital tools comes from." This situation calls for enhanced involvement from indigenous South African language speakers in the development and application of these tools. Additionally, there is often inadequate motivation to embrace South African languages since many of these language speakers may not fully recognise the value of their languages, given the longstanding dominance of English in global communication. The research suggests that raising awareness about the importance of these languages, enhancing their perceived value, employing South African language speakers in ICT development, and securing increased investment are essential steps for successful integration.

This qualitative exploratory study will utilise secondary data, specifically through desktop research. The study is structured as follows: it includes an introduction that provides background information, defines the research problem, and outlines the research significance and objectives. A literature review that situates the study within existing scholarship, while the discussion and findings section presents and interprets the collected data. A methodology section that explains the research approach and the theoretical or conceptual framework, defines key concepts, and theoretically grounds the study. Finally, the recommendation and conclusion encapsulate the findings and propose actionable steps.

Literature Review

The literature review compiles and analyses existing research, emphasising prior discussions surrounding the study's topic. During the literature collection process, two main themes emerged: the relationship between ICT and economic growth, and the role of ICT in integrating indigenous languages.

ICT and Economic Growth

Adeleye and Eboagu (2019:32) indicate that following the year 2000, the African continent emerged as the third-fastest growing region globally. This significant achievement can

be largely attributed to enhanced domestic investment in the services sector, which was notably influenced by improved ICT usage and efficiency. ICT services in rural areas of Africa have played a crucial role in transforming the lives and livelihoods of local populations. These changes have manifested in various ways such as job creation, income increases, cost reductions, decreased uncertainty and risk, and strengthened familial relationships, among other advantages. Furthermore, Saba et al. (2025: 444) noted that, as ICT continues to infiltrate every sector of the economy, its historical, contemporary, and potential future impacts on Africa's economy are significant, despite the various challenges it encounters. This highlights the significant challenges associated with fully integrating ICT. However, these challenges must be addressed considering that ICT is becoming essential in every sector. Without basic ICT skills or an understanding of ICT tools, individuals will miss out on the benefits of this transformation. The International Trade Administration (2024) highlighted that South Africa is a leader in Africa's digital economy, thanks to its strong infrastructure and supportive government. The sector is growing rapidly, driven by improved connectivity, e-commerce, and emerging technologies like AI and Block chain. By 2025, it is expected for the sector to contribute 15-20% to Gross Domestic Product (GDP), up from 8-10% in 2020 with an annual growth rate of 10-15% over the next five years. As a growing sector with an increasing contribution to GDP, South Africans need to engage with these tools. It must be ensured that no one is excluded due to language barriers or a lack of resources, not only for individual benefit but also for the economy. If more South Africans had access to and the ability to engage with these tools, the GDP would further increase.

In alignment with this, Andrianaivo and Kpodar (2010: 6) highlighted the rising interest in harnessing ICT to enhance financial inclusion in Africa, particularly through mobile financial services. They underscored the necessity of evaluating whether these initiatives can genuinely promote financial inclusion and, in turn, contribute to economic growth. Additionally, Lund and Cruz (2025) stressed the pivotal role of ICT in economic transformation, asserting that the adoption of digital technologies by businesses is a key driver of productivity and economic growth in developing economies. However, despite substantial progress in digitalisation, especially regarding mobile payments, the overall utilisation of digital technologies among African enterprises remains relatively low. The relatively low use of these tools can be attributed to individuals' distrust, lack of access, and difficulties in understanding the digital technologies due to language barriers. As a result, many people are left behind and are not engaging, not because they do not want to, but because of these obstacles.

Lund and Cruz (2025) identified several factors contributing to this gap. First, the cost of hardware and software is disproportionately high, as digital devices, machinery, and software in sub-Saharan Africa are roughly one-third more expensive than in the United States, even before accounting for the lower purchasing power. Second, the limited digital infrastructure continues to hinder digital adoption, with around 600 million Africans lacking access to electricity, alongside a similar number without 4G mobile coverage. Third, internet connectivity costs remain excessively high; for instance, fixed broadband internet in sub-Saharan Africa averages approximately 20% of per capita gross national income, in contrast to less than 6% in other developing regions and just 1% in North America.

In this context, Fadipe (2024) examined the issue of ICT and economic growth through the lens of digital language barriers, arguing that such barriers can obstruct the financial progress of numerous African countries by preventing speakers of indigenous languages from gaining access to job opportunities and online markets. According to Onamadike

(2025), language inclusion in civic tech is frequently overlooked, as design choices are made without acknowledging users' linguistic realities. This gap is exacerbated by insufficient language policies and a lack of government mandates for digital platforms to support indigenous languages. Furthermore, the digital divide has shifted from infrastructure to language barriers, as noted by Primus (2025). This linguistic exclusion limits access to e-commerce, online education, and digital banking, leading to low adoption of digital public services among non-English-speaking communities. This results in a less competitive workforce and isolates entrepreneurs using indigenous languages from broader digital platforms. As AI drives economic growth, Africa risks falling behind due to the neglect of its languages in the digital landscape. This shows the link between digital economic development and language integration, highlighting the significance of language inclusion to ensure increased economic activity.

ICT and Language Integration

Osborn (2006: 86) addressed the digital divide in Africa, emphasising that the continent is often viewed as being on the disadvantaged side of this divide due to low connectivity and access to the internet and computers. While most discussions primarily focus on technical issues like connectivity, Osborn (2006: 87) also drew attention to the linguistic aspect of the divide and its connection to literacy and access. He pointed out that the use of African languages in information and communication technology (ICT) is essential for several reasons: if a language is spoken and utilised elsewhere, it should be compatible with modern technologies. Moreover, African languages are crucial for knowledge generation and expression; however, structured educational support remains minimal. Consequently, relying solely on English, French, and Portuguese places non-speakers at a disadvantage. Using diverse languages in ICT can enhance engagement by making it easier for users to access and understand the knowledge created. Non-English speakers often feel more comfortable participating in their native languages, which fosters a deeper understanding and encourages more significant contributions to ICT tools.

Osborn raised concerns about the future of languages that lack active use in digital spaces, noting that while African languages are present online, they are not prominently used as communication media. Notably, email has emerged as a key application of the internet in Africa, with platforms such as Africast.com and Mailafrica.net facilitating communication in various African languages. Additionally, various email lists exist, often dominated by languages like Hausa and Swahili. However, measuring offline usage such as word processing is more challenging. Anecdotal evidence from telecentres in Senegal, for instance, indicates the use of Pulaar and Wolof. The NGO ANAFA has been conducting computer literacy programs in national languages, while the localisation of software and web interfaces for African languages is gaining traction (Osborn, 2006: 88). This indicates that there are some platforms attempting to utilise these languages in Africa, but their efforts are limited. In South Africa, there is little to no record of platforms using the official languages for ICT communication or other tools. While South Africa may be a leader in the digital transition, it faces the risk of widening the digital divide due to exclusion stemming from a lack of resources and language barriers. Therefore, South Africa needs to prioritise these issues to fully lead and benefit from the digital wave.

Furthermore, Osborn (2006: 89) identified motivational barriers that affect ICT usage. He noted that individuals who have access to ICT are typically educated in official languages used such as English or French, making them less likely to engage in their native languages

online. In contrast, native speakers who are not proficient in official languages often find themselves unable to engage with ICT, even if they desire to do so. Many individuals want to use ICT tools but struggle due to language barriers. As a result, only the elite and those educated in official languages can effectively engage with these tools. This situation is contributing to a significant digital divide. Additionally, he highlighted structural factors that impede the usage of African languages in ICT such as inconsistent orthography standardisation, changes in written forms, and the absence of formal writing systems for some lesser-spoken languages (Osborn, 2006: 90). Similarly, Dia (2013: 113) observed that the potential of African indigenous languages to foster sustainable development has frequently been overlooked. He noted that efforts to advance African languages in cyberspace have encountered various challenges, including infrastructural, financial, and sociocultural barriers, along with a lack of strong political commitment. Nevertheless, there is a growing recognition among stakeholders regarding the developmental potential at the intersection of ICTs and indigenous languages. Furthermore, Primus (2025) points out that in South Africa, colonial and apartheid policies favoured English and Afrikaans, sidelining indigenous languages to informal contexts. Even in the post-apartheid era, English continues to dominate business, academia, and digital communication, perpetuating the marginalisation of many indigenous languages and reinforcing existing socio-economic inequalities.

Furthermore, Dia (2013: 118) highlighted that in South Africa, both civil society and corporate initiatives, have sought to enhance women's access to affordable ICTs. Initiatives include training programs designed to empower women to report gender-based violence through mobile and radio platforms, as well as to search online for artisan information and marketing opportunities. While participation in these initiatives has been enthusiastic, challenges such as illiteracy and language barriers remain significant obstacles to utilising ICT tools aimed at promoting gender equality, livelihood, and education in rural areas. In addition, Dia (2013: 121) pointed out that both state and non-state actors have been implementing educational programs that combine ICT with indigenous languages. An example is the ANAFA project (Alf@net), launched in 1997, which focused on female literacy in Senegal and Mali, incorporating language localisation and ICT training. Its subsequent phases such as the Batik program merged literacy, ICT skills, and vocational training. Notably, the interplay between ICT and indigenous languages has also played a role in political transitions within Sub-Saharan Africa, despite ongoing language barriers that hinder ICT-supported decentralisation efforts.

Moreover, Xulu (2024: 1) underscored the transformative potential of ICT within the educational sector, asserting that it can enhance both teaching and learning in African languages while also serving as a crucial tool for their preservation and promotion. Xulu noted that while educators in other disciplines can source online materials and adapt lessons, those teaching African languages face difficulties due to the underrepresentation of these languages on ICT platforms. In a related update, Fadipe (2024) reported on recent advancements in digital language inclusion efforts by global technology companies. For instance, on October 28, Google announced the incorporation of 15 African languages into its services, including Voice Search, Gboard, and Translate Dictation. The languages added include Chichewa, Hausa, Igbo, Kikuyu, Oromo, Rundi, Shona, Somali, South Ndebele, Seswati, Tigrinya, Twi, Setswana, Nigerian Pidgin, and Yoruba. This milestone is anticipated to allow 300 million additional Africans to engage with the web through voice interaction. However, despite this progress, significant work still lies ahead. Although AI tools are increasingly automating daily tasks, access to these technologies in African

languages remains limited, further exacerbating the digital divide. Across the continent, African start-ups are vigorously working to bridge this gap. For example, in Nigeria, Awarri is developing a multilingual Large Language Model. Meanwhile, Code Vast in Kenya is providing healthcare advice using AI in local languages like Swahili, Meru, Kikuyu, and Luo. In South Africa, there is a project and organisation called "Masakhane", which is an isiZulu word meaning "We build together" and is aimed at attempting to solve the issues of low-resource languages in machine translation (Gitau et. al., 2021: 8).

Williams (2023) contextualised the technological advancements by noting that by the year 2050, a significant portion of the global population, one quarter, will be African. This underscores the urgent need for the inclusion of native languages in all technological innovations. However, despite the rich linguistic and cultural diversity present in Africa, global tech companies have largely neglected these languages. For instance, Google's recent decision to expand Gboard to accommodate nine African languages, which includes eight official languages from South Africa and one from Rwanda reflects a growing acknowledgment of Africa's impending role in the global economy. Nevertheless, Primus (2025) cautioned about the potential threats to African linguistic diversity posed by the rise of artificial intelligence (AI). She states that with over 2,000 languages spoken across the continent, the risk of being excluded from AI tools could perpetuate historical inequalities and deepen the digital divide. In the case of South Africa, colonial and apartheid-era policies promoted English and Afrikaans while marginalising indigenous languages. This trend of exclusion continues to the present day, with English predominantly dominating digital and academic landscapes.

Moreover, educational outcomes improve when children are taught in their mother tongues, yet AI-driven educational platforms frequently overlook this crucial aspect. Echoing this concern, Nkhwashu (2025) highlighted the ramifications of excluding African languages from AI systems. During the ITWeb Artificial Intelligence Summit, Neda Smith articulated the importance of incorporating African languages into AI development, emphasising that language encompasses not merely words but also culture, heritage, and identity. Currently, AI datasets are predominantly comprised of Western languages, thereby marginalising countless African users. This exclusion adversely impacts access to education, communication, and economic development. Additionally, Nkhwashu (2025) identified major challenges that hinder this inclusion such as data scarcity, insufficient funding, inconsistent orthographies, and the lack of AI-compatible data in African languages. The exclusion of African or South African languages in ICT is a significant concern, especially given the increasing integration of ICT into educational, social, and economic sectors. This literature review emphasises the importance of incorporating these languages to enhance engagement and understanding which can ultimately lead to improved financial outcomes.

Methodology

The research study at hand titled, "Bridging the Digital Divide: Exploring the Integration of African Languages in ICT for Economic Growth", is particularly relevant considering the increasing use of digital tools in economic settings. This study is significant as it examines the necessity of incorporating all South African official languages into Information and Communication Technology (ICT) to enhance economic participation and foster growth. Characterised as qualitative research, the study emphasises the collection and interpretation of verbal data rather than numerical statistics (Ugwu and Eze, 2023: 20). The context of the

study is South African; however, some of the literature is drawn from the overall African context. This is due to the limited availability of literature in the South African context.

The qualitative aspect of this research addresses the challenges that language barriers create when it comes to digital tools. It investigates the reasons behind the slow adoption of South African official languages besides English, as it is already being used in ICT, and highlights the importance and advantages of such integration. As an exploratory study, it seeks to answer key questions. Exploratory research is defined as a study aimed at answering a question or addressing a phenomenon (Singh, 2021: 2). The specific questions guiding this research include: What problems are associated with language barriers concerning digital tools? What factors contribute to the sluggish uptake of South African indigenous languages in ICT? Furthermore, what are the benefits of integrating these languages?

The study utilises secondary data, primarily relying on existing literature and information collected through desktop data collection methods, which involve internet searches. According to Ugwu and Eze (2023: 27), secondary research, also known as desk research, uses data that has already been compiled. This approach improves the overall effectiveness of the research by compiling and summarising previously collected information. This desktop study used publicly accessible sources, including journal articles, websites, opinion pieces, and newspaper reports focusing on ICT and language inclusion. Data were collected through targeted keyword searches such as "ICT, language inclusion and economic growth in South Africa," using general search engines (e.g., Google or ResearchGate), institutional repositories, and online media platforms. No single academic database was used exclusively, as the study prioritised thematic breadth and contextual relevance over systematic database coverage. Sources were selected based on their relevance to themes of linguistic accessibility, digital infrastructure, and socio-economic participation. Documents lacking sufficient detail on language use or digital access, or those focused solely on non-African regions were excluded. A thematic analysis was conducted to synthesise insights from journal articles, websites, and newspaper reports. Data were organised under emergent themes such as ICT and the Promotion of Economic Growth, Inconsistent Orthographies Disrupting Language Integration, Lack of Motivation to Integrate or Use Indigenous Languages in ICT, and issues with English being the key language in ICT. This approach enabled a nuanced exploration of recurring patterns and contextual dynamics across African and South African digital landscapes.

However, the study does acknowledge certain limitations. A primary issue is that the reliance on secondary data may restrict the representation of individual experiences and insights, as authentic and direct data collection from those affected by these language barriers would provide more depth. This would have been collected if primary data collection had also been utilised. Additionally, a significant limitation is the scarcity of disaggregated, language-specific data regarding ICT access and the usage of digital services among indigenous language speakers in South Africa. While national broadband statistics are available, they frequently lack granularity concerning linguistic diversity, rural connectivity, and levels of digital literacy. This gap in data limits the ability to conduct meaningful comparative analyses across language groups and undermines the precision of policy recommendations. This study serves as an entry-level exploration into the intersection of digital integration and language inclusion in the South African context, with special attention to implementation and economic development. While these limitations exist, they are acknowledged as inherent to the exploratory nature of the desktop review. Rather than undermining the

study's value, these limitations highlight areas for deeper investigation and provide a foundation for future, more targeted research.

Conceptual and Theoretical Framework

This section provides detailed definitions of significant concepts in the study. The concepts are as follows: Information Communication Technology (ICT), the digital divide, and indigenous languages. It further emphasises two pivotal theories that underpin the study of how incorporating indigenous languages into ICT can foster economic participation and growth. These theories are the linguistic justice theory and the digital inclusion theory. They will be explicated in detail after the conceptual framework discussion.

Conceptual Framework

This study examines the impact of language marginalisation on the digital divide in South Africa with a specific focus on speakers of indigenous languages who face significant obstacles to economic opportunities and essential services, due to their exclusion from Information and Communication Technology (ICT) platforms. Three foundational concepts frame this investigation; namely: ICT, the digital divide, and indigenous languages.

Information Communication Technology (ICT)

Mwiinga (2023: 4) defines ICT as the amalgamation of technology used to manage, process, and share information. This definition encompasses a broad array of technologies, including computers, the internet, mobile phones, and various digital devices. Mwiinga (2023: 4) also highlights that ICT includes the hardware, software, and services crucial for the creation, storage, retrieval, and sharing of information. As a vital component of contemporary society, ICT is essential for individuals, businesses, and organisations to operate effectively and efficiently. Therefore, given the ongoing transition to a digital-centric world, it is pivotal to recognise that ICT plays a crucial role in fostering economic advancement and growth.

Digital Divide

The term "digital divide" describes the inequality in access to ICT, leading to a scenario where certain individuals enjoy availability while others do not. This disparity results in an unequal capacity to utilise and benefit from ICT resources, thereby exacerbating a global knowledge gap between those with adequate resources and those without (Lediga and Fombad, 2018: 298). Hence, the implications of this divide have wide-ranging effects on social equity and economic inclusion.

Indigenous Languages

According to Siziba and Maseko (2024: 4), the South African Constitution recognises twelve (12) official languages; namely, Sepedi, Sesotho, Setswana, Siswati, Tshivenda, Xitsonga, Afrikaans, English, isiNdebele, isiXhosa, isiZulu, and sign language. Finnan (2025) highlights that of all the official languages, nine are indigenous to the region, while Afrikaans and English are of European origin introduced through colonialism. Despite constitutional protections for all languages, indigenous languages frequently face marginalisation and are at risk. Therefore, this study highlights that the predominant language speakers in

the nation are indigenous, underscoring the vital role of ICT not only in preserving these languages, but also in ensuring access and participation in economic activities.

The exclusion of indigenous languages from ICT platforms significantly exacerbates the digital divide by restricting access to essential services such as e-governance, education, and entrepreneurial opportunities. This systemic gap hinders equitable involvement in South Africa's digital economy, ultimately obstructing overall economic growth and development.

Theoretical Framework

Linguistic Justice Theory

As indicated above, this study utilises Linguistic Justice Theory as its main analytical framework to investigate the incorporation of African languages into digital information and communication technologies (ICT) in South Africa. According to Morales-Gálvez and Riera-Gil (2017: 44), rooted in political philosophy, this theory tackles two essential questions: Why is it important to consider languages in discussions of justice? And how should the values that are associated with communication and language identity be fairly distributed within society? Linguistic Justice Theory argues that languages possess both instrumental value, facilitating communication, mobility, and access to socioeconomic opportunities, as well as identity value which upholds cultural dignity, autonomy, and a sense of belonging (Morales-Gálvez and Riera-Gil, 2017: 45). It is also founded on two normative principles: moral individualism, which views individuals as the primary agents whose linguistic choices deserve protection, and ethical pluralism, which asserts that various linguistic identities must be acknowledged and embraced within society's legal and policy frameworks (Morales-Gálvez and Riera-Gil, 2017: 42).

A key concept within this theory is the fair background conditions position, which, according to Lewis (2016: 3), posits that the fundamental requirement for justice in the realm of language is the creation of conditions that allow individuals the opportunity to use their preferred language and to work towards its viability and survival. Providing such opportunities necessitates the elimination of any unjust pressures that discourage certain linguistic behaviours while favouring others. These unjust pressures may include overt coercion, such as bans on using one language and mandates to use another. In South Africa, a constitutionally multilingual country with 12 official languages, the supremacy of English on digital platforms has led to the exclusion of numerous indigenous language speakers from equitable participation in digital and economic spaces. As the nation progressively integrates ICT systems into government services, education, and financial activities, the marginalisation of African languages poses a threat to both communicative equity and cultural preservation.

Thus, Linguistic Justice Theory serves as a strong normative framework to direct inclusive digital design and policy reform. Fostering the integration of African languages into ICT transcends cultural or political aspects; it is a justice-oriented necessity that promotes democratic engagement, socioeconomic mobility, and administrative effectiveness. By acknowledging both the functional and symbolic aspects of language and advocating for equitable background conditions, this theoretical lens empowers South Africa to close the

digital divide in a manner that respects its linguistic diversity and encourages sustainable, inclusive development.

Digital Inclusion Theory

To complement the framework of Linguistic Justice Theory, this research uses Digital Inclusion Theory to explore how language obstacles intersect with digital access and engagement in South Africa's diverse linguistic landscape. Digital inclusion acknowledges that merely having access is not enough, genuine inclusion necessitates that individuals can effectively interact with digital tools, possess the skills to navigate them, and harbour positive attitudes towards their usage (Wiley and Goulding, 2023: 527). A study by Wiley and Goulding (2023: 527) highlights Helsper's framework that identifies four essential elements: use, access, skills, and attitudes, each crucial for comprehending the complex nature of digital equity.

The four pillars are outlined as follows:

- Use refers to digital engagement, incorporating both the nature of engagement, how people utilise technologies, and the extent of engagement – how much time individuals spend on various ICTs.
- Access involves where and how individuals utilise ICTs, including location (where internet access is available), quality (the type of connection such as broadband, wireless, or dial-up), and platforms (the diversity and richness of available media).
- Skills pertain to the ability to utilise ICTs, encompassing transferable skills that enable participation in a digital environment and self-efficacy, which reflects a person's confidence in their capability to use ICTs.
- Attitudes concern perceptions regarding the usefulness and potential risks associated with ICTs, including general attitudes toward ICTs and evaluations of their impact on individuals or society, as well as views on the opportunities and dangers of digital engagement and their significance in everyday life.

In the context of the discourse at hand under the title "Bridging the Digital Divide: Exploring the Integration of African Languages in ICT for Economic Growth", three of these pillars; use, skills, and attitudes are particularly affected by linguistic exclusion. Firstly, use, which denotes digital engagement, involves both how people interact with technology and the amount of time spent online. Indigenous language speakers struggle to fully engage with digital systems when they lack comprehension of the platform's language, hindering participation in education, e-commerce, and e-governance. Secondly, skills encompass both transferable abilities and self-efficacy, or the confidence users have when navigating ICT environments. Language barriers negatively impact both technical skills and users' confidence, resulting in alienation from digital processes. Thirdly, attitudes reflect beliefs regarding the usefulness, risks, and societal relevance of ICTs. If technologies fail to support indigenous languages, users may distrust them or view them as irrelevant, exacerbating the digital divide.

While access (including location, quality of infrastructure, and the variety of platforms) is foundational, this study underscores that linguistic accessibility is equally important. Language exclusion serves as a hidden barrier even when infrastructure exists. Consequently, incorporating South African languages into ICT tools actively promotes digital inclusion by enhancing engagement, improving navigational skills, and cultivating trust in technology.

When the use, skills, and attitudes toward digital tools are inclusive and supportive, they facilitate greater participation, economic empowerment, and collective advancement, particularly among linguistically marginalised communities. Thus, Digital Inclusion Theory offers a dynamic perspective for identifying intersections between language and digital disadvantage, as well as how adaptive ICT policies can foster both equity and development in multilingual settings such as South Africa.

Findings and Discussion

This section presents the key findings regarding the incorporation of South African indigenous languages into ICT to bridge the economic digital divide. By enhancing economic participation, this approach aims to lead to economic growth. This qualitative study sought to explore the reasons for the slow adoption of South African languages in ICT and to propose potential strategies for their integration. Furthermore, the study emphasises challenges related to language translation, noting that many South African languages lack support from widely used translation tools. This issue underscores the necessity for greater involvement from indigenous language speakers in the development and application of these tools. Additionally, the motivation to adopt South African languages is frequently inadequate, as many speakers of indigenous languages may not fully recognise the significance of their languages due to the longstanding dominance of English in global communication. The research indicates that it is crucial to highlight the importance of these languages, enhance their perceived value, employ indigenous language speakers in ICT development, and secure increased investment to achieve successful integration.

The key findings of this study are as follows:

- ICT and the Promotion of Economic Growth

The findings confirm that ICT in South Africa holds substantial potential to foster economic growth. It is believed that ICT services and tools can enhance productivity, improve efficiency, and reduce costs. With the appropriate skills, knowledge, and accessibility, ICT has the potential to generate job opportunities. Moreover, ICT services have played a transformative role in rural areas of Africa, significantly impacting the lives and livelihoods of local communities. These transformations have been observed in various forms, including job creation, income growth, cost savings, reduced uncertainty and risk, and enhanced familial relationships, among other benefits (Adeleye and Eboagu, 2019: 32). In addition, Lund and Cruz (2025) emphasised the critical role of ICT in economic transformation stating that the adoption of digital technologies by businesses is a primary driver of productivity and economic growth in developing economies. Although the study acknowledges this potential, it emphasises that due to digital language barriers, the promise of promoting economic growth appears distant. Fadipe (2024) examined the interaction between ICT and economic development through the lens of digital language barriers, arguing that these barriers can hinder the economic advancement of many African countries by denying indigenous language speakers access to job opportunities and online markets.

The study believes that this situation arises because a significant portion of the population in South Africa is indigenous speakers; thus, the ICT tools and services geared toward economic activities often exclude them. Most aspects are now conducted through online markets or e-governance; therefore, the South African government must invest more in ICT to boost economic development and overall growth. From the perspective of digital

inclusion theory, ICT and the promotion of economic growth rely on four pillars: use, access, skills, and attitudes. In a sense, ICT can foster economic growth in South Africa when more people have the necessary skills, attitudes, and access which also includes understanding the language used in these tools. Individuals will not be motivated to use tools they do not understand out of fear of embarrassment; likewise, one may know how to operate ICT tools, but without understanding the language, those skills are useless. Therefore, for individuals to effectively use ICT tools for the right reasons such as seeking job opportunities, advertising businesses, or accessing e-governance, they need to be motivated, possess the right skills, and have enough access.

- Inconsistent Orthographies Disrupting Language Integration

The ICT tools that are currently available do not adequately accommodate African or South African indigenous languages, and those few tools that do still struggle to represent these languages accurately. For instance, tools like Google Translate are not effective when it comes to indigenous languages. This situation highlights the potential requirement for indigenous language speakers to participate in the integration process of these languages. If the ICT tools that utilise indigenous languages do not accurately represent them, then their integration would be meaningless. Conversely, if integrated successfully, the benefits would extend beyond economic factors to include the preservation and promotion of these languages. Xulu (2024: 1) stressed the transformative potential of ICT within the educational sector, asserting that it can improve both teaching and learning in African languages, while also serving as a vital tool for their preservation and promotion. Moreover, Nkhwashu (2025) identified significant challenges hindering this inclusion, including data scarcity, insufficient funding, inconsistent orthographies, and the lack of AI-compatible data for African languages.

Incorporating indigenous languages into ICT is essential, but it must be done correctly. This means hiring individuals who are fully knowledgeable in these languages rather than relying on inadequate digital tools. Employing South African indigenous language speakers will not only facilitate successful integration but also create job opportunities for those individuals. The integration of these languages aligns with the principles of linguistic justice theory, which advocates for conditions that allow individuals to use their preferred languages, supporting their viability and survival. By promoting and sustaining these languages, many people are enabled to engage with economic platforms in their home languages. It is unjust for the digital realm to feature inconsistent orthographies or lack AI-compatible data for indigenous languages, particularly in South Africa. The language policy was established to rectify historical linguistic injustices and to promote African languages across various domains (Siziba and Maseko, 2024: 4). This situation contradicts the core objectives of the language policy. To uphold this policy, the South African government should advocate for greater integration of indigenous languages into ICT tools.

- Lack of Motivation to Integrate or Use Indigenous Languages in ICT

The study reveals that the failure to utilise or incorporate indigenous languages into ICT tools and services sometimes stems from a lack of motivation. As English has been the predominant and official language in South Africa, indigenous languages have been overlooked, especially in the current digital era. Privileged speakers of indigenous languages often place greater value on English than on their languages, resulting in demotivation when it comes to engaging with ICT tools and services in their native tongues. This lack of motivation consequently affects the motivation of creators of ICT tools and service

providers, leading them to be less inclined to incorporate indigenous languages into their offerings. Osborn (2006: 89) identified motivational barriers influencing ICT usage, noting that individuals who access ICT tools generally receive their education in official languages such as English, making them less likely to interact with online content in their native languages. The educational, economic, and social sectors all utilise English, and ICT is increasingly adopting English in these fields as well. As a result, individuals feel demotivated to use their native languages in these contexts and often do not question why their languages are not included in these tools. They have simply become accustomed to this situation. Moreover, what is the point of feeling motivated when all the available tools exclude native languages? These languages are often left out in translation, interpretation, and communication provided by these tools.

- Issues with English being the Key Language in ICT

ICT tools and services predominantly utilise English as their primary language. The study recognises that not everyone is capable of understanding or engaging effectively with these ICT tools and services when English is the chief language. This situation arises primarily because, in many African nations, English is regarded as a second language and not all individuals possess proficiency in it. According to Primus (2025), particularly in South Africa, policies from the colonial and apartheid periods favoured English and Afrikaans while simultaneously marginalising indigenous languages. Moreover, this trend of exclusion persists today as English continues to dominate both digital and academic spheres. Consequently, without the incorporation of linguistic diversity in AI tools, millions of individuals find themselves unable to access essential services, which include e-commerce, online education, or digital banking.

- The findings in connection with the theories that frame the study

The two central theories that frame the study are Linguistic Justice and Digital Inclusion. One key concept within linguistic theory is the fair background conditions position. According to Lewis (2016: 3), this position maintains that a fundamental requirement for justice within the field of language is to establish conditions that enable individuals to use their preferred language while also working toward its viability and survival. In addition, Digital Inclusion recognises that simply having access to digital resources is insufficient; true inclusion demands that individuals can not only interact effectively with these digital tools but also possess the requisite skills to navigate them and develop positive attitudes toward their use (Wiley and Goulding, 2023: 527). Thus, the study argues that the reliance on English as the primary language in ICT tools and services contributes significantly to both a digital divide and a situation of linguistic injustice. This is primarily because it confines individuals to using English, a language that many indigenous language speakers do not understand. Therefore, this limitation hinders their ability to interact effectively with digital tools and comprehend the language used to navigate these services. Consequently, language barriers severely impact individuals' abilities to thrive and leverage digital tools for their benefit, particularly in today's digital economy.

Conclusion

The study investigated the reasons behind the slow integration of African languages into ICT and proposed potential strategies for their inclusion. It emphasised the significance of incorporating African indigenous languages for fostering economic development. The study underscored the need to involve speakers of these languages in the integration

process to prevent inconsistencies in orthography, which can hinder effective and efficient incorporation. It also pointed out the lack of motivation to use indigenous languages, which results in disheartened creators of ICT tools and services when trying to incorporate these languages. Additionally, the dominance of English as the primary language used in ICT tools is highlighted as a challenge since it excludes those who do not comprehend English. This lack of understanding prevents individuals from fully engaging with the tools and services, even if they have access to digital technologies. Besides other obstacles such as infrastructure, devices, and costly data for using ICT tools and services, a significant yet often overlooked barrier is the impact of language. Therefore, it is crucial to recognise that language is fundamental and should be a primary focus when addressing barriers to digital access.

The study is grounded in two theoretical frameworks: linguistic justice and digital inclusion. Through the examination of these theories, the study argues that prioritising English in ICT tools while disregarding indigenous languages perpetuates linguistic injustice and contributes to a digital divide, thereby hindering economic involvement and growth. Consequently, the study contends that indigenous languages should be integrated into ICT tools to facilitate economic participation, ultimately driving economic development. It acknowledges that this integration will be challenging and time-consuming, but it asserts that it is feasible, given sufficient interest and emphasis on its importance.

For successful incorporation, the study recommends several actions. First, it is essential to recognise the value and significance of indigenous languages and their inclusion in ICT tools and services, not only for economic advancement but also for their preservation and promotion. This recognition would motivate users to engage with these languages in their interactions with ICT tools, consequently encouraging creators to incorporate them into their services. Second, it is advisable to employ speakers of indigenous languages during the integration process to avoid inconsistencies in orthography and to ensure proper representation of these languages. Lastly, the government should increase investment in ICT tools and services, as well as in the integration of indigenous languages into them. This investment would enhance the promotion of ICT tools, and incorporating indigenous languages would boost economic participation, leading to economic growth. Thus, given their substantial importance, indigenous languages must be acknowledged in the transition to the digital era, ensuring their inclusion.

References

Adeleye, N., & Eboagu, C. (2019). Evaluation of ICT development and economic growth in Africa. *NETNOMICS: Economic Research and Electronic Networking*, 20(1), 31–53. https://doi.org/10.1007/s11066-019-09131-6.

Andrianaivo, M., & Kpodar, K. (2010). *ICT, financial inclusion, and growth: Evidence from African countries (Draft)*. African Economic Conference Proceedings. United Nations Economic Commission for Africa. Available at: https://archive.uneca.org/sites/default/files/uploaded-documents/AEC/2010/Papers/session_i.2.1_1._ict_financial_inclusion_and_growth.pdf. Accessed: 05 July 2025.

Dia, I. A. (2013). 'African languages, and information communication technologies and development'. *International Journal of the Sociology of Language*, 2014(225), 113–130. https://doi.org/10.1515/ijsl-2013-0068.

Fadipe, A. (2024, November 19). 'Why tech giants must do more for African language inclusion'. *Global Voices Lingua*. Available at: https://lingua.globalvoices.org/2024/11/why-tech-giants-must-do-more-for-african-language-inclusion/. Accessed: 05 July 2025.

Finnan, A. (2025, July). 'South Africa: Rediscovering Indigenous languages'. *Aswica.co.za*. Available at: https://aswica.co.za/south-africa-rediscovering-indigenous-languages/. Accessed: 08 July 2025.

FoodBev SETA. (2022). *Bridging the digital literacy divide in rural and township communities*. Available at: https://foodbev.co.za/wp-content/uploads/2023/04/Bridging-the-Digital-Literacy-Divide.pdf. Accessed: 05 July 2025.

Gitau, C., Kabongo, S., Modupe, A., Marivate, V., Sindane, T. A., Lastrucci, R. and Dzingirai, I.J. (2021). *Masakhane Web: A Machine Translation Platform for African Languages*. University of Pretoria, South Africa.

International Trade Administration. (2024, September 19). *South Africa - Digital Economy*. U.S. Department of Commerce. Available at: https://www.trade.gov/country-commercial-guides/south-africa-digital-economy. Accessed: 21 October 2025.

Lediga, M.M. & Fombad, M.C. (2018). 'The use of information and communication technologies in public libraries in South Africa as tools for bridging the digital divide: the case of the Kempton Park public library'. *Public Library Quarterly*, 37(3), 296-305. https://doi.org/10.1080/01616846.2018.1471964

Lewis, D. H. (2017). 'Realising linguistic justice: Resources versus capabilities'. *Journal of Multilingual and Multicultural Development*, 38(7), 595-606. https://doi.org/10.1080/01434632.2016.1192178

Lund, S., & Cruz, M. (2025, April 29). 'Bridging the digital divide in Africa: Enhancing technology adoption for economic growth'. *Brookings*. Available at: https://www.brookings.edu/articles/bridging-the-digital-divide-in-africa-enhancing-technology-adoption-for-economic-growth/. Accessed: 05 July 2025.

Mahlaule, M., Marwa, N., & Akinrinde, O. (2024). ‚Constraints, challenges, and opportunities for digital business in rural and township economy'. *Indonesian Journal of Business and Entrepreneurship*, 10(3), 685-696.

Morales-Gálvez, S., & Riera-Gil, E. (2021). ‚What are just language policies? The current paradigms of linguistic justice'. *Catalan Social Sciences Review*, 11, 39-61. https://doi.org/10.2436/20.3000.02.56.

Mwiinga, P. (2023). *Introduction to Information Communication Technology*. Preston Mwiinga Enterprise. Available at: https://www.researchgate.net/publication/371986689_introduction_to_information_communication_technology. Accessed: 08 July 2025.

Nkhwashu, N. (2025). 'Bridging the digital divide: Why African languages must be included in AI'. *ITWeb*. Available at: https://www.itweb.co.za/article/bridging-the-digital-divide-why-african-languages-must-be-included-in-ai/lwrKxv3YkKBMmg10. Accessed: 08 July 2025.

Ohamadike, N. (2025, July 6). Digital technology must speak African languages. *Mail & Guardian.* https://mg.co.za/thought-leader/opinion/2025-07-06-digital-technology-must-speak-african-languages/. Accessed: 08 July 2025.

Osborn, D. Z. (2006). *African languages and information and communication technologies: Literacy, access, and the future*. In J. Mugane, J. Hutchison, & D. Worman (Eds.), *Selected proceedings of the 35th annual conference on African linguistics (pp. 86–93)*. Somerville, MA: Cascadilla Proceedings Project. https://www.lingref.com/cpp/acal/35/paper1299.pdf.

Primus, M. (2025). 'Breaking the digital divide: Why Africa's future depends on embracing its languages in AI'. *IOL Technology Opinion*. Available at: https://iol.co.za/technology/opinion/2025-02-11-breaking-the-digital-divide-why-africas-future-depends-on-embracing-its-languages-in-ai/. Accessed: 05 July 2025.

Saba, C. S., Ngepah, N., & Odhiambo, N. M. (2025). 'The nexus between ICT diffusion, financial development, industrialization and economic growth: Evidence from Sub-Saharan African countries'. *Journal of African Business*, 26(2), 454–478. https://doi.org/10.1080/15228916.2025.2451550.

Singh, A. (2021). 'An introduction to experimental and exploratory research'. *Social Science Research Network (SSRN)*. https://doi.org/10.2139/ssrn.3789360.

Siziba, L. and Maseko, B. (2024). 'Enabling the Economics of African Languages in Language Policies of South Africa and Zimbabwe'. *African Journal of Inter/Multidisciplinary Studies*, 6(1), 1–11. https://doi.org/10.51415/ajims.v6i1.1351

Ugwu, C. N., & Eze, V. H. U. (2023). 'Qualitative research'. *IDOSR Journal of Computer and Applied Sciences*, 8(1), 20–35.

Wiley, K., & Goulding, A. (2023). *The conceptualization of digital inclusion in government policy: A qualitative content analysis*. In I. Sserwanga, L. E. Freund, C. C. Marshall, M. H. Lim, & A. Grove (Eds.), *Information for a better world: Normality, virtuality, physicality, inclusivity*. iConference 2023 (Lecture Notes in Computer Science, Vol. 13972, pp. 527–543). Springer. https://doi.org/10.1007/978-3-031-28032-0_40.

Williams, C. (2023). 'African language inclusion in technology good for growth'. *IOL*. Available at: https://www.iol.co.za/lifestyle/opinion/african-language-inclusion-in-technology-good-for-growth-3311e8d6-b8da-4e53-8515-e83d65730772. Accessed: 08 July 2025.

Xulu, Z. C. (2024). 'Integrating technology in teaching African languages in South African universities: A call for digitalisation'. *Journal of the Digital Humanities Association of Southern Africa*, 5(1),1–14. https://doi.org/10.55492/dhasa.v5i1.5030.

The Use of Vernacular on Social Media for Citizen-Based Monitoring of Municipal Service Delivery in South Africa

Lesedi Senamele Matlala
Department of Public Management and Governance
University of Johannesburg
Johannesburg, South Africa
lesedi.matlala@uj.ac.za

Abstract

In South Africa, citizens are increasingly using social media to monitor and report failures in municipal service delivery. While Citizen-Based Monitoring (CBM) is gaining traction in public sector reforms, its application through social media remains informal, fragmented, and largely unrecognised by municipal authorities. A notable feature of these practices is the use of African indigenous languages, allowing citizens to articulate grievances authentically within their cultural contexts. Guided by Digital Public Sphere theories and the Uses and Gratifications Theory, this study examines how citizens employ digital platforms – particularly through vernacular expression – to participate in public deliberation, seek accountability and fulfil information, expressive, and community-oriented needs. Using a qualitative design, the research combines content analysis of vernacular posts from municipalities with strong social media presences and in-depth interviews with government officials, community-based organisations (CBOs), and community radio representatives. The findings show that indigenous language use amplifies marginalised voices, enhances accessibility, and fosters communal legitimacy around service delivery grievances. However, municipalities often lack the linguistic capacity, digital tools, and institutional frameworks to monitor or respond effectively to vernacular complaints. Meanwhile, community radio and CBOs play intermediary roles without formal support. The study contributes to digital governance, language justice, and participatory democracy scholarship by proposing a framework to formalise vernacular CBM. Key recommendations include developing multilingual digital engagement strategies, formally partnering with community media, and investing in African-language natural language processing (NLP) tools. Recognising vernacular citizen monitoring practices offers a pathway to more inclusive, responsive, and linguistically representative local governance aligned with South Africa's democratic ideals.

Keywords: Citizen-Based Monitoring (CBM), Digital Public Sphere, Vernacular Counterpublics, African Indigenous Languages, Digital Governance, Language Justice, Service Delivery, Participatory Democracy.

Introduction

The persistent failures in municipal service delivery remain among the most critical governance challenges in post-apartheid South Africa. Across urban and rural spaces, communities experience recurring issues such as water shortages, electricity cuts, waste management breakdowns, and deteriorating infrastructure (Mudzusi et al., 2024). Studies such as Breakfast et al. (2019) and Mamokhere (2020) link the escalation of violent service delivery protests to citizens' frustrations with ineffective local governance. This crisis has triggered increasing reliance on alternative accountability mechanisms outside formal government structures as traditional participatory forums – like ward committees and imbizos – often prove insufficient (Friedman, 2006; Matlala, 2025). In this context, digital technologies, particularly social media, have emerged as critical platforms for citizens to express grievances, share experiences, and mobilise collective action. These developments represent a significant shift in how civic participation and governance oversight are conducted with profound implications for the future of local democracy.

The rise of social media as a citizen-driven monitoring tool has been well-documented both globally and locally. Bonsón et al. (2015; 2017) show that municipal social media accounts in Western Europe increasingly serve as channels for information dissemination, citizen feedback and complaint escalation. Similarly, Gu et al. (2020) illustrate how Chinese municipalities leverage WeChat and Weibo to manage citizen relations and service reporting. South African studies reveal a more citizen-initiated dynamic; while municipalities maintain an online presence, it is often citizens themselves who drive conversations around accountability (Fashoro & Barnard, 2017; Okeke-Uzodike & Dlamini, 2019). Research by Fashoro and Barnard (2021) further emphasises the limited effectiveness of government-driven social media initiatives for genuine citizen participation, echoing Arshad and Khurram's (2020) findings that transparency and responsiveness are crucial for genuine online engagement. The global trend of evolving from traditional e-government to "social network government" (Halpern & Katz, 2012) underscores the need for a more responsive, participatory digital state, particularly in contexts where conventional governance structures are weakened.

For the purposes of this study, social media is defined as interactive digital platforms that enable users to generate, share, and exchange content in real time. The analysis focuses on three platforms most relevant in the South African context: X (formerly Twitter), which facilitates real-time public debates; Facebook, which provides communal forums for discussion and sharing of grievances; and WhatsApp, which supports more localised, community-driven exchanges through group messaging. These platforms represent distinct but overlapping communicative spaces where citizens employ African indigenous languages to document service delivery failures and demand accountability.

A crucial but underexplored dimension of digital civic participation is the role of African indigenous languages in shaping citizen engagement. Haro-de-Rosario et al. (2018) argue that content relevance and accessibility are key determinants of citizen engagement rates on platforms like X and Facebook. Yet, studies such as Ellison and Hardey (2014) and DePaula et al. (2018) have focused primarily on Western or monolingual contexts, neglecting multilingual societies like South Africa. Vernacular language use on social media enables authenticity, cultural resonance, and emotional immediacy which are critical factors for building trust and collective identity online (DePaula & Dincelli, 2018). In South Africa, where over 80% of the population speaks an indigenous language at home (We Are Social & Hootsuite, 2024; Worldwideworx & OrnicoGroup, 2024), linguistic accessibility is not

merely a matter of inclusion but a democratic imperative. Nevertheless, most municipal digital strategies remain overwhelmingly English-centric, thus marginalising the vernacular voices that dominate informal citizen-driven CBM activities.

Despite the proliferation of citizen grievances in African languages on social media, their integration into formal governance systems remains fragmented and ad hoc. Studies by Eom et al. (2018) and DePaula and Dincelli (2018) highlight that governments that fail to adapt their digital strategies to accommodate citizen-led, affectively charged communications risk widening the trust gap between themselves and their constituents. In South Africa, local governments often lack the institutional mechanisms, linguistic competencies, and technological tools to systematically monitor or respond to vernacular complaints (Fashoro & Barnard, 2017; Mudzusi et al., 2024). Community-based organisations (CBOs) and community radio stations frequently act as informal intermediaries, translating citizen concerns into formal channels, yet they operate without structured municipal support (Okeke-Uzodike & Dlamini, 2019; Mamokhere, 2020). This disjuncture underscores a missed opportunity - vernacular digital expressions, if formally recognised and incorporated into CBM systems could substantially enhance the inclusivity, responsiveness, and legitimacy of municipal governance.

This study aims to critically examine how citizens use African indigenous languages on social media platforms to monitor, report, and demand accountability for municipal service delivery failures. Specifically, it investigates how vernacular-language digital expressions serve as informal mechanisms of CBM, how intermediaries like CBOs and community radios engage with these expressions, and what institutional innovations are necessary to formalise and integrate this citizen-generated feedback into governance frameworks. The study addresses these questions and contributes to digital governance, participatory democracy, and language justice. The paper proceeds as follows: Section 2 reviews the relevant literature on CBM, digital participation, and language rights; Section 3 outlines the theoretical frameworks guiding the analysis; Section 4 presents the research design and methodology; Section 5 discusses the empirical findings; Section 6 offers a critical discussion; Section 7 proposes policy and practical recommendations; and Section 8 concludes with reflections on future research directions.

Literature Review

Citizen-Based Monitoring (CBM)

Citizen-Based Monitoring (CBM) originates from participatory development traditions and the rise of citizen science, both of which challenged the exclusive authority of experts in assessing public problems. Danielsen et al. (2005) highlight how locally-based monitoring empowered communities to produce knowledge influencing policy decisions, particularly in natural resource governance. Conrad and Daoust (2008) similarly demonstrate that community-based frameworks enhanced environmental stewardship by elevating local knowledge alongside scientific expertise. Dickinson et al. (2012) observe the expanding role of non-experts in systematic data collection for academic and policy influence, while Carlson and Cohen (2018) stress that successful CBM hinges on linking citizen data with policy responsiveness. Kouril et al. (2015) underscore CBM's relevance in marginalised regions, an insight echoed by Wiseman and Bardsley (2016) in their study on indigenous environmental monitoring. Lewandowski et al. (2017) note a growing recognition of citizen science's value in democratising knowledge production. In South Africa, CBM gained formal

policy recognition post-apartheid as part of deepening participatory democracy efforts (DPME, 2013; 2015; 2016). Initiatives like the Black Sash Trust's Community Monitoring and Advocacy Programme (DPME, 2013) demonstrated CBM's frontline oversight potential, with Matlala (2024a) identifying its integration of citizen feedback loops into health, education, and municipal service monitoring as a major governance innovation. Yet, challenges around sustainability, scalability, and institutional integration remain persistent (Bester, 2015), reflecting broader global shifts towards citizen-driven accountability while highlighting enduring systemic limitations.

Despite these advances, a major gap in CBM literature, especially in South Africa, lies in insufficient engagement with how digital technologies, particularly social media, reshape citizen monitoring practices. Traditional CBM models – such as structured surveys and community forums (Matlala, 2024b) – are increasingly supplemented by informal, real-time expressions on digital platforms, often articulated in African indigenous languages. While DPME (2016) and Matlala (2024a) acknowledge the potential of ICTs to strengthen CBM, they overlook the vernacular and affective dimensions of digital citizen feedback. In contrast, environmental monitoring globally (Danielsen et al., 2005; Conrad & Daoust, 2008) has more rapidly integrated mobile technologies, highlighting South Africa's relative lag in leveraging digital tools for inclusive CBM. Existing public service monitoring remains predominantly English-centric and institutionally siloed, failing to harness the democratic potential of vernacular digital activism fully. This study addresses that gap by critically examining African-language citizen expressions on social media as an emerging yet under-theorised mode of CBM, raising crucial questions about linguistic justice, informal civic agency, and the future of digital governance in South Africa's democratic landscape.

Social Media and Civic Accountability: Informal Complaint Mechanisms and Digital Activism

The rise of social media platforms has reconfigured traditional civic engagement channels by enabling informal, real-time mechanisms for voicing grievances and demanding public accountability. Scholars such as Kreiss et al. (2020) argue that digital platforms have expanded the "participatory repertoires" available to citizens, allowing them to bypass institutional barriers and engage directly with service providers and authorities. In contexts characterised by governance deficits, social media becomes a communication tool and an alternative monitoring system where ordinary citizens document, publicise, and pressure governments to address service delivery failures (Earl et al., 2022). Tufekci (2020) emphasises that such informal digital activism thrives in environments where formal participatory structures are weak or inaccessible, highlighting social media's role in shaping bottom-up accountability movements. Recent empirical work by Surya et al. (2023) shows that citizen complaints posted on platforms like Facebook and X often lead to faster government responses compared to formal reporting mechanisms, especially in municipal governance contexts. However, Rumbul (2020) cautions that the efficacy of digital complaint-making is contingent on governments' capacity and willingness to institutionalise social media monitoring into their service delivery systems. Without structural integration, citizen-generated digital grievances risk being relegated to symbolic acts rather than substantive participatory governance tools (Eom et al., 2021). These developments position social media as a key, albeit contested, site for informal CBM practices.

In South Africa, these dynamics are reinforced by the rapid rise of smartphone penetration and social media use over the past decade. According to ICASA (2024), more than 80% of

South Africans out of an estimated population of 63 million now own or have access to a smartphone, reflecting the country's growing digital inclusion and expanding opportunities for mobile-based civic engagement. Moreover, in South Africa, WhatsApp is the dominant social media and messaging platform: recent data show that roughly 93.8% of active social media users in the country report using WhatsApp (Statista, 2024). Furthermore, estimates suggest there are approximately 28–29 million South African WhatsApp users (Statista, 2024). Among adults aged 18–34, social media penetration exceeds 85%, compared to around 55% among those aged 35 and older, underscoring generational divides in digital activism. These trends confirm that social media has become a central arena for everyday communication and collective action, making it an indispensable site for analysing vernacular CBM in the South African context.

The evolution of digital activism further complicates traditional understandings of civic participation by foregrounding the interplay between language, identity, and power in online spaces. Studies by Maharaj and Pillay (2022) in the South African context reveal that digital protests against poor service delivery often unfold in indigenous languages, amplifying localised experiences of marginalisation and exclusion. Similarly, Ingram and Licona (2023) demonstrate that the vernacularisation of digital spaces fosters community solidarity while simultaneously challenging dominant narratives framed in official or elite languages. Despite these promising trends, Boulianne et al. (2022) argue that digital activism remains uneven across demographic lines with access to technologies, digital literacy, and language barriers influencing who participates and how effectively they are heard. Consequently, without deliberate municipal strategies to engage vernacular-language digital activism, informal complaint-making may reinforce existing exclusions rather than democratise governance. Jamil and Awal (2021) emphasise that municipalities must transition from passive monitoring to the active incorporation of citizen-generated data into planning and decision-making processes, if social media is to serve as a genuine accountability mechanism. Thus, recognising the complex, linguistically diverse, and informal nature of digital CBM practices becomes critical for building more inclusive and responsive local governance systems in South Africa and beyond.

Indigenous Languages in the South African Context

South Africa is constitutionally recognised as a multilingual democracy with 12 official languages: isiZulu, isiXhosa, Afrikaans, Sepedi, Setswana, Sesotho, Xitsonga, SiSwati, Tshivenda, isiNdebele, English, and South African Sign Language (Republic of South Africa, 1996). While this framework affirms linguistic diversity and equality, in practice, English continues to dominate formal governance, policy communication, and digital platforms (Alexander, 2019). Indigenous African languages, though widely spoken in households and community spaces, are frequently marginalised in institutional contexts, particularly in local government service delivery communication. This mismatch creates systemic barriers to inclusive participation as large segments of the population are unable to express grievances or interact with authorities in the languages most meaningful to them.

In this study, the term "vernacular" is used to refer specifically to African indigenous languages such as isiZulu, isiXhosa, Sepedi, Setswana, and others recognised within the official language framework. It does not refer to slang, colloquialisms, or informal hybrid expressions, although these sometimes appear in online interactions. By focusing on indigenous languages, the analysis emphasises linguistic justice as a central component of democratic participation, highlighting how citizens employ their home languages to

articulate service delivery failures and assert their right to recognition in governance processes. As scholars such as Banda and Mwanza (2023) and Kaya et al. (2016) note, the inclusion of indigenous languages in digital and civic spaces is critical for dismantling entrenched linguistic hierarchies rooted in colonial legacies. Clarifying the distinction between indigenous languages and slang ensures conceptual precision and aligns the study with broader debates on language policy, multilingualism, and digital participation in South Africa.

African Indigenous Languages and Language Justice in Digital Spaces

The marginalisation of African indigenous languages in the digital sphere remains a critical obstacle to inclusive digital citizenship and participatory governance. Mpofu and Salawu (2020) demonstrate how even well-intentioned localisation efforts such as Google's vernacular webpages often fail to accommodate African users' sociolinguistic diversity and functionality needs fully. Similarly, Jongbloed-Faber et al. (2016) show that language use on social media among bilingual communities often reflects deeper socio-political hierarchies, with dominant languages crowding out indigenous tongues. Ragnedda (2018) frames this phenomenon as part of broader digital inequalities where linguistic exclusion compounds social exclusion in the digital age. In this context, Aiyegbusi (2018) calls for decolonising digital spaces to ensure that African epistemologies, including indigenous languages, are embedded within digital infrastructures rather than merely appended as superficial inclusions. Chonka et al. (2022) further highlight the structural biases of algorithmic systems like search engine autocompletes, which systematically privilege dominant global languages over African ones, reinforcing linguistic hierarchies online. This linguistic digital divide undermines the capacity of African citizens to participate fully in digital public spheres, thereby excluding vernacular voices from policy debates, civic activism, and governance monitoring processes. Therefore, addressing language justice is a cultural or educational imperative and central to building equitable and participatory digital governance systems.

In civic activism and citizen monitoring, African indigenous languages have become potent tools for asserting identity, solidarity, and political agency in digital spaces. Nyabola (2018a) discusses the #FreeBobiWine movement as an example of Pan-African digital activism where vernacular expressions were crucial in mobilising solidarity across linguistic and national boundaries. Hernandez and Roberts (2018) similarly emphasise that ensuring linguistic inclusion is vital for "leaving no one behind" in digital development, arguing that marginalised groups often rely on local languages for digital engagement. Yet, the architecture of most digital platforms remains poorly aligned with these realities, usually privileging English and other dominant languages in their user interfaces, algorithms, and content moderation practices (Chonka et al., 2022). Without deliberate interventions to promote African languages online, digital spaces risk replicating offline inequalities and silencing the populations that participatory governance seeks to empower. Consequently, in contexts like South Africa, formalising vernacular digital expressions within CBM frameworks becomes an act of inclusion and a necessary step toward democratising digital governance and ensuring that indigenous linguistic identities are not erased in the emerging digital public sphere.

Theoretical Framework: Digital Public Sphere Theories & Uses and Gratifications Theory

This study is anchored in Digital Public Sphere theories, particularly the strands of networked publics, affective publics, and counter publics to critically examine how communication technologies reshape democratic participation, voice, and accountability. While Habermas' (1962) original conception of the bourgeois public sphere emphasised rational-critical debate in physical spaces, digital-era theorists highlight the fragmentation of that sphere into multiple, overlapping publics shaped by the affordances of online platforms. Papacharissi (2015), for example, argues that contemporary networked environments produce "affective publics" where emotion is a central driver of engagement and mobilisation. This is especially relevant for vernacular CBM; posts written in African indigenous languages often embed culturally resonant idioms, metaphors, and emotional tones that differ from traditional deliberative norms but nevertheless embody authentic political expression. Fraser's (1990) critique of Habermas provides further grounding by showing how marginalised groups construct "subaltern counterpublics" to resist exclusion from dominant spaces. In the South African context, citizens' use of indigenous languages on X functions as a vernacular counterpublic, directly contesting the dominance of English in both municipal governance and digital communication. These theories together allow the study to interpret vernacular complaints not as disorderly or informal, but as structured forms of civic action within a pluralistic digital democracy.

By explicitly aligning the study with affective publics, networked publics, and counterpublics, the theoretical framing highlights the relevance of vernacular CBM in shaping alternative civic spaces and strengthening democratic accountability. Dahlgren (2005) stresses that the quality of democratic participation online depends not just on access to platforms but on whether diverse publics can shape discourse meaningfully - a condition directly linked to linguistic inclusion. Dean (2003) similarly notes that digital publics are often characterised by "affective intensities", meaning that emotional and cultural registers carry political weight in shaping collective action. Vernacular CBM reflects this dynamic by translating service delivery grievances into affectively charged expressions that mobilise solidarity and pressure government actors. Thus, rather than treating indigenous-language complaints as noise or informal protest, this study frames them as legitimate contributions to South Africa's multilingual democratic project. Positioning vernacular digital expressions within the Digital Public Sphere literature enables the analysis to foreground linguistic justice, citizen agency, and participatory accountability as interdependent dimensions of digital governance in a postcolonial, multilingual society.

In addition to Digital Public Sphere theories, this study also draws on the Uses and Gratifications Theory (UGT) to explain citizen behaviour in digital spaces. UGT posits that individuals actively select media platforms to satisfy specific needs such as information seeking, social interaction, identity expression, and problem-solving (Katz et al., 1974). Applying this lens helps illuminate why South African citizens use platforms like X to broadcast grievances publicly, Facebook to share experiences within community networks, and WhatsApp to circulate localised complaints in closed group settings. These platforms meet different communicative needs: X provides visibility and immediacy, Facebook offers communal validation and discussion, while WhatsApp enables rapid coordination and vernacular intimacy. By combining Digital Public Sphere theories with UGT, the study situates vernacular citizen monitoring as both a collective political practice and an individualised choice shaped by the gratifications citizens seek from specific digital platforms.

Methodological Design and Research Strategy

This study adopts a qualitative, interpretive, and exploratory research design to critically examine how citizens use African indigenous languages on social media platforms to monitor municipal service delivery in South Africa. A qualitative approach is appropriate given the study's focus on meaning-making, linguistic expression, and contextual nuance rather than quantifiable metrics (Denzin & Lincoln, 2018). Interpretivism underpins the methodological orientation, recognising that citizens' vernacular complaints are socially constructed narratives shaped by cultural, political, and historical contexts (Schwandt, 2015). Given the emergent and under-theorised nature of vernacular digital citizen-based monitoring, an exploratory design was selected for inductive insights rather than hypothesis testing (Stebbins, 2001). The study seeks to uncover how indigenous language use on social media functions not merely as communication but as a participatory accountability practice within an evolving digital public sphere. This flexible design enables deep engagement with the lived experiences, perceptions, and digital behaviours of both citizens and intermediaries such as community-based organisations and community radio stations. By prioritising participants' voices and the socio-cultural textures embedded in their linguistic choices, the research design aligns methodologically with the study's broader commitment to inclusivity, linguistic justice, and democratic deepening.

Content analysis

The primary data source for this study's content analysis comprised vernacular-language posts on X from 2020 to 2025 (see Table 1). X was selected due to its public accessibility and the feasibility of obtaining ethical approval for research involving publicly available data (Townsend & Wallace, 2016). The platform allows researchers to access user-generated content without violating privacy expectations, provided appropriate ethical safeguards are in place such as anonymisation of user handles and paraphrasing of direct quotations to prevent traceability (Williams et al., 2017). The content analysis focused on posts directed at or mentioning official municipal accounts, hashtags associated with service delivery complaints (e.g., #NoWater, #Loadshedding, #FixOurRoads), and posts geotagged within selected South African municipalities with strong digital engagement records.

Purposive sampling was used to select municipalities with high levels of social media activity, diverse linguistic representation, and strong patterns of citizen engagement. This ensured that the dataset reflected a wide variety of vernacular expressions across urban and provincial contexts. Unlike random or systematic sampling, purposive selection was deemed more appropriate for the study's qualitative aims, which prioritised thematic richness and cultural diversity over statistical generalisability.

In total, 92,750 vernacular-language posts were collected and analysed across the nine municipalities listed in Table 1. While large-scale by qualitative standards, this corpus was not intended for statistical generalisation but for thematic depth. A dataset of this size is sufficient and appropriate in qualitative research because it ensures saturation of themes (Braun & Clarke, 2006) while still allowing for close, context-sensitive coding. The emphasis was on identifying patterns of meaning rather than quantifying frequencies, which aligns with the interpretive orientation of the study.

Data collection utilised X's Academic Research Product Track API, allowing for systematic retrieval of historical posts while complying with platform terms of service. For the analysis phase, sentiment analysis was conducted using a combination of manual thematic coding

and automated linguistic tools. VADER (Valence Aware Dictionary and sEntiment Reasoner) was applied for initial polarity detection, given its robustness in short-text social media contexts (Hutto & Gilbert, 2014), and was supplemented by TextBlob for cross-validation of sentiment scores (Loria, 2018). However, because standard sentiment tools are limited in their treatment of African indigenous languages, the automated outputs were refined manually to ensure cultural and linguistic accuracy. The thematic analysis identified key service delivery issues, emotional tones (anger, frustration, solidarity), and linguistic markers of collective mobilisation. This methodological approach enabled a robust, ethically sound exploration of how African language expressions on X function as informal citizen-based monitoring in South Africa's evolving digital governance landscape.

Table 1: Selected Municipalities[1] for X Content Analysis

Municipality	Province	Platform	Status	Followers	Last Post Date	Posts (2020–2025)
City of Johannesburg	Gauteng	X	Active	1200000	2025-04-28	18000
City of Cape Town	Western Cape	X	Active	950000	2025-04-27	16250
eThekwini Municipality	KwaZulu-Natal	X	Active	850000	2025-04-26	15400
Buffalo City Metropolitan Municipality	Eastern Cape	X	Active	500000	2025-04-25	9400
Mangaung Metropolitan Municipality	Free State	X	Active	450000	2025-04-24	8800
Polokwane Municipality	Limpopo	X	Active	400000	2025-04-23	8200
Mbombela Municipality	Mpumalanga	X	Active	350000	2025-04-22	7600
Sol Plaatje Municipality	Northern Cape	X	Active	300000	2025-04-21	6900
Rustenburg Municipality	North West	X	Active	250000	2025-04-20	6200

Source: Author's own compilation (2025)

Semi-structured interviews

To complement the content analysis of vernacular digital expressions, the study conducted semi-structured interviews with three key groups of stakeholders directly engaged in citizen monitoring and service delivery oversight: (1) government officials responsible for CBM initiatives in South Africa, which include DPME, DPSA, and PSC; (2) representatives from CBOs actively involved in service delivery advocacy; and (3) staff from community radio stations serving linguistically diverse constituencies. These actors were selected purposively because of their frontline engagement with citizen feedback and municipal accountability processes.

Interview participants were identified through purposive sampling, targeting individuals with demonstrable experience in managing or interpreting vernacular citizen feedback.

1 Only the main verified official municipal X accounts were included in the analysis. Subsidiary or department-specific accounts (e.g., electricity, transport, waste services) were excluded to maintain consistency and comparability across municipalities.

A snowballing technique was subsequently employed where initial interviewees referred additional participants working in similar roles. This approach ensured that the final sample reflected a mix of institutional, civil society, and media actors with practical knowledge of vernacular citizen-based monitoring, thereby strengthening the relevance of the findings to the study's objectives.

The semi-structured interview format was chosen to balance thematic consistency and flexibility, enabling participants to elaborate on their experiences interpreting, amplifying, and responding to vernacular citizen complaints on digital platforms (Brinkmann, 2013). A total of 12 interviews were conducted (see Table 2). The interviews explored participants' perceptions of the significance of vernacular digital complaints, institutional capacities for monitoring and responding to such expressions, and perceived barriers to integrating this feedback into formal governance structures. By capturing the perspectives of both institutional actors and community intermediaries, the interview component enriches the study's understanding of how vernacular citizen-based monitoring practices intersect – or fail to intersect – with formal service delivery governance mechanisms.

Table 2: Sector Specifications of the Respondents

Organisation	Province	Sector	Respondent Code	Experience (yrs)
A	Gauteng	Government	ARE1	5 years +
B	Gauteng	Government	BRE2	5 years +
C	Gauteng	Government	CRE3	5 years +
D	Gauteng	Government	DRE4	5 years +
E	Mpumalanga	Community Media	ERE5	5 years +
F	Free State	Community Media	FRE6	5 years +
G	Gauteng	Community Media	GRE7	5 years +
H	Limpopo	Community Media	HRE8	5 years +
I	Western Cape	CBO	IRE9	5 years +
J	Gauteng	CBO	JRE10	5 years +
K	Gauteng	CBO	KRE11	5 years +
L	KwaZulu-Natal	CBO	LRE12	5 years +

Source: Author's own compilation (2025)

Data Analysis: Thematic Integration and Ensuring Rigour

Data analysis followed a thematic approach aimed at identifying patterns across both the vernacular-language X content and the semi-structured interview transcripts. Initially, content analysis of tweets involved a two-stage coding process: an inductive open-coding phase to capture emergent service delivery themes, emotional tones, and linguistic characteristics, followed by a deductive phase where codes were refined against the research questions and theoretical framing around the digital public sphere (Braun & Clarke, 2006). Sentiment scores generated through VADER and TextBlob were triangulated with manual coding to mitigate algorithmic biases against African indigenous languages. For interview data, transcripts were analysed using NVivo 14 software to facilitate systematic coding and thematic development across cases. A combination of descriptive coding (for institutional practices and challenges) and process coding (for actions like interpreting, amplifying, and

responding to vernacular complaints) was employed (Saldaña, 2021). Rigour was ensured through multiple strategies: coder triangulation (independent coding by two researchers with reconciliation sessions), an audit trail of coding decisions, and detailed memo writing to document analytical reflections. Validity and trustworthiness were further strengthened by conducting member checks with a subset of interview participants who reviewed preliminary findings for resonance and factual accuracy (Birt et al., 2016). Data saturation was monitored and confirmed once additional interviews yielded no substantially new themes. This layered analytical strategy ensured that findings were grounded in the lived realities of citizens and institutional actors and met the standards of credibility, dependability, and confirmability expected of qualitative research. By integrating digital content and stakeholder perspectives thematically, the analysis captures the complex dynamics through which vernacular citizen expressions on social media shape, challenge, or bypass formal mechanisms of municipal accountability.

Ethical Considerations

This study complied with ethical research standards as approved by the University of Johannesburg's Faculty Ethics Committee. All interview participants provided informed consent and their identities have been anonymised using coded references (e.g., ARE1, BRE2). Direct quotations are paraphrased where necessary to prevent traceability. Regarding the content analysis of X, only publicly available posts were collected through the Academic Research Product Track API. As these posts are in the public domain, no private data were accessed, and appropriate safeguards (such as paraphrasing of vernacular expressions) were applied to protect user confidentiality.

Empirical Findings

Amplification of Local Voices through Vernacular Language on X

The content analysis revealed that X has become a vital digital arena where South African citizens use indigenous languages to document and protest municipal service delivery failures, including potholes, water shortages, electricity outages, and waste collection breakdowns. Across the sampled municipalities, vernacular posts frequently employed culturally specific expressions, metaphors, and idioms that carried emotional and historical resonance, often framing service delivery grievances in ways that English translations would fail to capture fully. For example, isiZulu expressions likening potholes to "traps for cattle" or Sepedi phrases describing dry taps as "the death of the home" imbued the complaints with deep communal significance and urgency. Sentiment analysis confirmed that the dominant emotional tones across vernacular posts were anger, frustration, and solidarity. Rather than isolated complaints, these expressions signalled a shared sense of injustice where anger and frustration were most common but often intersected with solidarity. Solidarity was particularly evident where users retweeted, amplified, or added collective calls for action, often using indigenous-language hashtags such as #Asikhulume (isiZulu for "let's talk") and #ReBatlaMetsi (Setswana/Sepedi for "we want water"). These hashtags functioned not only as rallying cries but also as vernacularised spaces of mobilisation, allowing users to frame shared grievances in culturally grounded terms that municipal authorities could not easily ignore.

Importantly, vernacular expressions fostered emotional proximity and authenticity, signalling to fellow citizens that service failures were a community-wide injury rather than

isolated inconveniences. This collective dimension was evident in how grievances framed through indigenous languages created resonance, legitimacy, and traction among local audiences. As one CBO representative explained, "When people post in Sepedi or isiZulu, they are not just complaining; they are speaking in the language of the community. It carries more weight because everyone understands the struggle in their own words" (JRE10). Such voices highlight how indigenous language use amplifies marginalised perspectives by producing affectively charged narratives that position service delivery failures as collective injustices. By embedding grievances in local linguistic and cultural repertoires, citizens shaped a counter-discourse that both challenges municipal legitimacy and strengthens grassroots solidarity in digital public spheres.

Digital Vernacular Counterpublics: Claiming Civic Space Beyond Formal Structures

The analysis demonstrates that indigenous language posts on X function as more than individualised complaints; they coalesce into vernacular counter publics that challenge the dominance of English in both formal governance communication and mainstream digital discourse. By using isiZulu, Setswana, isiXhosa, and other African indigenous languages, citizens carved out parallel spaces of civic expression where grievances were voiced on their own cultural and linguistic terms. These counterpublics reflect Fraser's (1990) notion of subaltern publics that arise when dominant forums exclude marginalised groups, and they were particularly visible in contexts where municipal responses were absent or dismissive. The persistence of vernacular digital voices in these spaces illustrates how citizens collectively claim legitimacy and recognition, despite systemic barriers that marginalise their languages in both policy and practice. In doing so, they invert expectations of who holds communicative authority online by foregrounding lived experiences articulated in languages that municipalities are ill-equipped to interpret or respond to.

These vernacular counterpublics also served as sites of mobilisation and solidarity. Hashtags and phrases in indigenous languages became symbolic anchors for collective struggles, framing grievances in terms of shared community struggles rather than isolated incidents. Emotional resonance and cultural intimacy strengthened this dynamic, enabling users to bypass the alienation often produced by English-only bureaucratic discourse. As a community radio staff member explained, "We often see posts in Setswana or isiXhosa tagged to mayors. Even if the municipality does not respond, the people online rally around those words. It creates a movement that is bigger than a single complaint" (FRE6). Such reflections underscore how vernacular CBM operates as both cultural resistance and political engagement. By asserting space for African Indigenous languages in digital civic arenas, citizens transform grievances into collective claims for recognition, effectively broadening the boundaries of participatory democracy in South Africa's multilingual context.

Institutional Blind Spots and Language Barriers in Municipal Digital Engagement

The findings highlight significant institutional blind spots in how municipalities manage vernacular digital complaints. Despite the volume of posts in African indigenous languages, most municipal communication systems are designed primarily for English, with limited or no capacity to process multilingual content. Municipal digital teams often lack linguistic resources, translation tools, or training to recognise vernacular expressions as legitimate channels of accountability. As a result, complaints expressed in isiZulu, isiXhosa, or Sepedi are frequently ignored or misclassified, perpetuating cycles of exclusion. This gap reinforces the dominance of English in governance processes, effectively silencing large sections of

the population who may be more comfortable expressing their grievances in indigenous languages. It also exposes a structural weakness in South Africa's democratic framework where linguistic diversity is constitutionally protected but rarely operationalised in practice.

Institutional actors interviewed acknowledged this weakness but noted bureaucratic and technical barriers to reform. Digital monitoring tools currently in use cannot easily process African indigenous languages, leaving municipal staff reliant on ad hoc strategies such as translating posts manually or depending on intermediaries. As one municipal official admitted, "We do not have systems that can pick up or translate vernacular complaints. If someone writes in isiZulu or Sesotho, it disappears from our radar. That is a major gap" (ARE1). Another official added that the lack of institutional ownership for social media monitoring further compounds the problem given that the responsibility is often fragmented across departments. These blind spots illustrate how vernacular CBM is undermined not by citizen apathy but by institutional incapacity to respond inclusively. Unless municipalities develop multilingual monitoring systems and strengthen accountability mechanisms, vernacular expressions will continue to be overlooked, reinforcing perceptions of state indifference and eroding public trust in local governance.

Community-Based Organisations and Community Radio as Intermediaries

The findings show that CBOs and community radio stations act as crucial intermediaries between vernacular citizen voices and municipal authorities. Many citizens prefer to lodge complaints in isiZulu, Sepedi, Setswana, or isiXhosa through familiar digital platforms such as WhatsApp and Facebook. However, these messages rarely reach government officials directly due to both technological barriers and the limited linguistic capacity within municipalities. CBOs often step into this gap by systematically collecting vernacular complaints from local networks and repackaging them into English-language summaries that can be forwarded to officials or included in advocacy reports. This intermediary role positions CBOs as vital translators of community voices, ensuring that grievances framed in indigenous languages do not remain invisible in governance processes. In this way, CBOs provide an informal but highly effective accountability mechanism, amplifying the civic power of vernacular expressions.

Community radio stations similarly extend the reach of vernacular CBM by broadcasting grievances to wide audiences and relaying them to municipal officials. These stations often operate call-in programmes where residents express service delivery concerns in their own languages, creating a participatory forum that municipal structures fail to provide. As one CBO leader explained, "We check WhatsApp groups and Facebook posts in isiZulu or Sepedi, and then we compile them into English summaries for the municipality. Without us, those voices would never be heard at the council level" (KRE11). A community radio presenter echoed this, noting, "When a caller says in isiXhosa that there is no water for three days, we broadcast it to thousands of listeners and also send the clip to the local ward office. It is our way of amplifying their voice" (GRE7). These intermediaries play a dual role: amplifying vernacular voices within the public sphere and bridging the institutional disconnect by relaying complaints into formal governance spaces. Their work demonstrates that inclusive accountability in South Africa is not simply a matter of technology but also of cultural mediation and grassroots advocacy.

Barriers to Institutionalising Vernacular Citizen-Based Monitoring Digitally

A primary challenge identified in the study is the absence of technical infrastructure capable of detecting, translating, and categorising African indigenous language complaints on digital platforms. Although municipalities have experimented with CBM frameworks, these remain largely paper-based or dependent on English-language call centres and reporting tools. Digital citizen monitoring systems that could process vernacular inputs are almost entirely absent. Current municipal monitoring platforms are often designed by external vendors with limited localisation capacity, meaning that languages such as isiZulu, Setswana, or Sepedi are excluded from system interfaces and analytics functions. This not only undermines inclusivity but also results in the systematic erasure of grievances expressed in vernacular forms. As one community-based organisation leader explained during interviews, "The municipality says they have a citizen feedback system, but if you write in isiZulu or Sepedi, the system doesn't even register it. It only understands English, so those complaints go nowhere." Such technological blind spots reveal a disconnect between constitutional commitments to multilingualism and the reality of digital governance in practice.

Even where municipalities engage with citizen feedback, existing CBM practices are not conducted digitally but remain manual and fragmented. Reports are compiled from ward committee meetings, paper-based forms, or sporadic hotline calls which are then translated and filtered before reaching officials. This reliance on non-digital, centralised methods sidelines the dynamic and immediate nature of vernacular expressions on platforms such as X, Facebook, or WhatsApp. Moreover, institutional culture compounds these barriers. Some officials continue to perceive vernacular complaints on social media as informal and therefore less credible, reinforcing English as the de facto language of governance. A municipal respondent candidly admitted, "We still treat social media complaints as noise. If they are not submitted through the official form in English, they are ignored." Such attitudes not only perpetuate historical linguistic hierarchies but also prevent the integration of digital vernacular monitoring into formal accountability systems. Without deliberate investment in multilingual digital tools, cultural shifts within governance institutions, and policies that recognise the legitimacy of vernacular voices online, efforts to institutionalise citizen-based monitoring digitally will remain aspirational rather than achievable.

Discussions

Reframing Vernacular Citizen-Based Monitoring as Democratic Participation

The findings of this study challenge dominant assumptions that citizen complaints on social media, particularly those articulated in African indigenous languages, are informal, fragmented, or politically marginal expressions. Instead, they reveal that vernacular citizen monitoring practices constitute a vital form of democratic participation, asserting both visibility and voice within South Africa's evolving digital governance landscape. Drawing on Papacharissi's (2015) notion of affective publics and Fraser's (1990) theory of subaltern counterpublics, vernacular digital complaints can be understood as deliberate political acts that contest the exclusionary boundaries of the formal public sphere. Citizens' use of indigenous languages to articulate service delivery failures is not simply a reflection of cultural preference but an assertion of civic entitlement – the right to hold government accountable in one's own linguistic register. This repositions vernacular digital activism as a participatory practice deeply rooted in struggles for linguistic recognition, service delivery

equity, and democratic inclusion. Importantly, the emergence of vernacular counterpublics on platforms like X underscores the adaptability and resilience of citizen agency in contexts where formal participatory mechanisms remain inaccessible, unresponsive, or linguistically exclusionary. By framing vernacular CBM as a legitimate and necessary extension of democratic life, the study contributes to broader debates on how digital technologies reconfigure the modes, spaces, and languages through which citizenship is enacted and accountability is demanded. Recognising these practices as authentic forms of civic engagement necessitates a paradigmatic shift in how governments, policymakers, and scholars conceptualise digital participation – not merely as rational deliberation in dominant languages but as diverse, affectively charged, and linguistically pluralistic acts of democratic expression.

Language as Infrastructure for Digital Inclusion and Civic Accountability

The findings underscore that language operates as a communication medium and a foundational infrastructure for digital inclusion and civic accountability. In the context of South Africa's multilingual democracy, the failure to integrate African indigenous languages into digital governance spaces effectively disenfranchises large segments of the population from participating fully in democratic processes. Hernandez and Roberts (2018) argue that digital inclusion must go beyond access to devices or connectivity, extending to the cultural and linguistic frameworks that shape meaningful participation. Similarly, Ragnedda (2018) warns that digital spaces risk replicating and even amplifying existing social inequalities without addressing linguistic and cultural barriers. The vernacular citizen monitoring practices uncovered in this study demonstrate that African languages are critical infrastructures through which grievances are articulated, collective identities are forged, and demands for accountability are expressed. Yet, government digital engagement practices remain anchored in a narrow, English-dominated paradigm, effectively rendering vernacular complaints invisible within official service delivery monitoring systems. This linguistic exclusion not only violates principles of language justice but undermines the very goals of participatory governance by systematically marginalising non-English speakers from digital accountability mechanisms. As scholars such as Mpofu and Salawu (2020); Chonka et al. (2022); and Haile (2022) have shown, algorithmic and institutional biases that privilege dominant languages create structural inequities in the digital public sphere. Addressing these exclusions requires conceptualising language not as a secondary concern but as core governance infrastructure – essential to ensuring that the digital transformation of civic engagement genuinely democratises access, voice, and power.

Institutional Resistance and the Persistence of Bureaucratic Gatekeeping

The study's interviews with government officials revealed a persistent pattern of institutional resistance to integrating vernacular CBM into formal governance systems, reflecting deep-seated bureaucratic gatekeeping tendencies. Despite recognising the growing prevalence of citizen grievances voiced in indigenous languages on social media, officials expressed hesitations rooted in concerns over procedural control, verification challenges, and the perceived legitimacy of informal digital complaints. This reluctance mirrors broader critiques by scholars such as Eom et al. (2021) and Rumbul (2020) who argue that public institutions often treat informal citizen engagement on digital platforms as disruptive rather than constructive. In the South African context, these anxieties are compounded by historically entrenched administrative cultures favouring hierarchical, paper-based, and English-dominated modes of interaction, as highlighted by Okeke-Uzodike and Dlamini

(2019). The persistence of such bureaucratic gatekeeping undermines the transformative potential of digital technologies to democratise governance and widen participation. Instead of embracing the vernacularisation of the digital public sphere as an opportunity to enhance inclusivity and responsiveness, institutions retreat into proceduralism, thereby excluding a large portion of the citizenry from meaningful dialogue. This dynamic reflects what Fraser (1990) describes as the structural marginalisation of subaltern counterpublics, wherein the communicative practices of historically disenfranchised groups are deemed invalid within dominant institutional frameworks. By failing to adapt to the realities of vernacular digital engagement, government actors reinforce existing linguistic and socio-economic inequalities and weaken the broader goals of participatory democracy and service delivery accountability that CBM initiatives are intended to promote.

Toward a Vernacular-Centred Digital Governance Framework

The findings of this study point toward the urgent need for a vernacular-centred digital governance framework that recognises institutionalise and strengthens CBO practices expressed through Indigenous languages. Rather than treating vernacular digital complaints as peripheral or informal, municipalities and government agencies must formally integrate these expressions into service delivery monitoring systems. This would require the development of multilingual digital engagement strategies that enable the detection, classification, and response to citizen grievances articulated in African languages across social media platforms. Building partnerships with CBOs and community radio stations – already operating informally as civic intermediaries – could provide an immediate and culturally resonant bridge between citizen voices and formal governance structures. In the longer term, investment in NLP technologies, as advocated by Chonka et al. (2022), would enhance municipalities' ability to automate the monitoring of vernacular expressions without sacrificing linguistic nuance or cultural specificity. Policy innovations, such as incorporating vernacular citizen feedback into IDPs and municipal performance management frameworks, would institutionalise these practices and lend them procedural legitimacy. Moreover, piloting vernacular CBM initiatives within selected municipalities would serve as proof-of-concept projects, helping to build political will and foster a culture of linguistic inclusivity within digital governance. Such a framework would address current technological and institutional deficits and advance the broader goals of participatory democracy by affirming that all citizens – regardless of linguistic background – have an equal right to be heard, seen, and responded to within South Africa's digital governance landscape.

Implications and Lessons Learned for Digital Governance and Citizen-Based Monitoring

The findings of this study carry significant implications for the theory and practice of digital governance and CBM in South Africa and comparable multilingual democracies. Theoretically, they demand a rethinking of Digital Public Sphere frameworks by foregrounding vernacular, affective, and linguistically diverse forms of civic engagement as central rather than peripheral to democratic participation. Digital citizen voice cannot be conceptualised solely through English or rational deliberation but must account for the emotional, cultural, and political significance of indigenous language expressions in digital spaces. From a policy perspective, excluding vernacular complaints from formal governance systems represents not a technological inevitability but a policy failure, highlighting the

need for deliberate institutional reforms that centre on multilingual digital inclusion. CBM policies, such as those developed by the DPME, must move beyond offline and English-centric modes of engagement to formally recognise, monitor, and act upon digital citizen feedback articulated in indigenous languages. Practically, the study illustrates that citizen monitoring is already happening in informal digital spaces and being mediated through CBOs and community radio stations; governments must act swiftly to support and integrate these networks into formal accountability systems. Reliance on English-only digital tools leaves municipal authorities deaf to large population segments, undermining service delivery responsiveness and deepening democratic deficits. Overall, the lessons learned suggest that institutionalising vernacular CBM is both a governance imperative and an opportunity to revitalise citizen participation, linguistic justice, and democratic inclusion in the digital age.

Moreover, the study highlights the critical importance of reconceptualising digital governance infrastructures as technical systems and socio-linguistic ecosystems that must be intentionally designed to accommodate diversity. Lessons from the findings suggest that participatory governance will continue to exclude marginalised voices structurally unless linguistic pluralism is embedded at the design stage of digital monitoring platforms. This calls for reimagining civic technology procurement and design processes, including mandatory multilingual functionality, indigenous language sentiment analysis capabilities, and consultation of language and community experts during platform development. Lessons also point to the need for broader public sector innovation strategies that view digital civic engagement not as an isolated administrative task but as a central pillar of developmental local governance. Vernacular CBM practices reveal that digital citizenship is relational, affective, and culturally situated; policies and systems that fail to recognise these dynamics risk deepening the democratic disconnection between citizens and the state. Therefore, strengthening vernacular CBM is not simply about adding translation layers to existing frameworks, but fundamentally realigning governance institutions to better mirror and respond to the complex, multilingual realities of the publics they serve.

Policy and Practice Recommendations

This study's first primary policy recommendation is the urgent need for municipalities to develop and implement multilingual digital engagement strategies that actively incorporate African indigenous languages into their citizen monitoring frameworks. The findings showed that while citizens increasingly use vernacular languages to voice service delivery grievances online, municipal digital infrastructures remain overwhelmingly English-centric, effectively silencing a large segment of the citizenry. To correct this exclusion, municipal communication and CBM teams must be trained and resourced to systematically monitor, analyse, and respond to indigenous-language expressions across social media platforms such as X, Facebook, and WhatsApp. This would require redesigning sentiment analysis tools, keyword tracking systems, and citizen response dashboards to include multilingual capabilities. Language specialists and citizen linguists could be formally incorporated into digital engagement teams to ensure culturally sensitive interpretation of grievances. Furthermore, municipal websites, online complaint forms, and digital reporting mechanisms should be redesigned to allow users to submit feedback in their preferred languages. In addition to internal capacity-building, municipalities must establish clear operational protocols that treat vernacular citizen complaints as equally valid and actionable inputs within broader service delivery monitoring processes. Developing these multilingual

strategies would help address both technological and institutional blind spots revealed in the findings. It would represent a concrete step toward institutionalising vernacular citizen-based monitoring as a legitimate element of digital governance.

The second major recommendation is establishing formal partnerships between municipalities and intermediary organisations such as CBOs and community radio stations. The findings revealed that CBOs and community radios already perform crucial, informal roles in translating, amplifying, and escalating vernacular citizen complaints despite operating without institutional support or recognition. Municipalities should develop Memorandums of Understanding (MOUs) or partnership agreements that formally incorporate CBOs and radio stations into municipal CBM frameworks to harness and enhance these efforts. These partnerships should define clear roles, including real-time issue escalation processes, co-management of multilingual digital complaint platforms, and participation in service delivery forums. Municipalities could also fund and support capacity-building initiatives for these intermediaries, including training in digital monitoring, advocacy skills, and basic data management.

Furthermore, community radio stations could be leveraged as key platforms for multilingual civic education campaigns, informing citizens about how to lodge complaints, track service delivery promises, and hold local governments accountable using digital and offline methods. Formalising these partnerships would strengthen the feedback loop between citizens and government and ensure that vernacular voices, currently thriving in informal spaces, are institutionally integrated into official governance processes. Such an approach would operationalise the inclusivity and responsiveness that participatory governance frameworks, such as the DPME's CBM policy, aspire to achieve but have thus far struggled to realise in practice.

The third recommendation addresses the technological infrastructure gap by calling for substantial investment in African language NLP technologies and amendment of existing governance frameworks to institutionalise vernacular digital feedback mechanisms. The findings highlighted that existing sentiment analysis and social media monitoring tools are ill-equipped to detect, classify, and interpret indigenous-language citizen expressions, resulting in significant gaps in municipal responsiveness. To address this, municipalities, national government agencies, and research institutions should collaborate to develop AI-driven, African language–compatible NLP systems tailored to local governance needs. Pilot projects could focus initially on high-frequency service delivery complaints (e.g., water, electricity, potholes) across selected languages, expanding as the models improve. Parallel to technological innovation, policy frameworks such as IDPs, Municipal Performance Management Systems, and CBM guidelines should be revised to include vernacular citizen feedback as a formal performance indicator explicitly. Municipal scorecards could be expanded to track service delivery outputs and responsiveness to citizen complaints articulated across multiple languages and platforms. Additionally, municipalities should pilot multilingual digital CBM initiatives in select wards or regions, using them as proof-of-concept models to build political will, refine operational procedures, and demonstrate the governance benefits of linguistically inclusive citizen monitoring. Without these systemic technological and policy shifts, vernacular CBM will remain confined to informal spaces, reinforcing exclusion rather than advancing South Africa's democratic project.

Conclusion and Directions for Future Research

This study has demonstrated that African indigenous languages are central to how South African citizens articulate grievances, demand accountability, and participate in municipal service delivery monitoring within digital spaces, particularly on platforms like X. Far from being peripheral or informal, vernacular digital expressions represent a vibrant, politically significant form of CBM that challenges traditional bureaucratic assumptions about civic engagement, language, and legitimacy. However, institutional resistance, technological shortcomings, and the absence of supportive policy frameworks have collectively marginalised these practices, relegating vernacular CBM to the informal peripheries of governance. Addressing this exclusion requires deliberate reforms that recognise indigenous languages as core infrastructures of democratic participation, invest in multilingual technological solutions, and formalise partnerships with community-based organisations and community radio stations that already act as vernacular civic intermediaries. By centring vernacular voices within digital governance systems, municipalities can build more inclusive, responsive, and democratic service delivery frameworks aligned with South Africa's constitutional values of equality, dignity, and participation.

Future research should extend this work by examining the longitudinal impacts of institutionalising vernacular digital citizen monitoring on governance performance and citizen trust in local government. Comparative studies across different municipalities, provinces, or countries could shed further light on how linguistic diversity shapes digital civic participation in varied governance contexts. There is also a pressing need for interdisciplinary research combining computational linguistics, public administration, and civic technology to accelerate the development of African language NLP tools tailored to public sector needs. Finally, further exploration of the emotional and affective dimensions of vernacular digital activism is needed, particularly how affect-laden vernacular expressions mobilise civic solidarity, shape public opinion, and influence political accountability in multilingual societies. In pursuing these avenues, future scholarship can continue to advance more inclusive models of digital democracy that genuinely reflect the cultural and linguistic realities of the citizens they aim to serve.

References

Aiyegbusi, B.T. (2018). Decolonizing digital humanities: Africa in perspective. In: E. Losh & J. Wernimont (eds.). Bodies of information, Minneapolis: University of Minnesota Press, pp. 75-90.Alexander, N. (2019). Language policy and national unity in South Africa/Azania. Cape Town: Buchu Books. https://doi.org/10.5749/j.ctv9hj9r9.26

Arshad, S. & Khurram, S. (2020). 'Can government's presence on social media stimulate citizens' online political participation? Investigating the influence of transparency, trust, and responsiveness', Government Information Quarterly, 37(3), 1-18. https://doi.org/10.1016/j.giq.2020.101486

Banda, F., & Mwanza, D. S. (2023). 'Introduction. Multilingual Margins', A Journal of Multilingualism from the Periphery, 7(3), 5-13.

Bester, A. (2015). Scoping Study on Monitoring, Review, and Accountability for Development Cooperation to support implementation of a Post-2015 Development Agenda, United Nations Development Programme. Available at: https://share.google/7DZ9ZEOdTU05Pszqs. Accessed: 09 August 2025.

Birt, L., Scott, S., Cavers, D., Campbell, C. & Walter, F. (2016). 'Member checking: A tool to enhance trustworthiness or merely a nod to validation?', Qualitative Health Research, 26(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Bonsón, E., Royo, S. & Ratkai, M. (2015). 'Citizens' engagement on local governments' Facebook sites. An empirical analysis: The impact of different media and content types in Western Europe', Government Information Quarterly, 32(1), 52–62. https://doi.org/10.1016/j.giq.2014.11.001

Bonsón, E., Royo, S. & Ratkai, M. (2017). 'Facebook practices in Western European municipalities: An empirical analysis of activity and citizens' engagement', Administration and Society, 49(3), 320–347. https://doi.org/10.1177/0095399714544945

Boulianne, S., Theocharis, Y., & van Deth, J.W. (2022). 'Digital media and political engagement: A review and future research agenda', New Media & Society, 24(6), 1285–1306. https://doi.org/10.1177/1461444820968210

Braun, V. & Clarke, V. (2006). 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp0630a

Breakfast, N., Bradshaw, G. & Nomarwayi, T. (2019). 'Violent service delivery protests in post-apartheid South Africa, 1994–2017: A conflict resolution perspective', African Journal of Public Affairs, 11(1), 106–126.

Brinkmann, S. (2013) Qualitative Interviewing: Understanding Qualitative Research. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199861392.001.0001

Carlson, T. & Cohen, A. (2018). 'Linking community-based monitoring to water policy: Perceptions of citizen scientists', Journal of Environmental Management, 219, 168–177. https://doi.org/10.1016/j.jenvman.2018.04.077

Chonka, P., Diepeveen, S. & Haile, Y. (2022). 'Algorithmic Power and African Indigenous Languages: Search Engine Autocomplete and the Global Multilingual Internet', Media, Culture & Society, 45(2), pp. 246–265. https://doi.org/10.17863/CAM.94298

Conrad, C.T. & Daoust, T. (2008). 'Community-based monitoring frameworks: Increasing the effectiveness of environmental stewardship', Environmental Management, 41(3), 358–366. https://doi.org/10.1007/s00267-007-9042-x

Dahlgren, P. (2005). 'The Internet, public spheres, and political communication: Dispersion and deliberation', Political Communication, 22(2), 147–162. https://doi.org/10.1080/10584600590933160

Danielsen, F., Burgess, N.D. & Balmford, A. (2005). 'Monitoring matters: Examining the potential of locally-based approaches', Biodiversity and Conservation, 14(11), 2507–2542. https://doi.org/10.1007/s10531-005-8375-0

Dean, J. (2003) 'Why the Net is Not a Public Sphere, Constellations', 10(1), 95–112. https://doi.org/10.1111/1467-8675.00316

DePaula, N. & Dincelli, E. (2018). 'Information strategies and affective reactions: How citizens interact with government social media content', First Monday, 23(4), 1–15. https://doi.org/10.5210/fm.v23i4.8414

DePaula, N., Dincelli, E. & Harrison, T.M. (2018). 'Toward a typology of government social media communication: Democratic goals, symbolic acts and self-presentation', Government Information Quarterly, 35(1), 98–108. https://doi.org/10.1016/j.giq.2017.10.003

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J. & Purcell, K. (2012). 'The current state of citizen science as a tool for ecological research and public engagement', Frontiers in Ecology and the Environment, 10(6), 291–297. https://doi.org/10.1890/110236

DPME (Department of Planning, Monitoring, and Evaluation). (2013). Independent impact assessment of the community monitoring and advocacy programme of the Black Sash Trust, Government

of South Africa, Pretoria. Available at: https://share.google/Sx3H6yZqPYNgNZSBT. Accessed: 09 August 2025.

DPME (Department of Planning, Monitoring, and Evaluation). (2015). Annual report 2014/15, Government of South Africa, Pretoria. Available at: https://share.google/iKz2yGNTV7r7IUfIE. Accessed: 09 August 2025.

DPME (Department of Planning, Monitoring, and Evaluation). (2016). Improving public services: Citizen-based monitoring of service delivery, Government of South Africa, Pretoria. Available at: https://share.google/Gi6HugjelGyNxaRI9 Accessed: 09 August 2025.

Earl J, Maher TV, Pan J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. Sci Adv. 2022 March 11, 8(10), 81–98. https://doi.org/10.1126/sciadv.abl8198

Ellison, N. & Hardey, M. (2014). 'Social media and local government: Citizenship, consumption and democracy', Local Government Studies, 40(1), 21–40. https://doi.org/10.1080/03003930.2013.799066

Eom, S.J., Hwang, H. & Kim, J.H. (2018). 'Can social media increase government responsiveness? A case study of Seoul, Korea', Government Information Quarterly, 35(1), 109–122. https://doi.org/10.1016/j.giq.2017.10.002

Eom, S.J., Kim, J.H., & Nam, S. (2021). 'Citizen engagement and government responsiveness on social media: A comparative study of Korean and US local governments', Government Information Quarterly, 38(1), 1–15. https://doi.org/10.1016/j.giq.2020.101547

Fashoro, I. & Barnard, L. (2017). 'Challenges to the successful implementation of social media in a South African municipality', in Proceedings of SAICSIT 2017, ACM Digital Library. https://doi.org/10.1145/3129416.3129426

Fashoro, I. & Barnard, L. (2021). 'Assessing South African Government's Use of Social Media for Citizen Participation', The African Journal of Information Systems, 13(1), Article 3. Available at: https://digitalcommons.kennesaw.edu/ajis/vol13/iss1/3. Accessed: 09 August 2025.

Fraser, N. (1990). 'Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy', Social Text 25/26, 56–80. https://doi.org/10.2307/466240.

Friedman, S. (2006). Participatory governance and citizen action in post-apartheid South Africa, Occasional Paper No. 164, United Nations Research Institute for Social Development. Available at: https://share.google/vWUZYU2n9yapdR502. Accessed: 09 August -2025.

Gu, T., Harrison, T.M. & Zhu, Y. (2020). 'Municipal government use of social media: An analysis of three Chinese cities', in Proceedings of the 53rd Hawaii International Conference on System Sciences, vol. 3, pp. 1803–1812. https://doi.org/10.24251/hicss.2020.223

Habermas, J. (1962). The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society, MIT Press (1989 English translation). Available at: https://share.google/Cpr0BhaumxRyBI9w5. Accessed: 09 August -2025.

Halpern, D. & Katz, J.E. (2012). 'From e-government to social network government: Towards a transition model', Proceedings of the 3rd Annual ACM Web Science Conference, pp. 119–127. https://doi.org/10.1145/2380718.2380735

Haro-de-Rosario, A., Sáez-Martín, A. & del Carmen Caba-Pérez, M. (2018). 'Using social media to enhance citizen engagement with local government: Twitter or Facebook?', New Media and Society, 20(1), 29–49. https://doi.org/10.1177/1461444816645652

Hernandez, K. & Roberts, T. (2018). Leaving No One Behind in a Digital World, K4D Emerging Issues Report. Institute of Development Studies. Available at: https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/13968. Accessed: 09 August 2025.

HSRC (Human Sciences Research Council). (2019). Code of Research Ethics. Human Sciences Research Council, South Africa. Available at: https://hsrc.ac.za/wp-content/uploads/2021/07/HSRC-Research-Ethics-Guidelines.pdf. Accessed: 24 January 2025.

Independent Communications Authority of South Africa (ICASA). (2024). The State of ICT Sector Report in South Africa: 2024. ICASA. Available at: https://www.icasa.org.za/pages/reports. Accessed: 24 January 2025.Ingram, M. & Licona, A.C. (2023). 'Vernacular activism and resistance in the digital age: Language as power in civic engagement', Journal of Language and Politics, 22(1), 98–120. https://doi.org/10.1075/jlp.21095.ing

Jamil, S. & Awal, M.A. (2021) 'Social media and citizen engagement: Redefining government–citizen interaction during COVID-19 crisis', Government Information Quarterly, 38(4), 101622. https://doi.org/10.1016/j.giq.2021.101622

Jongbloed-Faber, L., van der Velde, H., van der Meer, C. & Klinkenberg, E. (2016). 'Language use of Frisian bilingual teenagers on social media', Linguistics and Sociology, (26), 27–54. https://doi.org/10.2436/20.2504.01.107

Kaya, H. O., Kamwendo, G. H., & Rushubirwa, L. (2016). 'African Indigenous languages in higher education'. Studies of Tribes and Tribals, 14(2), 121–128. https://doi.org/10.1080/0972639X.2016.11886739

Kreiss, D., Lawrence, R.G. & McGregor, S.C. (2020). 'Political communication and the affordances of social media platforms', Journal of Communication, 70(2), 214–217. https://doi.org/10.1093/joc/jqaa003

Lewandowski, E., Caldwell, W., Elmquist, D. & Oberhauser, K. (2017). 'Public perceptions of citizen science', Citizen Science: Theory and Practice, 2(1), 1–9. https://doi.org/10.5334/cstp.77

Maharaj, B. & Pillay, U. (2022). 'Urban governance, service delivery and digital citizen participation in South Africa', Urban Forum, 33(4), 703–725. https://doi.org/10.1007/s12132-022-09450-8

Mamokhere, J. (2020). 'An assessment of reasons behind service delivery protests: A case of Greater Tzaneen Municipality', Journal of Public Affairs, 20(2), 20–49. https://doi.org/10.1002/pa.2049

Matlala, L.S. (2024a). 'Improving citizen-based monitoring in South Africa: A social media model', African Evaluation Journal, 12(1), 7–19. https://doi.org/10.4102/aej.v12i1.719

Matlala, L.S. (2024b). 'Factors affecting effective citizen-based monitoring of frontline service delivery in South Africa', Africa's Public Service Delivery and Performance Review, 12(1), 8–21. https://doi.org/10.4102/apsdpr.v12i1.851

Matlala L.S. (2025). 'Harnessing citizen-based monitoring for sustainable governance in South Africa: a framework for achieving SDGs through enhanced governance practice', OIDA International Journal Sustain Dev., 18(8), 109–128.

Mpofu, P. & Salawu, A. (2020). 'African language use in the digital public sphere: Functionality of the localised Google webpage in Zimbabwe', South African Journal of African Languages, 40(1), 76–84. https://doi.org/10.1080/02572117.2020.1733833

Mudzusi, T.N., Munzhedzi, P.H. & Mahole, E. (2024). 'Governance challenges in the provision of municipal services in the Vhembe District Municipality', Africa's Public Service Delivery and Performance Review, 12(1), 27–49. https://doi.org/10.4102/apsdpr.v12i1.749

Nyabola, N. (2018a). '#FreeBobiWine and Today's Pan-Africanism for the Digital Age', African Arguments, 23 August, viewed 15 September 2024. Available at: https://africanarguments.org/2018/08/free-bobiwine-and-todays-pan-africanism-for-the-digital-age/. Accessed: 09 August 2025.

Okeke-Uzodike, O.E. & Dlamini, B. (2019). 'Citizens' e-participation at local municipal government in South Africa', Journal of Reviews on Global Economics, 8, 458–468. https://doi.org/10.6000/1929-7092.2019.08.39

Papacharissi, Z. (2015). Affective Publics: Sentiment, Technology, and Politics. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199999736.001.0001

Ragnedda, M. (2018). 'Tackling Digital Exclusion: Counter Social Inequalities Through Digital Inclusion', in G.W. Muschert et al. (eds.). Global Agenda for Social Justice: Volume One, Bristol: Bristol University Press. pp. 151–158. https://doi.org/10.56687/9781447352204-021

Republic of South Africa. (1996). Constitution of the Republic of South Africa, 1996. Government Printer. Available: https://share.google/8JZiwsi7x5GGJyszG. Accessed: 09 August 2025.

Rumbul, R. (2020). 'Citizen engagement, digital government and the new civic infrastructure', Digital Government: Research and Practice, 1(1), 7. https://doi.org/10.1145/3391801

Saldaña, J. (2021). The Coding Manual for Qualitative Researchers, 4th edn. Thousand Oaks, CA: SAGE Publications Limited.

Surya, R., Rahmat, A. & Setyaningrum, R. (2023). 'The impact of social media citizen complaints on local government responsiveness: Evidence from Indonesia', Information Polity, 28(2), 207–224. https://doi.org/10.3233/IP-220181

Tufekci, Z. (2020). Twitter and tear gas: The power and fragility of networked protest. (Updated ed.). Yale University Press. Available at: https://doi.org/10.12987/9780300228175

We Are Social & Hootsuite. (2024). Digital 2020: South Africa. Available at: https://datareportal.com/reports/digital-2020-south-africa. Accessed: 24 March 2025.

Wiseman, N.D. & Bardsley, D.K. (2016). 'Monitoring to learn, learning to monitor: A critical analysis of opportunities for indigenous community-based monitoring of environmental change in Australian Rangelands', Environmental Reviews, 24(2), 151–163.

Worldwideworx & OrnicoGroup. (2024). South African social media landscape 2020. Available at: http://www.worldwideworx.com. Accessed: 24 January 2025.

Braun, V. & Clarke, V. (2006). 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp0630a

Saldaña, J. (2021). The Coding Manual for Qualitative Researchers (4th ed.), SAGE Publications.

Birt, L., Scott, S., Cavers, D., Campbell, C. & Walter, F. (2016). 'Member checking: A tool to enhance trustworthiness or merely a nod to validation?', Qualitative Health Research, 26(13), 1802–1811. https://doi.org/10.1177/1049732316654870

Exploring the Impact of Centralised Translation Memory in OmegaT Translation Software

The Case of the Xitsonga Language

Hlamalani Dollence Baloyi
Department of African Languages
University of South Africa
Pretoria, South Africa
dollenzo@gmail.com

Abstract

The role of Information Communication Technology (ICT) is crucial in language development. As the technological era evolves, it is important for the language space to keep up with current trends. Computer Assisted Translation (CAT) tools are still inadequately designed for African indigenous languages, which hinders their precision, usage, and digital participation, ultimately perpetuating linguistic disparities. Using a corpus-based translation approach, the study aims to explore the impact of centralised translation memory in OmegaT translation software, specifically in the Xitsonga language. CAT applications are used to enhance, accelerate, enrich, and deepen language skills. A qualitative research approach was adopted which comprised desktop research of secondary materials. A comparative analysis was used for analysing data as it involves a direct comparison that methodically examines the OmegaT translation memories and translated documents to identify their similarities and differences. Inconsistencies were noted in the terminologies used by different translators due to the absence of centralisation of translation memories and glossaries. The lack of translation equivalents in the standardised terminology list indicates that the terms have not been included for development; it is essential to welcome new words into the corpus and offer definitions for these terms to support the evolution of language and terminology. Centralised translation memory could lead to reduced translation time and quicker turnaround for translation projects. It is recommended that OmegaT should include a sharing feature or a repository for storing translated terms, and this should be facilitated by Department of Sport, Arts and Culture, as they carry the mandate.

Keywords: OmegaT, Translation Equivalent, CAT Tool, Software, Terminology, Glossary

Introduction

Although the global use of Computer-Assisted Translation (CAT) tools is increasing, the incorporation and centralisation of translation memory (TM) systems for African indigenous languages, especially Xitsonga is still insufficiently advanced and researched. OmegaT CAT tool offers TM centralisation capabilities that allow translators to effectively share, and reuse translated segments. Even though that is the case, its implementation in

Xitsonga translation workflows encounters various obstacles. The introduction of Machine Translation (MT), Computer-Assisted Translation (CAT) tool, and Artificial Intelligence (AI) in the language landscape is a major development since technology is evolving and gaining momentum daily. Efforts to preserve languages are often multifaceted, involving documentation, educational programs, community initiatives, and increasingly the use of technology (Sharon, 2024). Different software programs have emerged in the language field to improve, speed up, enrich, and intensify language abilities. These computer applications are called Computer-Assisted Translation (CAT) tools, some refer to them as Computer-Aided Translation or Computer-Aided Human Translation (CAHT). CAT falls under an interdisciplinary field in language studies called Computational linguistics, the study of computer systems for understanding and generating natural language (Grishman, 1986).

The application of Computer-Assisted Translation (CAT) tools has progressed notably in multilingual settings, yet their efficient incorporation of African indigenous languages, such as Xitsonga, is still insufficiently examined, especially concerning the centralisation of translation memory (TM) in tools like OmegaT application. Current research on CAT tools in Africa has mainly centred on user adoption, interface design, and overall translation effectiveness with minimal exploration of the linguistic and technical functioning of centralised TMs in languages with fewer resources. A significant gap is present in the access to and use of Xitsonga linguistic resources, including aligned bilingual corpora and standardised terminology databases, which are crucial for creating and sustaining dependable translation memories. In the absence of these, centralised TM systems cannot guarantee accuracy, consistency, or terminological uniformity. Additionally, prior studies have not empirically investigated how Omega's TM centralisation addresses Xitsonga-specific linguistic features, dialectal differences, or culturally rooted expressions. A notable gap exists in assessing the quality of translations generated by centralised TM systems in Xitsonga.

Throughout the years, CAT tools have developed to provide numerous features in addition to Translation Memory. They have begun to incorporate terminology management, glossaries, machine translation integration, quality assurance checks, and collaborative features, along with various other functionalities. CAT tools have continued to evolve, integrating artificial intelligence and machine learning technologies to improve translation precision, increase efficiency, and streamline workflows for translators and localisation specialists

Over the years, several CAT tools were developed and now available. Some are free and some are paid. The free applications include OmagaT, Autshumato, MateCat, Wordfast (online), CafeTran, Anywhere, Espresso, and SmartCat. The paid ones include Wordfast (PC installation), XTM Cloud, SDL Trados Studio, Crowdin, MemoQ, Memsource, Wordfast, Wordbee, and déjà vu, amongst others. Flórez & Alcina (2011) declare that there is another difference between free and payment software that in many cases is not considered. Unlike the proprietary model, which views the user as a customer purchasing a completed product, users in the free software community are an essential part of the development process. One of the easiest ways to contribute to a free software project is by using the program and notifying the developers about any bugs encountered or features that are lacking (Flórez and Alcina, 2011).

In the South African context, Autshumato Machine Translation Web Service has been developed and introduced in the translation field by the Centre for Text Technology at the North-West University. Autshumato is an initiative supported by the Department of Sports,

Arts and Culture to create, launch, and maintain open-source translation technologies with the purpose of facilitating translation and ultimately enhancing information accessibility for all South Africans.

The initiative began in 2007, with the goal of aiding the development of documents in all official South African languages within the public sector. This was done by creating user-friendly open-source technologies that facilitate the translation process, encourage terminology standardisation, and reduce translation time. Additionally, the project encompasses machine translation (MT) systems. MT involves the automated translation of a given text from one language (like English) to another by a computer. The current available MT system capable of translating from English into Afrikaans, isiZulu, Sepedi, Xitsonga, Sesotho, and Setswana. This is a limitation of some sort; South Africa is a diverse multilingual country with twelve (12) official languages, including South African sign language. However, one of the primary objectives of creating Machine Translation Web Service (MTWS) was to promote multilingualism and access to information to all South Africans (Nemutamvuni, 2018).

The Autshumato Integrated Translation Environment (ITE) consists of various plugins that tailor the well-known OmegaT software. It offers a unified space featuring translation memory, machine translation (MT), and a glossary to aid in the translation procedure. While Autshumato ITE is designed for the eleven (11) official languages of South Africa, OmegaT operates independently of language, allowing translation between any two languages. However, the Department of Sport, Arts and Culture provided training for both Autshumato and OmegaT which caters for any language.

Literature Review

A literature review surveys books, scholarly articles, and any other sources relevant to a particular issue, area of research, or theory, and by so doing, provides a description, summary, and critical evaluation of these works in relation to the research problem being investigated (Fink, 2004). Many scholars have written and shared their views regarding CAT tools, including their advantages, importance, and benefits. Many scholars and researchers posit that machine translation systems or CAT tools (computer-assisted translation, computer-aided translation) are software that help to translate faster (Garcia, 2014).

Digital learning tools such as machine translation (MT) and creating online dictionaries can also contribute to preserving these languages. Each of these strategies offers benefits on how technology could be employed effectively and facilitate the preservation of indigenous languages. The use of technology is used to adequately preserve minority indigenous languages of South Africa; namely, Xitsonga, Siswati, Tshivenda, and isiNdebele. These languages do not receive much attention in terms of preservation using technology in South Africa (Mlambo and Matfunjwa, 2024).

The view mentioned above is supported by Sundani (2023) when he mentions that there is a connection between indigenous languages of South Africa and digital technologies, emphasizing the difficulties and possibilities in access, promotion, and preservation. Digital technologies can greatly support the promotion and preservation of South African indigenous languages. This can be credited to their multimedia functions, storage abilities, and communication tools. Nevertheless, restricted access to digital tools that aid these languages adversely affects their promotion and preservation. Barriers like insufficient

expertise, collaboration, fair digital services, and effective advocacy obstruct access to digital technologies for indigenous languages in South Africa (Sundani, Op cit.).

The lack of ICTs in indigenous languages reduces the opportunities for producing and accessing content in local languages on the internet. This influences the culture and knowledge that might be accessible to researchers keen on examining them. ICTs ought to be localised by translating and culturally adapting applications and software graphical user interfaces into native languages (Osborn, 2006).

Terminology management is an essential element in CAT environments as it encompasses the development, maintenance, and access of terms specific to a domain to guarantee precision and consistency. Integrated terminology databases within CAT software offer translators established vocabularies that minimise confusion and promote consistent word application across translations, especially in specialised areas like financial, legal, medical, and technical translation. CAT tools have translation memories that record previously translated phrases and terminology for future use, reducing repetitive work. OmegaT also provides customisable glossaries and potential translation suggestions, enhancing its utility for translators.

CAT tools not only act as improvements to productivity but also facilitate quality control, particularly in fields that demand high localisation standards such as software localisation projects that illustrate the benefits of CAT environments in handling intricate projects with various stakeholders (Yixin, 2024).

The primary purpose of CAT tools is to retain translated text segments for use in future translations. This technology is referred to as Translation Memory (TM), which serves as a database that keeps track of past work for potential reuse, along with enabling rapid searches through the existing content. The capability to reuse translations allows CAT systems to conduct a search in memory and provide reused matches, not just at the sentence level but also at the phrase and individual word levels, which is beneficial in documents that have repeated phrases (Barrachina et al., 2009).

CAT tools can acquire knowledge by assembling a vast collection of source texts alongside their corresponding translations, which need to be processed with a highly capable multiprocessor. When translating new texts, fragments from this collection are identified and selected to create the target text (Olohan, 2011).

Computer-aided translation refers to a method of translating from one language to another where a human translator utilises computer technology to enhance the process and carry out the translation (Bowker and Fisher, 2010). This is supported by Han (2020) when he mentions that in contrast to machine translation systems, utilising a CAT program necessitates the involvement of a translator in the translation process. It does not translate for you, but it allows you as a translator to fully carry out your responsibility.

Methodology

The researcher adopted a qualitative research approach in this article, that comprised desktop research of secondary materials. Desk research or secondary research is a pivotal method for analysing existing data and literature to generate actionable insights (Gupta, 2024). The researcher has chosen this methodology because data is readily available to understand

the phenomenon in order to achieve the aim of the research. The Promotion of Access to Information Act (PAIA) documents translated by different organisations were compared whereby terms were selected and compared for consistency against the standardised terminology. It is challenging for organisations and government entities to have identical content due to their differing mandates. The translations will always reflect the nature of the content or mandates they uphold, which vary from one organisation to another. The researcher has selected the Promotion of Access to Information Act (PAIA) document because very few organisations have translated it. This motivated the researcher to choose it for comparative analysis. When several translations are evaluated, it becomes easier to identify the accuracy and consistency of the terminology used in the different translations against standardised terminology. Comparing translation versions also helps detect spelling errors and evaluate the quality of the translations by examining the consistency of terms and translations. Terms that were appearing in all, if not most of the available OmegaT translation memories or glossaries, were selected to make comparisons easier.

Utilising the available data from the current translation memories and PAIA translated documents is both economical and efficient, as it eliminates the need for costly fieldwork or data gathering. This study indicates that conducting desktop research allows research conclusions to be shaped by the pre-existing terminology glossaries in the OmegaT translation tool and already translated PAIA documents, particularly for under-resourced languages like Xitsonga where empirical data is scarce. Comparative analysis was employed for data analysis. A comparative analysis is a side-by-side comparison that systematically compares two or more things to pinpoint their similarities and differences (Kaluza, 2023). The researcher opted for comparative analysis as it provides a systematic approach to assess and juxtapose various terminologies used by different translators in their OmegaT glossaries with the chosen translated terms in the PAIA documents, all in relation to the standardised terminology to derive insightful conclusions. Data or evidence was gathered from already translated texts (from English to Xitsonga) and OmegaT saved translation memories from Xitsonga for translators to verify the uniformity and/or variations of translation equivalents of the same term. Validity and reliability have been ensured during data collection through triangulation. Triangulation refers to using different data sources, investigators, and methods of data collection (Lincoln and Guba, 1985). The researcher uses two methods of collecting data, that is by comparing translated PAIA documents by different translators, and OmegaT glossaries or translation memories against the authenticated terminology respectively. The aim of this study is to explore the impact of centralised translation memory in OmegaT translation software in the case of Xitsonga language. To achieve the aim, the following research questions are explored:

- In what ways does centralised translation memory aid in the preservation and dissemination of terminology?
- How does centralised translation memory influence the overall quality of translations?
- In what manner does centralised translation memory enhance collaboration among translators?

Theoretical framework

Swanson (2013: 122) avers that the theoretical framework is the structure that can hold or support a theory of a research study. The Skopos theory and functional translation theories are among the most recognised approaches impacting translation studies. However, these theories did not consider the advancements in technology. Balkul (2016) pointed out that

technological elements were not adequately incorporated into established translation theories, which have yet to recognise the cognitive impact of technology.

To achieve the purpose of the research, the Corpus-based Translation approach is adopted. Corpus-based translation approach, often referred to as 'reference translation', comprises text and its corresponding translations in the target language (TL). The source language (SL) and the TL are connected, and their respective translations are derived through an extraction based on statistical models. The corpus encompasses electronically stored texts in a single language or multiple languages. It can be categorised into two distinct approaches: the statistical method and the example-based method. These can be better understood in this manner:

- The statistical approach focuses on bilingual text corpora and statistical models. A sentence from the source language can be translated in several ways.
- The example-based approach relies on a bilingual corpus to operate. Examples are fetched and selected based on their proximity in pairs of sentences. If there are no sufficiently close matches in the bilingual corpus, the example-based method may not successfully find the appropriate pair, leading to a low-quality result.

The corpus-based translation approach was incorporated into the comparative evaluation of translation memory (TM) within the OmegaT CAT tool to establish an empirical foundation for assessing the quality, consistency, and efficacy of preserved terminology. Translation memories aim to leverage previously translated segments, yet their effectiveness and dependability differ among various CAT tools. By utilising both parallel and monolingual corpora, the research was able to compare TM outputs with authenticated and standardised terminology, ensuring that the retrieved segments were not only precise but also contextually suitable.

Discussion/Analysis of Findings and Recommendations

The analysis of findings will be divided into part A and B.

Part A (PAIA translated documents)

Data was obtained through evaluating the already translated texts from English to Xitsonga to check variations and uniformity of translation equivalents of the same terms. The four translations of Promotion of Access to Information Act (PAIA) document from four translators were evaluated and compared. They were analysed using a side-by-side comparison that systematically compares two or more things to pinpoint their similarities and differences, hence the comparative analysis. The two translations were from two government departments and two from state owned entities. Pseudonyms were used for privacy purposes, these documents were named document A, B, C, and D. The Promotion of Access to Information Act document was chosen because even though every department or state-owned entity has customised their own documents according to their mandates and duties, some information will remain formal or similar. The researcher wanted to check similarities and differences as far as terminology is concerned. These translated documents were readily available online for download. Only one document (PAIA) was chosen because almost every department or state-owned entity has translated it, the researcher did not want to evaluate or compare different documents. In these documents, some words and terms, or phrases were selected for the purpose of this study. Different published Xitsonga

terminology lists were also checked to verify if the selected words or terms are available on those lists. Those are the terms that have been authenticated by the Xitsonga National Language Body of the Pan South African Language Board, and they are considered to be standardised terms. Below are the selected words or terms and the authenticated ones:

Table 1: indicates the extracted terms from the PAIA documents translated by different translators to check consistency and spelling of the terminology used against the standardized or authenticated terminology (The researcher's own collection)

Term	Document A	Document B	Document C	Document D	Authenticated / Standardised Term
data	datara	switiviwa	switiviwa	data	switiviwa/deyitha
prescribed (fee)	phirisikirabiweke	vekiweke	vekiweke	vekiweke	–
contract (law)	kontiraka	kontiraka	kondiraka	kontiraka	kontiraka
strategy	xitirateji	kungu	–	qhinga	qhinga
processed	purosesa	phurosesa	phurosesa	phurosesa	–
printed	pirintiweke	kandziyisiweke	kandziyisiweke	kandziyisiweke	–
deposit	diphoziti	–	diphoziti	dipoziti	dipoziti/diphoziti
electronic	xielekitironiki	xilekitironiki	xielekitironiki	xitironiki/ xielekitironiki	xiilekitironiki/ xitironiki/ xielekitironiki (not a stand-alone term)
third party	thedi phathi	vandla ra vunharhu	vandla ra vunharhu	munhu wa vunharhu	ntlawa wa vunharhu
Information officer	muofisiri wa mahungu	muofisiri wa vuxokoxoko	muofisiri wa vuxokoxoko	muofisiri wa mahungu	not a stand-alone term, information - mahungu/officer- muofisiri

Data

The term data denotes information or details concerning a specific topic. Document A provided a translation equivalent of datara; however, despite adhering to the appropriate consonant-vowel pattern in localising the term, it was still deemed inaccurate according to the authenticated term. Above all, the translation equivalent datara was made to conform incorrectly to the spelling and orthography rule from Xitsonga Spelling Rules and Orthography (2019: 61), stating that:

> *Marito lama helaka hi ntwariso wa /-**ter**/, /-**ner**/, /-**dar**/, /-ure/ na man 'wana lamo tano eka tindzimi to fana na Xinghezi ya fanele ku hela hi /-**ra**/ loko ya hundzuluxeriwa eka Xitsonga.*
>
> (Words that end with the /-**ter**/, /-**ner**/, /-**dar**/, /-**ure**/ sounds and others similar to those ones in languages such as English must end with the sound, /-**ra**/ when translated into Xitsonga). The term data ends with a suffix or sound of /-**ta**/ not, /-**ter/,** /-**ner**/, /-**ure**/, so it is incorrect to translate it into *datara.*)

However, there are some exceptions involved when dealing with the above-mentioned rule, some words/terms are not able to conform and take the sound /-**ra**/. This is substantiated by Xitsonga Spelling Rules and Orthography (2019: 61) when it avers that:

> *XIYA: Swi nga endleka man'wana marito ya nga pfumeli ku tirhisa xilandzi xa /-**ra** / xa nawu* lowu, xk. collar > *kholoro*; spanner > *xipanere*; archar > *acha*.
>
> (NOTE: It is possible that some words may not conform to the rule of adopting the suffix /-ra/, e.g., collar > *kholoro*; spanner > *xipanere*; archar> *acha.*)

Translators require knowledge and understanding of the linguistic mechanisms of word-formation processes (Baloyi, 2023). Xitsonga language translators and Xitsonga speakers should familiarise themselves with the new spelling and orthography. The *banginkulu** saga would have been avoided if people knew the spelling rules in their language. Inconsistencies like these in the language hinder the ongoing terminology development efforts and drag down the quality of the translated documents. However, implementing a centralised translation memory might come with challenges with issues like:

- Management and version control: inquiries might emerge regarding who oversees the centralised TM, who modifies entries, and how to handle version control.
- Standardisation of terms: Consistent formats, spelling and orthographies, and metadata are essential for translation memories. Xitsonga, just like any African language, frequently experience variations in spelling and a lack of standardised terminology, complicating harmonisation efforts.
- Inter-institutional coordination: Language units, universities, government departments, and entities may use different CAT tools or workflows. PanSALB through Department of Sport, Arts and Culture (DSAC) must regulate this.
- Technical Infrastructure: Centralised translation systems need dependable servers, safe storage, and consistent internet. Numerous institutions continue to encounter bandwidth restrictions, obsolete systems, or a deficiency of cloud-based solutions customised for African environments.
- Privacy, ownership, and access rights: Who possesses the translations of the terms? Should every independent translator, government department or entity, educational institution, given open or complimentary access? Sharing is sometimes complicated by sensitive or copyrighted content. However, DSAC, as the recognised custodian, can manage this.

Contract

Documents A, B, D and the verified list (the standardised term) translated the contract as *kontiraka*, while Document C rendered it as *kondiraka*. The equivalent provided by Document C is inaccurate when considering the stipulation of the rule in the Xitsonga Spelling Rules and Orthography (2019: 61) which proposes that:

> *Malombiwa lama cinciweke mimpfumawulo ya twarisiwa hi sisiteme ya Xitsonga ya tsariwa hi laha ya twarisiwaka hi Xitsonga.*
>
> (Words taken from other languages that have been altered for pronunciation according to the Xitsonga system must be spelled as they are pronounced in Xitsonga.)

To prevent the creation of a new term, it was intended for the term to be pronounced as it is in English. The writing system needs to be considered and different terminology lists should be referenced to ensure the quality of our translations is maintained. Nonetheless, reviewing every single terminology list can be a lengthy process.

Strategy

The verified term for strategy appears in one of the terminology lists as *qhinga*, and Document D aligns with the verified term. Document A, on the other hand, has rendered it into *xitirateji*, while Document B has translated it into *kungu*. Why must we transliterate a word when we already have an equivalent in our language? *Xitirateji* adheres to the spelling, and the writing system is accurate (Consonant + Vowel), yet it is a wrong equivalent and negatively impacts the quality of our translations, inviting criticism.

Prescribed

In Document A, the term prescribed is translated as *phirisikirabiweke*, while Documents B, C, and D use the equivalent translation *vekiweke.* No standardised or verified term exists in the terminology lists; however, Document A chose transliteration, whereas Documents B, C, and D depended on the term's context. Although no standardised term exists, the translator should have explored each option and used the accurate equivalent by consulting the terminology list provided. This approach wastes time due to uncertainty about which list to reference, relying solely on the availability of equivalents in the translation memory and glossary of the CAT tool, OmegaT.

Processed

Documents B, C, and D possess the same translation equivalent *phurosesa,* whereas Document A contains *purosesa*. The verified term is not easily accessible, yet this is a straightforward word to translate. The term *purosesa* does not adhere to the Xitsonga spelling and orthography rules. The identical guideline that applies to the term "contract" should also be applied to this term. Borrowed words altered for pronunciation according to the Xitsonga system should be spelled as they sound in Xitsonga. In summary, translators need to have a strong understanding of word formation processes and translation strategies.

Printed

The strategy applied to this word is transliteration. Document A chose to transliterate the word printed as *pirintiweke,* whereas the other documents used the translation equivalent *kandziyisiweke*. The official terminology lists do not contain the standardised term. The translation used here is inaccurate, given that the equivalent is relatively easy to find. This could be due to the fact that sometimes translators do not put more effort in what they, as a result, poor quality translations are created.

Deposit

Documents A and C provided the same equivalent of *diphoziti*, while Document D offered an equivalent of *dipoziti*. Two verified terminology lists include *dipoziti* and *diphoziti*. This creates a disparity in translations because the two are validated in separate terminology lists and are regarded as distinct terms rather than synonyms. PanSALB should allow board members to gather occasionally to align the verified terms with similar meanings to prevent problems like these. Possessing various translation options for the identical term in the same context is problematic.

Electronic

Documents A and C have rendered the word electronic as *xielekitironiki*, while Document B has rendered it as *xilekitironiki,* and Document D has used *xitironiki* and xielekitironiki. In certain validated terminology lists, the term appears as an independent term, but it is noted as *xiilekitironiki, xitironiki*, and *xitironiki.* Because the terminology lists cannot maintain a consistent word or synonymous terms, this may lead to discrepancies in translations and issues with quality.

Third party

The translation equivalent for the term third party in Document A is *thedi phathi*.Iin Documents B and C, it is *vandla ra vunharhu*; in Document D, it is *munhu wa vunharhu*, and for the authenticated term, it is *ntlawa wa vunharhu*. From the comparison of the provided translation equivalents, it is clear that each translator offered an equivalent they believe is appropriate. Nonetheless, the differences between the equivalents *vandla ra vunharhu* and *munhu wa vunharhu* might be a matter of context. Thus, centralised translation memory must include terms with all translation equivalents based on various registers to accommodate contextual scenarios. Such factors lead to inconsistencies in translations and diminish their quality.

Information Officer

The standardised terminology list does not include the term 'information officer' as an independent term; however, the verified term for information is *mahungu* and for officer it is *muofisiri*, indicating that the correct term should be *muofisiri wa mahungu.* Documents B and C translated information to *vuxokoxoko*, therefore they possess *muofisiri wa vuxokoxoko. Vuxokoxoko* and *mahungu* can sometimes be used interchangeably in certain situations based on the context; translators must be cautious when addressing contextual considerations in translations.

Part B (OmegaT translation memory/glossaries)

Data were gathered by assessing OmegaT glossaries or translation memories from various Xitsonga translations to analyse discrepancies and consistency of translation counterparts for identical terms. Regarding glossaries, the terms that will be included are at the translator's discretion. Various translators will possess distinct terms in their glossaries,

as what holds significance for one translator might not be significant for another. In this context, it is essential for translators to access each other's glossaries and translation memories. In this manner, translators will gain access to terminologies from various fields and prevent the duplication of terms that have already been established. The researcher analysed four OmegaT glossaries created by various translators from both the private and public sectors. Pseudonyms were utilised here for confidentiality reasons, referred to as Glossaries A, B, C, and D. Below are the chosen terms along with the verified terms for the chosen expressions:

Table 2: indicates the extracted terms from the OmegaT translation memories or glossaries by different translators to check consistency and spelling of the terminology used against the standardized or authenticated terminology

Term	Glossary A	Glossary B	Glossary C	Glossary D	Authenticated Term
state capture	lawuriwa ka mfumo hi mavandla ma le tlhelo	nhlohlotelo wa le handle wa mfumo	–	–	–
vaccine	nsawutiso	xisawutisi	–	–	nsawutiso
non-compliance	nkavulandzelelanawu	ku nga landzeleriwi ka milawu	–	–	nkavulandzelela-wu
e-hailing services	–	vukorhokeri bya mathekisi lama vitaniwaka eka inthanete	–	–	–
disclaimed audit opinions	–	mihlahluvo ya oditi leyi oditara yi nga nyikiki mavonelo	–	–	–
disclaimed and adverse opinions	–	mihlahluvo leyi oditara yi nga nyikiki vonelo na leyi tirhekodo na switatimente swi andlariweke hindela leyi hambanaka na milawu ya tinkota	–	–	–
load-shedding	ntirhiso wa gezi hi ku siyerisana	–	–	–	–
terror financing	mali yo seketela vutherorisi	–	–	–	–
catfishing	kanganyisa vuwena	–	–	nkangayiso	–

State capture

Glossary A offered a translated equivalent for state capture as *ku Lawuriwa ka Mfumo hi Mavandla ma le Tlhelo*, while Glossary B rendered it as *Nhlohlotelo wa le Handle wa Mfumo*. Glossaries C and D, along with the authenticated term, lack translation equivalents or terminology. The lack of the translated term in Glossaries C and D may indicate that the

translators have never come across the term previously or it was not significant enough for them to retain or include it in their glossaries. The lack of presence in the standardised terminology list indicates that it has not been included for development; however, it is essential to welcome new words into the corpus and offer definitions for these terms to support the evolution of language and terminology. State capture is a relatively new and intricate concept; the two translators choose to employ paraphrasing as a translation approach, since finding a single term is quite difficult. Due to the lack of access each translator has to other translators' glossaries, the term will still end up with multiple translation equivalents or terms for a particular term. This undermines the quality of translations and leads to inconsistencies in terminology.

Vaccine

Glossary A offered a translation equivalent for vaccine as *nsawutiso*, whereas glossary B provided a translation as *swisawutisi*. The other glossaries did not offer any equivalent. The authenticated list offered a parallel as *nsawutiso*. The two translation equivalents differ in their structure; one is in uncountable form (*nsawutiso*) while the other is in countable form, yet both terms essentially convey the similar meaning. Nevertheless, because the standardised translation equivalent exists, translators must adhere to it. It can be exhausting at times to look for a term in terminology booklets or to search for it in a terminology list on a computer or online. As translators utilise OmegaT for their work, the term would be easily obtainable in a centralised translation memory or glossary, resulting in saved translation time and shorter turnaround times for translation projects.

Non-compliance

Glossary A provides the translation equivalent for non-compliance as *nkavulandzelelanawu*, while Glossary B offers an equivalent translation of *ku nga landzeleriwi ka milawu*. The other glossaries failed to offer translation equivalents. Nevertheless, the authorised terminology corresponds to *nkavulandzelelanawu*, which is the recognised term. The translation equivalent of Glossary B is identical to Glossary A and the verified term; nevertheless, Glossary B chose to paraphrase the term, providing an explanation of it which is acceptable. The joint work of the consultative meetings, verification, and authentication process by the PanSALB facilitated translators in creating a strong term instead of needing to devise an explanation.

E-hailing services

Glossary B rendered e-hailing services as *vukorhokeri bya mathekisi lama vitaniwaka eka inthanete*, while the other glossaries offered no translation equivalent. There is no standardised translation equivalent for the term; therefore, the official term is missing. This term is quite recent in the vocabulary, it was introduced when this mode of transport began. Its lack in other glossaries may be due to being new and having less content requiring translation; only those in the transport sector might have encountered it. The equivalent provided by Glossary B restricts the transportation to a specific kind of transport, namely taxis, whereas e-hailing services can encompass various types of transport, including vans, sedans, or even trucks, though the intended message is clear. Collaboration might

be beneficial in this case, translators can share their ideas and subsequently establish a standardised term, conserving time and preventing the need to create new terminology.

Disclaimed audit opinions

The translation equivalent provided for disclaimed audit opinions by Glossary B is *mihlahluvo ya oditi leyi oditara yi nga nyikiki mavonelo*. The other glossaries containing various authenticated terms or terminology lack the equivalents. The translator for Glossary B attempted to paraphrase but the equivalent appears too lengthy, and there was little the translator could accomplish on their own. Accounting terminology consistently presents challenges for translators. If the translators could be observed by fellow translators, it is thought that others could have enhanced it by utilising available translation strategies and word formation strategies. This approach would allow translators to adhere to the superior translation and simplify its validation, as no new term would surface. This would remove superfluous duplication and redundancy.

Disclaimed and adverse opinions

Glossary B is the only one that succeeded in finding a translation equivalent, meaning, *mihlahluvo leyi oditara yi nga nyikiki vonelo na leyi tirhekodo na switatimente swi andlariweke hindela leyi hambanaka na milawu ya tinkota.* A reader cannot fail to comprehend this in a translated text. Due to the absence of a suitable equivalent, the translator opted for paraphrasing, and there is always potential for enhancement since the term is excessively lengthy and transformed into a sentence. If the OmegaT translation memory of glossary was automatically accessible to other translators without the need for manual sharing, they could examine the contents of Glossary B and attempt to enhance the equivalent; translation necessitates creativity, collaboration, and teamwork. Translators do not operate in isolation for the same reason. Translators work together to create new terminology and enhance the existing content in the corpus. A term can have one meaning today and another tomorrow, as times are changing. Fresh ideas, innovation, and experience are consistently required in translations.

Load-shedding

Glossary A proposed a translation equivalent for load-shedding as *ntirhiso wa gezi hi ku siyerisana*. The term is missing in the other glossaries, including the authenticated terminology. Translators determine what to add and what to omit in their glossaries. If other translators could view what Glossary A generated, they could easily create a new term to replace this one or add a new synonym. Primarily, translators may retain the established term if they find it acceptable. In this manner, they conserve time instead of squandering it and overthinking without reason and maintaining consistency in translations, thus enhancing quality.

Terror financing

Glossary A translated terror financing as *mali yo seketela vutherorisi*. The other glossaries did not include the term, resulting in the lack of translation equivalents. Its appearance

in Glossary A may indicate that the translator primarily works with financial terminology or that the term holds significance based on the judgment. Developing terminology for financial terms is quite challenging. In an effort to close this gap, the former Financial Service Board Multilingual (FSB), now known as The Financial Sector Conduct Authority (FSCA), developed its Financial Terminology List released in 2017. The financial sector is extensive and so is its vocabulary; thus, additional terminology development projects are required to address its wide range of terms.

Catfishing

Glossary A is the sole source that provided a translation equivalent for *kanganyisa vuwena*, and a paraphrasing translating approach was employed in rendering the term. Glossary D rendered it as *nkangayiso*. The two counterparts differ in all respects; the aspect of inconsistency is highlighted here. If an individual encounters both translations in a document, they might be puzzled because the term was rendered differently, possessing two distinct meanings. *Nkanganyiso* in an acknowledged definition referring to fraud. The phrase is recent as the age of social media has developed and accelerated over time. The other glossaries omit the term, suggesting that translators have never encountered it or that it has not affected them in any manner. The introduction of new terms should be acknowledged and incorporated during the development and creation of terminologies. Acknowledgment must also be made to the Limpopo Department of Sport, Arts and Culture, which has initiated its consultation procedures concerning the creation of social media guidelines in Xitsonga, Sepedi, and Tshivenda.

Recommendations

For language to grow, as new words frequently enter language space, terminologies must be developed, maintained and preserved. The centralisation of the OmegaT glossary or translation memory should be embedded to the General Theory of Terminology (GTT) by Eugen Wuster, the main proponent. His core idea regarding the theory is based on removing uncertainty from technical languages through standardising terminology to transform them into effective communication tools to persuade all users of technical languages regarding the advantages of consistent terminology, and to define terminology as a field for all practical aims and to elevate it to the level of a science (Wüster, 1968). The glossary shown in the translation memories within OmegaT software is at the translators' discretion and it is improbable that the terms one translator possesses differ from those of other translators. In light of this context, it is suggested that OmegaT should include a sharing feature or a repository for storing translated terms. Post-translation, the translator will enable the sharing feature to store the document in a central location. Activating the sharing feature will make the document containing the terms accessible in the OmegaT glossary for any translator who opens a document with the same terms. In this manner, translation memories and glossaries are centralised for all translators to access, preventing inconsistencies in translation terms and equivalents. In dictionary auto-search, CAT tools emphasise terms within the text recognised by the terminology management system (Krüger, 2016). Should the document be confidential, the translator will refrain from enabling the sharing feature ensuring that only they have access to the document for privacy reasons. This approach will preserve translation quality, save time, and foster collaboration. They also need to ensure that the available authenticated terminology lists are loaded into a centralised glossary for

easy access to prevent the duplication of terminologies already created for various fields or subjects. The recommendation stated above is the responsibility of the Department of Sport, Arts and Culture, which is tasked with developing, implementing, and promoting the Computer Assisted Translation tool. They need to ensure that all translators adopt OmegaT for translation by organizing additional awareness workshops and mandating its use for every translator. Requests for OmegaT Enhancements (in English) can be submitted at the SourceForge site: https://sourceforge.net/p/omegat/feature-requests/. Translators must familiarise themselves with the linguistic mechanisms of word formation processes, spelling rules, and orthography, as language is dynamic rather than static. The recognition and inclusion of new terms is essential in the development and creation of terminologies, advancing the objectives of terminology management.

Conclusion

A CAT tool is a crucial element of ICT that translators need to use to manage and preserve languages through terminology development and management. The article examined the influence of centralised translation memory and glossary in OmegaT translation software, focusing on the Xitsonga language. Centralising the translation memories and glossaries will reduce inconsistencies, redundancies, streamline translation time, and result in high-quality translations. Not only will it improve the quality of the translations, but it will also promote collaboration among translators and facilitate the sharing of information regarding terminology learning. However, certain translators possess the software on their devices but are not utilising it, they should reassess and employ this tool to ensure preservation and maintain quality in their everyday translations. The researcher wanted to evaluate more OmegaT glossaries and documentation but was limited by the fact that some translators were not comfortable sharing their glossaries with the researcher. Some translators are not utilising any CAT tool in their translations, everything is done manually. To find the same document translated by a large number of translators is hard since the content that is translated is guided by the organisation's mandate or subject matter. This subject matter is under researched especially in indigenous languages or low resourced languages, a lot can still be done for future research. Future studies may focus on exploring how CAT tools can effectively share translation memories, standardise and validate translation equivalents across various organisations, and how to manage quality control in stored glossaries while removing outdated terminology from the CAT tool.

References

Autshumato. (n.d.). *The Autshumato MT Web Service*. Retrieved June 20,2025, from https://mt.nwu.ac.za/#.

Balkul, H. I. (2016). 'Translation technologies: A dilemma between translation industry and Academia', *Online Submission*, 4(4), 100-108.

Barrachina, S., Bender, O., Casacuberta, F., Civera Saiz, J., Cubel, E., Khadivi, S., Lagarda A. L., Ney, H., Tomas, J., Vidal, E., & Vilar, J. (2009). 'Statistical approaches to computer-assisted translation', *Computational Linguistics*, 35(11), 3-28. https://doi.org/10.1162/coli.2008.07-055-R2-06-29

Fink, A. (2014). *Conducting research literature reviews: From the internet to paper* (4th ed.). Thousand Oaks, CA: SAGE.

Flórez, S., & Alcina, A. (2011). 'Catálogo de software libre para la traducción', *Revista Tradumàtica*, 9, 57–73. https://doi.org/10.5565/rev/tradumatica.5

Folaron, D. A. (2010). Translation tools. In Y. Gambier & L. van Doorslaer (Eds.), *Handbook of translation studies* (Vol. 1, pp. 429–436). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/hts.1.tra9

Garcia, I. (2014). Computer-aided translation: Systems. In I. Garcia (Ed.), *Routledge encyclopaedia of translation technology* (pp. 106–125). London, New York, NY: Routledge.

Grishman, R. (1986). *Computational linguistics: An introduction*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511611797

Gupta, R. (2024). 'Methodological and theoretical rigor in desk research', *ResearchGate*, https://www.researchgate.net/publication/386642850_Methodological_and_Theoretical_Rigor_in_Desk_Research.

Kaluza, J. (2023). *Comparative Analysis*. Retrieved July 10, 2025, from https://dovetail.com/research/comparative-analysis/

Han, B. (2020). 'Translation, from pen-and-paper to computer-assisted tools (CAT tools) and machine translation (MT)', *Proceedings*, 63(1), 56. https://doi.org/10.3390/proceedings2020063056

Krüger, R. (2016). 'Contextualising computer-assisted translation tools and modelling their Usability', *Journal of Translation and Technical Communication Research*, 9, 114–148.

Lincoln, YS. & Guba, EG. (1985). *Naturalistic Inquiry*. Sage Publications: Newbury, Park, CA. https://doi.org/10.1016/0147-1767(85)90062-8

Mahdi, H. (2018). 'A review of literature of computer-assisted translation', *ResearchGate*, https://www.researchgate.net/publication/328019119_A_Review_of_Literature_of_Computer-Assisted_Translation.

Mlambo, R. & Matfunjwa, M. (2024). 'The use of technology to preserve indigenous languages of South Africa', *Literator*, 45(1), 1–8. DOI: https://doi.org/10.4102/lit.v45i1.2007

Nemutamvuni, M. E. (2018). Investigating the effectiveness of available tools for translating into Tshivenda. Unpublished M. A. dissertation. University of South Africa, Pretoria.

Olaare, S. (2024). 'The role of technology in language preservation', *European Journal of Linguistics*, 3, 44–56. https://doi.org/10.47941/ejl.2046

Olohan, M. (2011). 'Translators and translation technology: The dance of agency', *Translation Studies*, 4(3), 342–357. https://doi.org/10.1080/14781700.2011.589656

OmegaT. OmegaT (n.d). *Translation Tool Overview*. Retrieved September 15, 2025, from (https://autshumato.nwu.ac.za/software/.

Osborn, D.Z. (2010). *African languages in a digital age: Challenges and opportunities for Indigenous languages computing*. Human Sciences Research Council: Cape Town.

Sundani, N.D. (2023). 'South African Indigenous languages and digital technologies: Access, promotion and preservation', *International Journal of Social Science Research and Review*, 6(8), 41–51. https://doi.org/10.47814/ijssrr.v6i8.1385

Swanson, R. A. (2013). *Theory building in applied disciplines.* San Francisco, CA: Berrett-Koehler.

Wang, Y. (2024). 'The impact of technology on human translators and translation quality: A study on machine translation and computer-assisted translation tools', *English Linguistics Research*, 13(1),19. https://doi.org/10.5430/elr.v13n1p19

Wüster, E. (1968). *The Machine Tool. An Interlingual dictionary of basic concepts*. London: Technical Press.

The Incorporation of Shikomori to Improve ICT Comprehension, Access, and Uptake by Comorian Communities

Issa Goba Roukiyat
Social Sciences of Development and Sustainable Tourism
University of Antananarivo
Antananarivo, Madagascar
roukiyatissa@gmail.com

Abstract

Despite the transformative potential of ICT, the exclusion of Shikomori from digital platforms significantly undermines digital inclusion in the Comoros. This paper explores the incorporation of Shikomori, the indigenous language of the Comoros Islands, into Information and Communication Technologies (ICT) to improve understanding, accessibility, and uptake within Comorian communities. Linguistic differences are acknowledged as a significant impediment to the adoption of ICT. It is posited in this discourse that by leveraging the familiarity and cultural relevance of Shikomori, the digital divide may be bridged and inclusive technological advancement fostered. In the Comorian context, the digital divide refers to the gap between not only urban and rural populations but also young and old users' access to affordable and reliable internet and digital technologies, as compounded by limited digital skills and infrastructure. Digital transformation, on the other hand, focuses on expanding digital access and services, such as e-government platforms and quality internet networks to overcome this divide and foster socio-economic development. This study examines the development and implementation of Shikomori-localised ICT solutions in four key sectors; namely, tourism, administrative documents, education, and national telecommunications. In the tourism sector, the study examines the development of multilingual mobile applications that provide tourist information and navigation in Shikomori, thereby improving accessibility for local guides and community-based tourism initiatives. In the context of administrative documentation, it explores the digitisation of forms and documents in Shikomori with the objective of enhancing bureaucratic processes and fostering greater citizen engagement. In the field of education, the analysis of educational software and online resources in Shikomori is undertaken with a view to enhancing students' comprehension and knowledge retention. Moreover, the study investigates the implementation of Shikomori voice interfaces and text messaging services in national telecommunications to improve accessibility for individuals with limited literacy or fluency in official languages. The study integrates a linguistic-cultural relevance framework, postulating Shikomori integration as essential for achieving inclusive digital transformation with methodologies including linguistic analysis, user-centered design, and participatory development to ensure the solutions' effectiveness and cultural appropriateness. The findings emphasise the significant potential of Shikomori integration to democratise access to ICT, promote digital literacy, and empower Comorian communities.

Keywords: Comoros Islands, ICT Solutions, Local Communities Inclusion, Shikomori

Introduction

The considerable influence of Information and Communication Technologies (ICT) provides opportunities for socio-economic development and societal transformation on a global scale. Nevertheless, the equitable distribution of these benefits is constrained by linguistic barriers, particularly in regions where indigenous languages are marginalised within the digital sphere. In the Comoros Islands, the limited integration of Shikomori, the nation's indigenous language, into ICT solutions presents a significant impediment to widespread digital comprehension, access, and uptake, thereby marginalising a substantial portion of the population with limited proficiency in official languages.

A central premise is that, in today's society, the ability to access, adapt, and create knowledge using information and communication technologies is critical to social inclusion. This focus on social inclusion shifts the discussion of the "digital divide" – from gaps to be overcome by providing equipment – to social development challenges to be addressed through the effective integration of technology into communities, institutions, and societies (Warschauer, 2003: 5).

The Comoros archipelago, influenced by Bantu, Arabic, and Swahili cultures, exhibits a distinctive linguistic landscape where Shikomori, a vernacular language, has a notable presence in daily life despite French and Arabic being the official languages. The Comorian language is called Shimasiwa by its speakers, a word that means 'language of the archipelago'. It is currently known as Shikomori, which means 'language of the Comoros' (Ahmed-chamanga, 2011: 19). Shikomori encompasses four main dialects corresponding to the archipelago's primary islands. This analysis focuses on the dialects prevalent in the Union of the Comoros; namely, ShiNgazidja (spoken on Grande Comore and often considered the standard and official compared to the other dialects due to Moroni as capital and center of Comoros), ShiNdzuwani (Anjouan), and ShiMwali (Mohéli). Though these dialects differ slightly, mainly in vocabulary (e.g., 'moja' and 'montsi,' both meaning 'one'), they are mutually intelligible, enabling seamless inter-island communication. This linguistic interconnectedness is crucial in developing ICT solutions.

Understanding these dialectal nuances allows for designing adaptable digital platforms tailored to each island community's linguistic preferences, ensuring broader adoption and effectiveness. Despite the presence of variations, the shared linguistic root presents a significant opportunity for the development of core ICT solutions that can be readily localised or adapted for specific dialects. This analysis focuses on Shikomori's current integration in ICT solutions within the Comoros, emphasising challenges encountered and the potential benefits of its broader adoption to enhance understanding, accessibility, and adoption within Comorian communities. The research question guiding this study is: How can the strategic integration of Shikomori into ICT effectively bridge the digital divide, foster digital literacy, and empower local communities in the Comoros Islands? It is argued that such linguistic inclusion is not just a convenience, but a fundamental prerequisite for achieving genuine digital transformation and sustainable development in the Comoros.

In order to examine this question, this research uses a multi-faceted methodology, including a comprehensive review of relevant literature, a socio-linguistic analysis of the Comorian context, and the application of user-centered design and participatory development principles to ensure the cultural appropriateness and effectiveness of proposed solutions.

This paper is structured to address the following aims: to examine the current use of Shikomori in ICT and national media, emphasising its underrepresentation in key digital domains; to analyse the principal challenges hindering the broader integration of Shikomori into ICT solutions with critical barriers such as the lack of technical expertise, insufficient funding, limited governmental support, and the absence of a standardised writing system being identified; and lastly, to propose concrete recommendations and actionable ICT solutions leveraging Shikomori to enhance accessibility, promote digital literacy, and empower local communities.

The following solutions are illustrated through specific examples of Shikomori integration across four vital sectors: tourism (e.g., multilingual mobile applications); public administration (e.g., digitised forms and documents); education (e.g., educational software and online resources); and national telecommunications (e.g., voice interfaces and texting services). The anticipated findings of this research emphasise Shikomori integration's potential to democratise ICT access, cultivate digital literacy, and significantly empower Comorian communities. Overall, this study aims to provide a comprehensive roadmap for leveraging Shikomori as a vital bridge toward a more equitable, inclusive, and technologically advanced future for the Comoros, contributing notably to global discussions on language technology and digital inclusion in developing nations.

Methodology

The study employs a mixed-methods research design, combining qualitative and quantitative approaches to comprehensively investigate the integration of Shikomori into ICT solutions in the Comoros Islands. The methodology is primarily descriptive and exploratory, with the objective of mapping the current landscape, identifying barriers, and proposing actionable solutions. The research process is structured into three interconnected phases: a literature review and documentation analysis, primary data collection through surveys and semi-structured interviews, and a final data analysis.

Documentation

The initial phase of the project entailed an extensive review of academic literature, including peer-reviewed journals, conference proceedings, and scholarly books. As reported by peer researchers, this specific review focused on several themes: indigenous language technologies, digital inclusion, language revitalisation through ICT, and the socio-linguistic dynamics of multilingual digital environments (Galla, 2018). Concurrently, an environmental scan of the Comorian digital ecosystem was conducted. This involved the analysis of publicly available documents and reports from a variety of national and international organisations (e.g., UNESCO, local NGOs, and companies), as well as existing online platforms (websites, social media groups, and mobile applications). The aim of this phase was to identify key stakeholders, prevailing languages in digital spaces, and the current ICT infrastructure, thereby establishing a robust theoretical framework, understanding global best practices, and contextualising the specific challenges and opportunities within the Comoros.

Primary Data Collection

Primary data was systematically collected through a combination of online surveys and semi-structured interviews, targeting diverse user groups and key stakeholders.

Online Survey

An online survey was designed and disseminated to gather quantitative and qualitative insights from local ICT users across the Comoros Islands. The survey's objective was to assess three key areas: (1) access to and usage of ICT among Comorian communities, (2) the need for Shikomori-localised ICT solutions, and (3) Shikomori usage in informal digital communication. The survey utilised a combination of multiple-choice, Likert scale, and open-ended questions to capture both statistical trends and nuanced user perspectives. The dissemination of the survey was conducted through local community networks, educational institutions, and social media platforms to ensure comprehensive reach and inclusive participation.

Semi-Structured Interviews

Semi-structured interviews were conducted with a purposive sample of key stakeholders and experts to gain in-depth qualitative data. The following representatives were included in the study:

- Linguistic and Cultural Organisations: Prominent examples include a linguist and cultural advocate, a Comorian teacher at INALCO, founder of ORELC, and Collective Shikomori (community initiatives promoting the language).
- ICT professionals and developers: These interviews aimed to establish the technical feasibility of the project considering the challenges with developing Shikomori-based applications.
- Sectoral experts: This group included officials from the Ministry of Tourism and the National Office for Tourism of Comoros, as well as educators, online course providers, and representatives from national telecommunications companies. In-depth interviews were conducted within four key sectors (tourism, administrative services, education, and telecommunications) focusing on identifying specific linguistic barriers within their operations, exploring existing informal adaptations, and measuring the potential and requirements for formal Shikomori-integrated ICT solutions.

The questionnaire was administered in two phases. Firstly, it was distributed in French and English. Secondly, the same questionnaire was redistributed, but this time it was translated into Shikomori. This approach was necessary to test the hypothesis that Comorians would be more willing and committed to answering questions in a language they are familiar with. However, it is important to note that the Comorian version of the questionnaire incorporated vocabulary from French or Arabic due to a lack of total mastery of writing, which is still a matter of debate within the oral user community itself.

Ethical considerations, sampling, and recruitment

The study was conducted in accordance with ethical standards, with informed consent obtained from all 60 participants from the Shikomori communities and confidentiality was ensured by anonymising all data. Participants were selected based on their professional roles and expertise in linguistic, technological or sector-specific fields relevant to the study; namely, in tourism, education, administration, and telecommunication. To gain deeper qualitative insights, purposive sampling was employed for semi-structured interviews with experts in linguistics and technology. The interviews used flexible thematic prompts or avenues for reflection to encourage open and honest dialogue, enabling participants to share insights that went beyond the preset questions and deeper highlight on issues that we might have overlooked such as the political role of a language in the case of administrative conflict between the Comoros (with Shingazidja, Shimwali, and Shindzuwani) and Mayotte (with Shimaore).

In addition, local university students were engaged as co-researchers in a focus group to help collect data from their peers and family members who were mostly Shikomori speakers. Meanwhile, an online survey was distributed through community networks and social media using convenience sampling to reach participants with varying technological skills, ages, and levels of education. A stratified sampling method was opted for, as the population was divided into four distinct subgroups based on the tourism, public administration, education, and telecommunications sectors. The questionnaire was distributed to voluntary agents within these subgroups, ensuring representativeness across all four sectors.

The survey questions were co-designed with students and linguistic experts to reflect user-centred design principles and ensure that the voices of end users were included. While the sample leaned towards younger, tech-savvy users, efforts were made to include all social groups by adapting communication channels (e.g., mobile audio and voice transcripts) for users with varying literacy levels, including those who did not receive formal education in reading French or Shikomori at school, as they only attended madrasas (Arabic schools). The data collection instruments, including the survey and interview questions (available on Appendix 1 and 2) were designed to assess ITC usage and the need for Shikomori-localised solutions. The questionnaire was translated and validated into French and English at first, and then into Shikomori, with the process reviewed by local linguistic experts to ensure cultural appropriateness and avoid bias. This two-phase approach was a deliberate methodological choice, designed to test the hypothesis that respondents would be more committed if the language was familiar, thereby ensuring consistency in the research. A pilot test of the questionnaire was conducted with peers, students during a focus group, and linguists to refine the survey tool before distributing the final version to a broader sample, strengthening trust in its reliability.

Data Analysis

The quantitative data obtained from the online survey was analysed using descriptive statistics (e.g., frequencies, percentages) to identify trends in ICT access, usage, and perceived needs. Cross-tabulations and correlation analyses via Xlstat examined relationships between demographic variables and digital inclinations related to Shikomori language use. Analysing the relationship between participants' dialect and their preferred ICT tool language revealed whether language preference varies significantly by dialect, helping to ensure that digital solutions are appropriately tailored. Exploring correlations

between digital literacy and attitudes towards Shikomori ICT tools helped to understand acceptance patterns. Thematic analysis was employed on the qualitative data derived from both the open-ended survey responses and the semi-structured interviews. This entailed the systematic coding of the transcribed data to identify recurring themes, patterns, and key insights related to the challenges and opportunities associated with Shikomori integration. To enhance the validity and reliability of the findings and to ensure a comprehensive understanding of the research problem, data triangulation from the literature review, survey, and interviews was performed.

Conceptual and Theoretical Framework

This study proposes a conceptual framework for the strategic integration of Shikomori into Information and Communication Technologies (ICT) within the Comoros Islands. This framework synthesises findings from an extensive literature review and a sectoral analysis, drawing on recommendations specifically tailored to the socio-linguistic and technological landscape of the Comorian context (e.g., Tahmasebi, Osborn, and Warschauer). By synthesising established best practices, it provides a compelling case for policymakers, developers, educators, and private sector stakeholders. The study is fundamentally grounded in two pivotal theoretical frameworks: User-Centered Design (UCD) and the Technology Acceptance Model (TAM). These models are adapted to explain the specific dynamics influencing technology adoption and usability within the Comorian socio-linguistic environment.

Prioritising the Comorian User

User-Centred Design (UCD) is an iterative design philosophy that posits the ultimate success of any technological solution is contingent upon its design being intrinsically aligned with the needs, wants, and inherent limitations of its end-users (Norman & Draper, 1986)

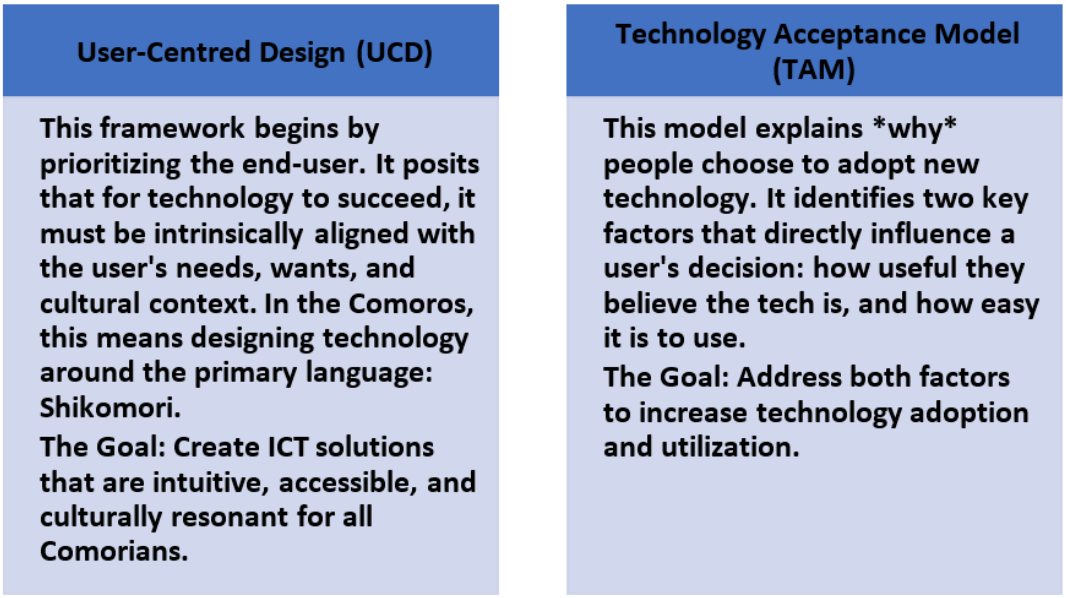

Figure 1: User-Centred Design and the Technology Acceptance Model Drive Shikomori Integration in ICT. Source: Model of The Framework for Digital Inclusion

In the Comorian context, where Shikomori serves as the primary language of daily communication for the vast majority of the population, a rigorous UCD approach necessitates the profound incorporation of the indigenous language into digital interfaces, content, and services. This principle extends beyond mere linguistic translation to advocate for comprehensive localisation, including culturally specific metaphors, nuances and communication practices. By prioritising the linguistic and cultural context of the Comorian user, UCD principles become instrumental in guiding the development of ICT solutions that are not only intuitive and accessible but also deeply resonant and culturally relevant, thereby fostering genuine engagement and utility. Therefore, the goal is to create ICT solutions based on Shikomori use that are intuitive, accessible, and culturally resonant for all Comorians.

The enhancement of usefulness and ease in Comoros context.

The Technology Acceptance Model (TAM) provides a comprehensive framework for analysing the multifaceted factors that collectively influence an individual's intention to adopt and consistently utilise a new technology (Davis, 1989). TAM identifies two principal beliefs that are critical determinants of technology acceptance: perceived usefulness, the degree to which an individual believes that using a particular system will enhance their job performance or improve their life; and perceived ease of use, the degree to which an individual believes that using a particular system will be free of effort (Davis et al., 1989: 985). The prevailing linguistic barrier, characterised by the dominance of official languages in contemporary ICT solutions, imposes a substantial cognitive burden and plays a pivotal role in the perception of challenges and apprehension associated with technology adoption in the Comoros (World Bank, 2021).

The integration of Shikomori into ICT is a strategic approach that directly addresses and enhances the fundamental tenets of the Technology Acceptance Model. This linguistic integration has been demonstrated to enhance the perceived usefulness of digital content by rendering it more comprehensible, relevant, and directly applicable to the daily lives of Comorian citizens. The provision of digital interfaces, applications, and content in Shikomori is intended to reduce the cognitive burden on the user. This linguistic alignment serves to render the technology less intimidating, more intuitive, and significantly easier to navigate and interact with, thereby lowering the barrier to entry for a substantial segment of the population (Duran et al., 2019).

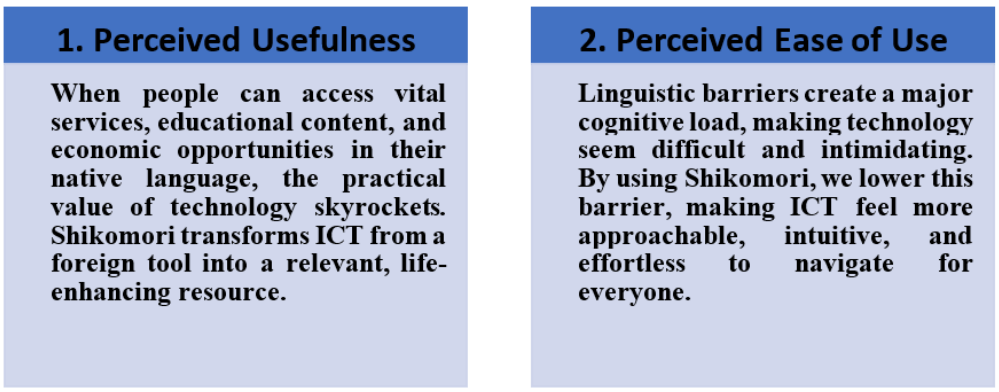

Figure 2: Integrating Shikomori directly strengthens both core tenets of the Technology Acceptance Model. Source: Model of The Pillars of Acceptance

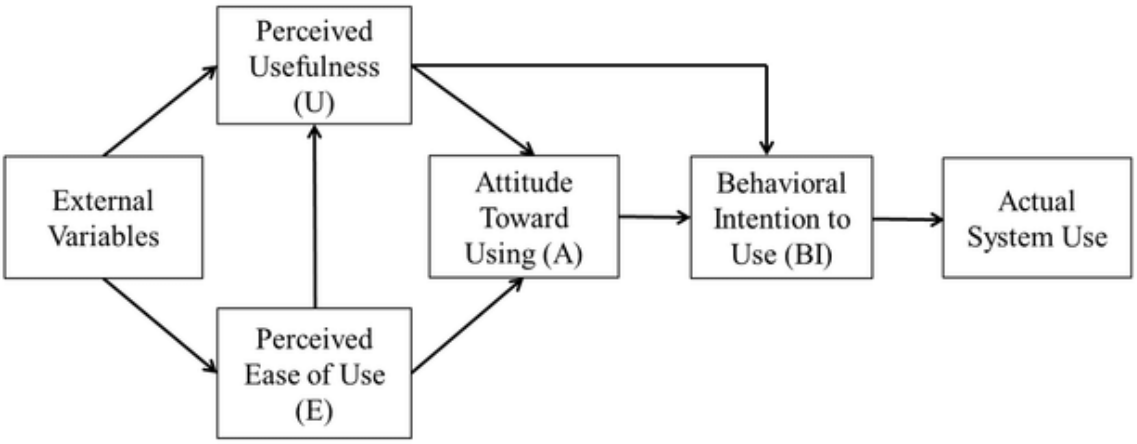

Figure 3: Technology acceptance model. Source : Technology acceptance model (Davis et al., 1989 : 985)

For Comorian people, the perceived usefulness of Information and Communication Technologies (ICTs) is hypothesised to increase substantially if the ICTs are perceived to deliver valuable information, essential services, and tangible opportunities in a language that is comprehensible to the population (i.e., Shikomori). For instance, the ability to seamlessly access critical government services, comprehend vital public health information, or engage with educational materials directly in Shikomori profoundly elevates the practical value and relevance of digital tools in their daily lives.

UCD and TAM limit critical for our research

The User-Centered Design (UCD) and Technology Acceptance Model (TAM) provide valuable frameworks for individual experience; however, they overlook the socio-cultural dynamics of perceived usefulness and ease of use in the Comorian context. Comorian society is communal and group-oriented, especially regarding traditions and changes in social behaviour and consumption. Most changes occur through community approval, meaning user experience alone is not sufficient within the community. Instead, if leaders promote the digital use of Shikomori in everyday solutions, for example, it is more likely to succeed, as role modelling is critical for community acceptance. Therefore, addressing collective dynamics – and simultaneously technical barriers such as the lack of a language policy, infrastructure limitations, or the variety of Shikomori dialects can lead to more effective digital inclusion. As this study also recommends, this requires a co-designed process and implementation that include local communities, the public and private sectors, and researchers. This strategic linguistic inclusion transforms technology from an unfamiliar tool into a practical resource that improves daily life, provides critical access to administrative and educational content, and ensures professional and economic improvements, as seen in the tourism sector, which supports sustainable development goals within Comorian communities.

Thematic Literature Review

The integration of indigenous languages into ICT is a subject addressed by various researchers who seek to explain the digital divide and promotion of sustainable development, particularly in linguistically diverse nations. A significant body of research underscores language as

a primary barrier to ICT adoption, asserting that true access encompasses the ability to interact with and create content in one's native tongue (Warschauer, 2003). Organisations like UNESCO (2003) consistently advocate for linguistic diversity in cyberspace, recognising that the dominance of a few languages online marginalises a vast portion of the global population and contributes to the erosion of cultural heritage.

Linguistic description and complexity of Shikomori

Research by Mohamed Ahmed-Chamanga (2011: 19-35; 2022: 79-98), including his manual on Comorian conjugation and other grammatical works provides an in-depth understanding of the linguistic features of Shikomori. Shikomori is an agglutinative language with complex syntactic structures that are closely tied to Comorian cultural identity. The main dialectical variations, such as Ngazidja and Ndzuwani, highlight this linguistic diversity. Djohar Abdou and the National Centre for Comorian Language Documentation (CNDRS) (2010) further expand on this by emphasising the language's rich lexical complexity and writing rules. This confirms that, despite external influences, Shikomori remains a fundamental pillar of local communication.

Shikomori Language in the Digital and Resource Scarcity

While specific academic research on the integration of Shikomori into ICT may be limited, the broader literature and recent scientific contributions provide an updated framework for the Comorian context. Key challenges include a lack of resources in local languages, including the three main dialects. Naira et al. (2024) have directly addressed this issue by focusing on dataset creation and empirical evaluations of cross-lingual learning for Comorian dialects. Their work, presented at the 18th Linguistic Annotation Workshop (LAW-XVIII) in "Datasets creation and empirical evaluations of cross-lingual learning on extremely low-resource languages: A focus on comorian dialects", underscores the foundational work required to enable digital linguistic processing for such languages. This demonstrates a promising approach to overcoming the scarcity of localised data. Their research supports the view that successful integration of local languages within ICT requires active community engagement in creating digital content, rather than merely consuming content in dominant languages such as French or English.

Cultural Valorisation and Sociopolitical Challenges

Despite being deeply rooted in Comorian culture, Shikomori faces challenges from globalisation and the dominance of languages such as French and Arabic. Comorian institutions, including the CNDRS, actively promote the language through education, literature, and cultural activities in order to preserve this linguistic heritage within a changing national and regional socio-political context. Additionally, in "Harnessing Transfer Learning from Swahili: Advancing Solutions for Comorian Dialects", at the 6th Deep Learning Indaba Conference (DLI 2024), Naira et al. (2024) explore the efficacy of transfer learning from resource-rich languages like Swahili to overcome data scarcity in Comorian dialects. These scientific efforts align with the broader understanding of local language integration in ICT. The Comorian community must actively create and engage with digital content in Shikomori, including its dialects to overcome the digital divide, rather than just consuming French or English-language content.

Integration of Shikomori in ICT and Digital Education

The literature emphasises the importance of adapting digital content for local languages, particularly to improve educational accessibility. Projects supported by international organisations such as UNICEF have experimented with integrating Shikomori into digital educational platforms, thereby enhancing access to education and helping to reduce social inequalities. Studies such as those by Obikudon (2022) demonstrate the importance of developing digital lexicons for African languages as a foundational step for ensuring their viability in the digital space. Furthermore, it has been shown that such efforts encourage greater engagement in the dissemination of local knowledge and cultural expressions.

Institutional and Political Framework

Official documents emphasise the importance of considering the political and administrative context of the Comoros to effectively implement ICT solutions in Shikomori. They emphasise the importance of involving public actors and linguistic communities through a participatory approach that is adapted to the country's sociolinguistic realities. This study also builds on Naira et al.'s (2004) work to analyse technology acceptance in specific cultural contexts. Based on Venkatesh and Bala's (2008) TAM3 model, their research provides a robust theoretical framework that highlights perceived usefulness and ease of use, emphasising adaptation to local cultural and socio-economic particularities. This approach enriches the current study by providing a contextualised understanding of the factors that encourage the sustainable adoption of digital technologies in Comorian communities.

Sustainable Development as a framework for Shikomori Integration in ICT

The integration of Shikomori into ICT solutions directly advance the Sustainable Development Goals (SDGs) by anchoring development in the Comorian context – prioritising cultural preservation and community benefit alongside economic growth. For example, providing educational content in Shikomori on digital platforms supports SDG 4 (Quality Education) by enabling Comorian citizens to learn in their native language, significantly boosting understanding and retention, mirroring the success observed in local Quranic schools (madrassas). Furthermore, granting Shikomori speakers access to digital public administration and tourism services promotes SDG 9 (Industry, Innovation, and Infrastructure) and SDG 10 (Reduced Inequalities) through inclusive service provision and enhanced school inclusion for children. Ultimately, fostering Shikomori in digital technologies strengthens social cohesion and local identity, supporting SDG 11 (Sustainable Cities and Communities) by cultivating united and supportive digital communities. These findings confirm that linguistic inclusion in technology is not a mere convenience, but a culturally rooted mechanism essential for driving sustainable development in the Comoros.

Research gap

Despite the general advancement of African language digitisation, a significant research gap remains because prior work, such as that by Naira et al. (2004) does not address the distinct sociolinguistic and political context of the Comoros. Specifically, there has been a lack of comprehensive, evidence-based studies demonstrating the necessity of

integrating Shikomori into ICT given the Shikomori dialects spoken across the islands and the marginalised socio-political perception of Ngazidja as the main island, therefore Shingazidja as the standard dialect to be referred into ICT solutions. This gap is critical as it shapes language standardisation and adoption in key sectors. Closing this research gap, which this study begins to address, is essential to enable future ICT solutions to be co-designed and properly tailored for effective communication, education, and digital inclusion within Comorian communities, ultimately paving the way for initiatives like a dedicated language and culture laboratory for Shikomori. This should start with its implementation as a language taught in local schools, alongside French and English, and ultimately as a language used to teach all subjects, as with any sovereign language in any sovereign country worldwide.

Analysis of Findings and Recommendations

The analysis of the Comorian digital landscape, informed by UCD and TAM, reveals that the dominance of French and Arabic in ICT creates a significant barrier to equitable digital participation for much of the population. Integrating Shikomori is a fundamental requirement, not just a linguistic preference, to ensure access to information and services. Despite being deeply embedded in the daily life and oral communication of Comorian society – even being the primary medium on national TV – Shikomori is largely excluded from international ICT solutions, hindering digital inclusion. This study uses survey and interview data focused on education, tourism, public administration, and telecommunications to evaluate the demand for and barriers to integrating this culturally significant language into digital platforms.

Findings from interviews

Interviews conducted with teachers, linguistics, IT developers, digital entrepreneurs, and professionals, in four key sectors, reveal significant findings regarding the integration of the Shikomori into ICT solutions. Analysis of the responses allows for conclusions to be structured around three main themes (high demand from the diaspora, standardisation challenges, lack of political will and investment in promoting Shikomori across various sectors).

Interest in ICT for Shikomori online learning as an identity driver for diaspora

There is a high diaspora demand on learning Shikomori due to quest of linguistic identity and cultural or community inclusion. This proves that ICTs are more than just learning tools for them and allow reconnection to culture and language that is not locally available or taught at schools. As stated by Nasser (Interview, September 26, 2025), ORECL founder, "ICT solutions, particularly ORELC's online courses are massively adopted by the diaspora (90% of learners) seeking to reconnect with their identity and overcome geographical distance". The success of platforms like ORELC (with its A1–B2 certifications) grants Shikomori *de facto* recognition and professional value, which is denied by the formal education system. In the case of Shikomori, digitisation precedes institutionalisation. It is paradoxical that Shikomori is an official language in the Comoros, yet it is not taught at local schools and in France – it is on high demand and taught through ORELC or INALCO. That reveals that the

legitimation of Shikomori is currently driven in a decentralised manner by the diaspora and the private sector, not by the state. According to Nusbati (Interview, September 18, 2025), linguist and teacher of Shikomori, "The prohibition or punishment for speaking Shikomori in class or at home contributes to its unconscious devaluation and perception as a 'non-legitimate language', reinforcing the appeal of French".

Digital Standardisation Challenges and Partnership

Although the spelling has been officially formalised, the acceptance of a single digital standard remains difficult due to dialectal bias and non-uniform writing practices (e.g., writing "oi" instead of "wa"). The major technical challenge has shifted from basic feasibility to scaling up digital resources and databases, requiring consensual linguistic harmonisation and innovation investment to integrate AI tools like voice translation. This advanced development is significantly hampered by the lack of financial means for acquiring technology and the massive data corpora needed (e.g., 10,000 *open source* words from ORELC), apps (Luha), and basic tools (keyboard) are accessible, creating an initial digital foundation for developers (Nasser, founder of ORELC). Furthermore, private initiatives face institutional instability among public stakeholders, making public-private partnerships unpredictable. The private sector's preference for private-to-private arrangements highlights a structural governance weakness hindering national linguistic innovation.

Integration and Social Impact

Integrating the mother tongue into technology is considered a fundamental tool for mental decolonisation, fostering cultural pride, and promoting effective learning. ICTs are uniquely positioned to redefine Shikomori's social role by making it professionally and culturally relevant, thereby helping to rebuild linguistic self-esteem proved by the high demand for our translation services by local companies such as Telma and Rotary (Nasser, founder of ORELC). Beyond utility, technology serves as a powerful vector for transmitting culture and literature (like poetry and slam). Successful digital integration is essential for human development, leading to improved cognitive efficiency and greater civic inclusion (e.g., decentralisation of administrative services). However, Mkandzile (Interview, July 10, 2025), Shikomori language and culture promoter, concludes that without a concerted effort to institutionalise these ICT tools within the public education and administration sectors, the current digital divide risks evolving into a new form of linguistic inequality.

In conclusion, although ICTs offer an opportunity for Shikomori integration, inclusion, and cultural recognition, the institutionalisation of digital tools and resources (e.g., through a national laboratory for Comorian language and culture) is requested to ensure long-term transmission.

Findings from Survey

From the collected data, the following approach was used to create pivot tables and dynamic charts for a study on the integration of Shikomori in ICT solutions across tourism, public administration, education, and telecommunications sectors. The results will be presented with interpretations based on 60 respondents.

Current Landscape and Language Use in ICT

This objective assesses how the community currently interacts with digital technology and the linguistic environment they operate in.

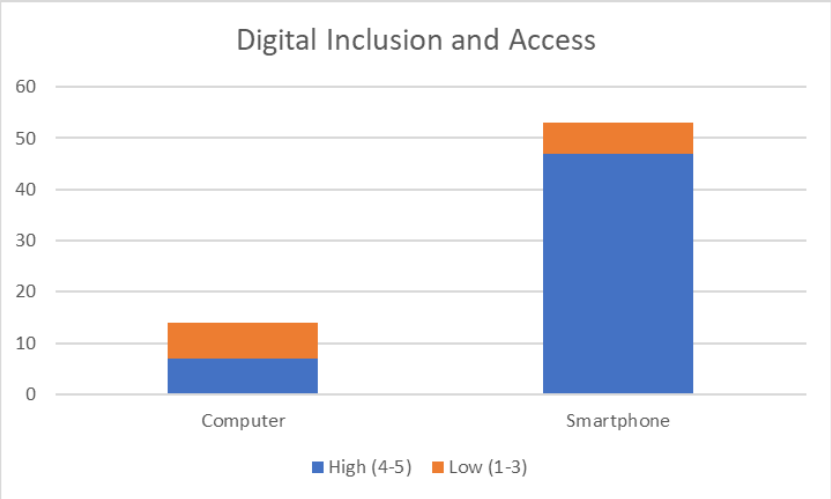

Figure 4: Digital inclusion and internet access device. Source: Author, September 2025

With 79% of usage citations belonging to the Smartphone, all ICT strategies for sectors like Education and Public Administration must be mobile-first, especially since the population most capable of engaging with Shikomori content shows even higher reliance on these devices. This dominance confirms that relying on traditional computer literacy or complex website interfaces will exclude the majority of the population.

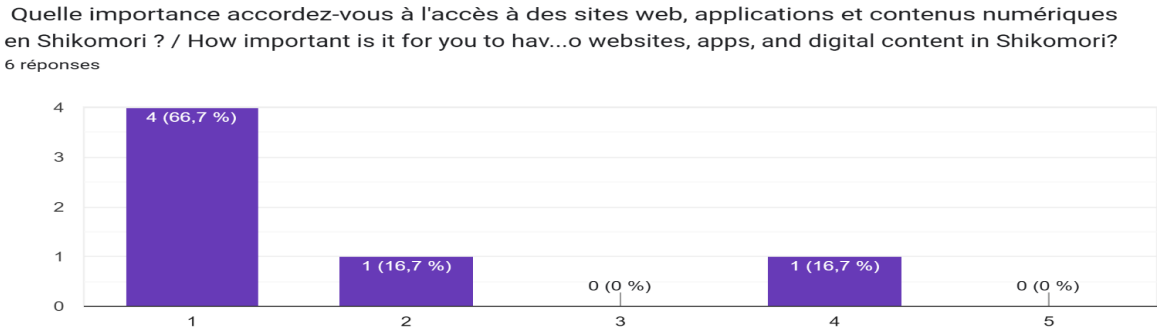

Figure 5: Importance of access to Shikomori for digital content. Source: Author, June 2025

A majority of respondents (66.7%) assigned the lowest rating (1), confirming a predominant and high degree of dissatisfaction with the current integration of Shikomori into ICT solutions. This distribution, lacking any neutral or highly positive feedback, indicates that the current level of integration is widely considered substandard and inadequate.

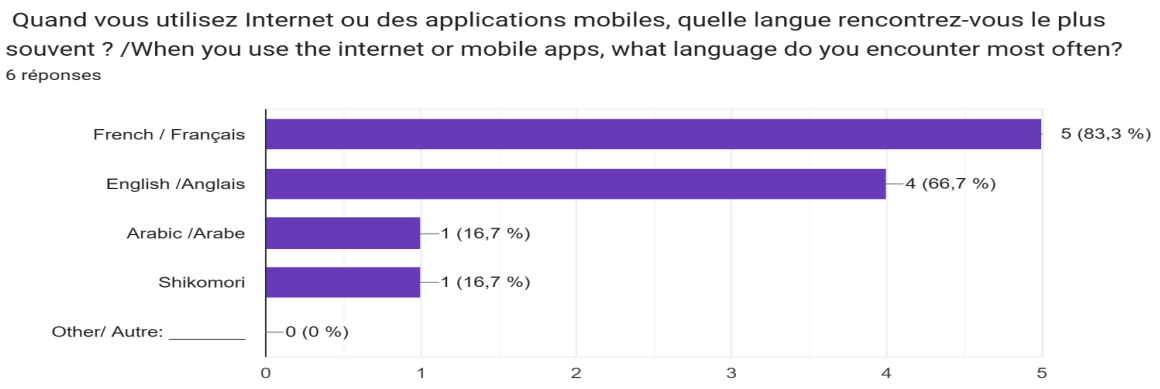

Figure 6: Language most encountered while using Internet. Source: Author, June 2025

The provided chart demonstrates the prevalence of French and English in ICT solutions, with 83.3% and 66.7% representation respectively. However, it also indicates a limited integration of Shikomori, which currently stands at only 16.7%. This finding suggests that, despite its recognition as a language, Shikomori is significantly underrepresented in current ICT offerings, which may potentially limit access and usage to a segment of the population.

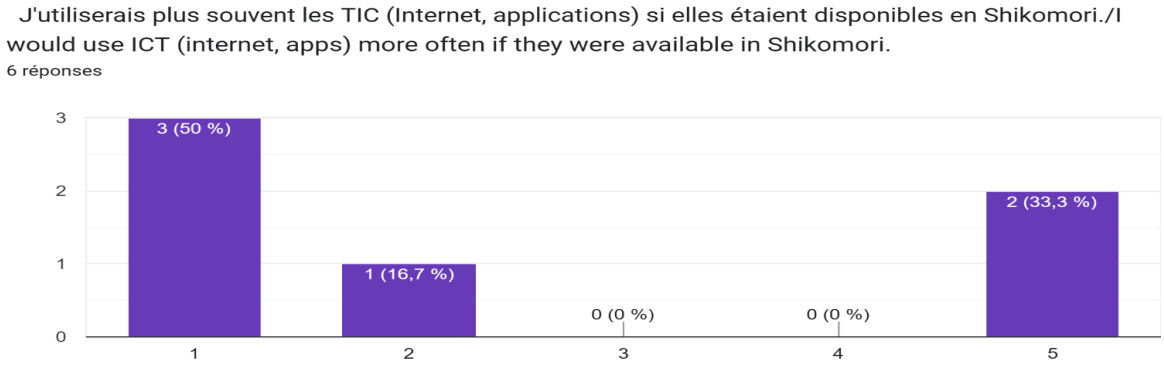

Figure 7: More use of ICT solutions if available in Shikomori. Source: Author, June 2025

The data shows a sharp divergence in satisfaction, with 50% of respondents assigning the lowest rating (1), while a significant 33.3% assigned the highest rating (5). This absence of moderate or neutral responses indicates that opinions are firmly split between those who see adequate integration and those who see little to none.

Despite the limited formal integration of Shikomori on many digital platforms, ICTs have permeated various aspects of daily life in the Comoros, often in informal ways. The term "Shikomori" is employed in a variety of informal settings, including residential areas, public spaces such as markets and transportation hubs, office environments, and even educational institutions. Its primary function is not in the context of formal education, but rather in facilitating conversation. For instance, the Shikomori language is utilised by customers and vendors in business transactions, by families in communication, and by educators in the clarification of lessons to ensure full comprehension, despite the official curricula being in French. It is evident that students are naturally drawn to Shikomori

as a forum for collaborative work and concept discussion. This common informal usage underscores a pressing need for formalised Shikomori-based ICT applications.

- **In transportation:** Mobile phones are used extensively for the organisation of transportation, the coordination of travel, and the dissemination of real-time road information. Despite the absence of formal ride-sharing applications, informal communication via telephone and text message (frequently in Shikomori) serves as a crucial means of facilitating transportation. However, there is a lack of formal public transport apps or digital navigation tools with Shikomori interfaces. While drivers may utilise French-based GPS systems, communication with passengers regarding routes or destinations is predominantly conducted in Shikomori. This underscores a functional disparity wherein informal linguistic interactions are lacking in support from formal digital tools.
- **In local markets**: Mobile money services are gaining traction for transactions, especially for larger purchases. Although negotiations frequently take place in Shikomori, the digital transaction itself is usually facilitated by a French or Arabic interface. Furthermore, vendors utilise social media platforms for the purpose of advertising. Informal communication among market vendors, including voice notes or text messages in Shikomori, is common for sharing information on product availability or market trends. This indicates a strong existing user base and a clear need for localised financial and commercial ICT tools.
- **At school:** Despite the predominance of French and Arabic in educational content and administrative communication, mobile devices are frequently utilised by students and teachers for informal learning, research, and communication. The utilisation of social media platforms, particularly WhatsApp groups, has become a dominant method for facilitating group discussions and the distribution of supplementary materials. It has been observed that communication frequently occurs in either Shikomori or a blend of Shikomori and French. However, there is a significant absence of dedicated educational software or online learning platforms with Shikomori content. It has been observed that students may engage in information seeking in French, yet their discourse with peers on the subject tends to occur in Shikomori. This informal adaptation of existing technologies for Shikomori communication demonstrates a strong demand for mother-tongue educational resources.

This exclusion of Shikomori **marginalises most of the population**, holding back education, economic growth, and civic participation. For true digital transformation to happen, it is essential to embrace the language of the people.

An overwhelming 81% of respondents rate access to Shikomori content as high or highest importance (scores 4 or 5), confirming a massive market demand that justifies immediate investment in localised digital content creation for sectors like Education and Tourism. This demand serves as a critical proxy for content needs, highlighting the importance rating across both the under 26 group (Education/Youth) and the 26+ group (Professionals/ General Public).

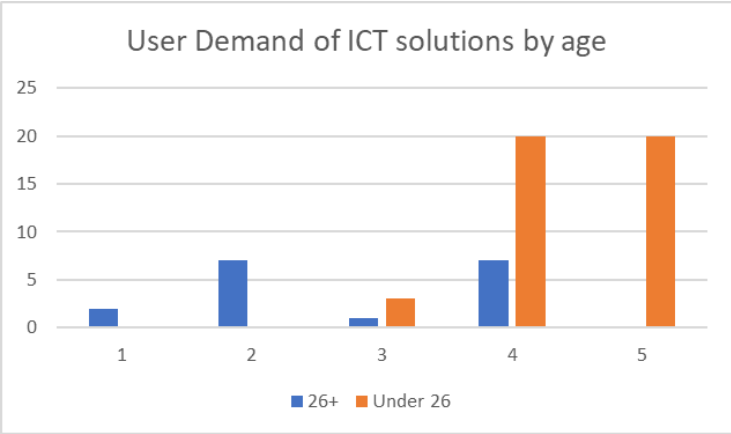

Figure 8: User Demand of ICT solutions by age. Source: Author, September 2025

The findings effectively mapped the current digital landscape by confirming the necessity of a mobile-first strategy and highlighting device dependency across sectors, thus directly answering the research question regarding current ICT access methods and platforms within Comorian communities. Users are forced to navigate system in languages that are not their mother tongue, particularly for formal applications. This informal adaptation highlights the pressing need for a more deliberate and comprehensive integration of Shikomori into ICT solutions to achieve true democratisation of access and enhance utility. The absence of culturally and linguistically appropriate digital tools in key sectors severely limits the potential for ICT to drive democratise digital access, empower communities, foster digital literacy, and significantly enhance the overall impact of ICT on national development.

Sector-Specific Gaps and Informal Adaptations

This objective identifies where the need for Shikomori is greatest and how the lack of integration affects access in key sectors like Education, Tourism, Public Administration, and Telecommunications.

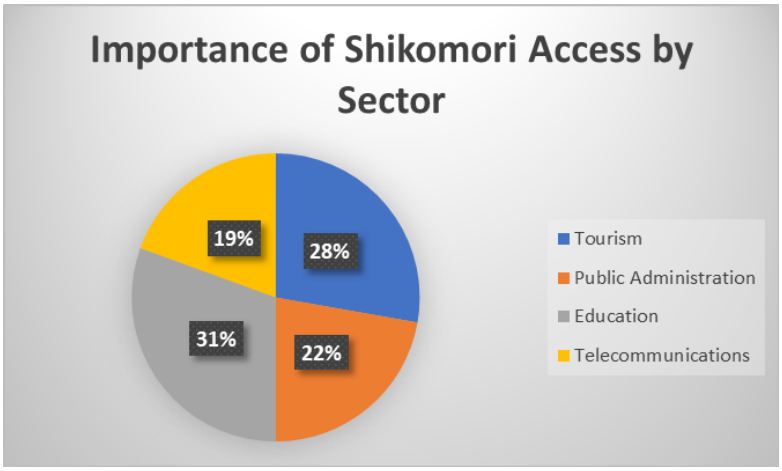

Figure 9: Importance of Shikomori Access by Sector. Source: Author, September 2025

Tourism and education sectors show highest importance for Shikomori access, suggesting these are top priority areas for digital language integration. While it is evident that

international visitors require information in global languages such as French and English, it is equally important to recognise the necessity of empowering local tourism stakeholders through Shikomori-based digital resources. The creation of mobile applications featuring localised maps, historical narratives, and service directories in Shikomori has the potential to substantially enhance the operational capacity of local guides, small-scale entrepreneurs, and community-based tourism initiatives. This approach not only enhances the visitor experience by facilitating more authentic and culturally immersive interactions but also ensures a broader and more equitable distribution of economic benefits within local communities.

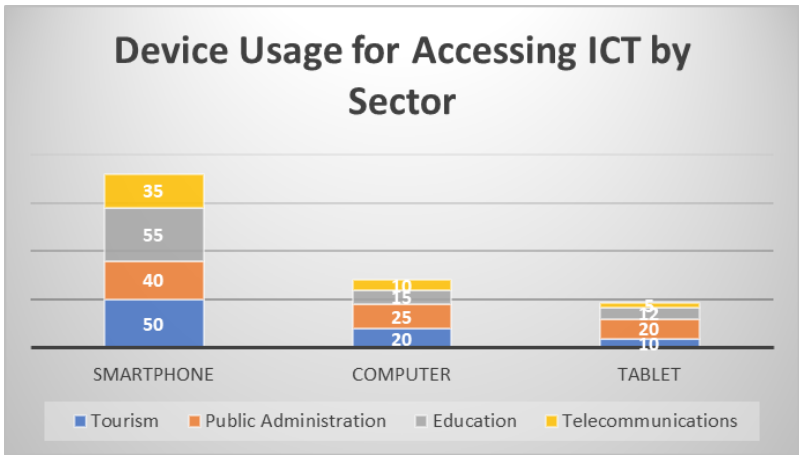

Figure 10: Device Usage for Accessing ICT by Sector. Source: Author, September 2025

Smartphones dominate as primary device for accessing ICT, especially in tourism and education, underlining the importance of mobile-friendly Shikomori content.

Barriers and Challenges to Integration

This objective isolates the key hurdles – technical, institutional, and linguistic – that prevent effective Shikomori ICT integration. The vast majority of ICT tools, software applications, and online content is developed in globally dominant languages such as French, Arabic, and English. This creates a considerable accessibility obstacle for the Comorian population, a significant proportion of whom demonstrate higher proficiency in Shikomori.

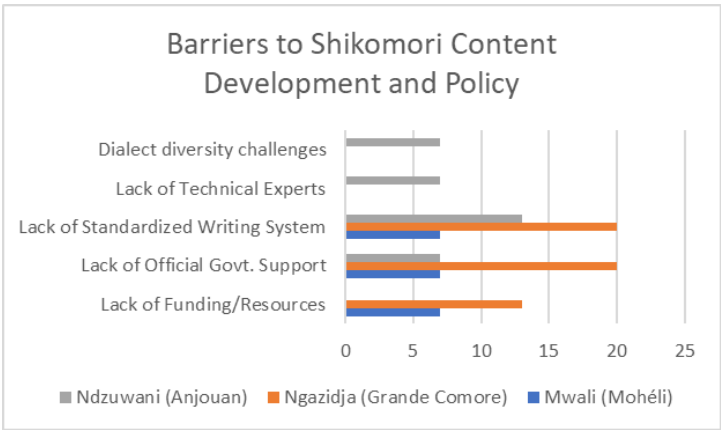

Figure 11: Barriers to Shikomori integration by Island. Source: Author, September 2025

The primary barriers are non-technical and institutional: Standardisation (40 citations) is the most critical technical barrier preventing advanced ICT integration, and the lack of official government support (34 citations) urgently demands a cohesive national policy. Furthermore, lack of funding is a significant financial constraint for local development, particularly in Ngazidja.

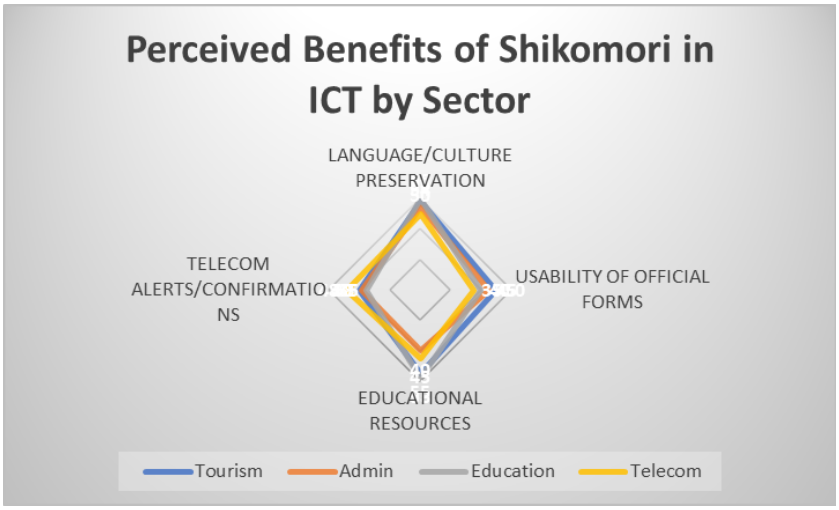

Figure 15: Perceived Benefits of Shikomori in ICT by Sector

The chart shows the perceived benefits of Shikomori in ICT by sector as consistent, high perceived value across all sectors (Tourism, Admin, Education, Telecom) for educational resources and language/culture preservation, suggesting these are the most widely agreed-upon benefits of integrating Shikomori into ICT solutions. Conversely, the sectors show a lower and more variable perceived benefit for usability of official forms and telecom alerts/confirmations.

By isolating both user-driven adoption patterns (driven by youth and literacy) and macro-level structural barriers (standardisation and political will), the analysis provided a comprehensive diagnosis required to answer the research question regarding challenges and adoption drivers in Shikomori ICT integration.

Community Needs and Aspirations

This objective captures the emotional, cultural, and practical goals of the community regarding Shikomori integration into technology. The community's aspiration for Shikomori integration is rooted in a unanimous mandate for cultural preservation, with 60 out of 60 respondents (100%) of respondents agreeing it is crucial for linguistic survival, not just convenience. This high demand is paired with the overwhelming dominance of the smartphone across all sectors, making a mobile-first strategy mandatory. However, achieving effective integration and bridging the digital divide requires immediately addressing key obstacles: the lack of specialised technical expertise, standardisation, and funding, through targeted, collaborative, and interdisciplinary interventions. The results reinforce that despite the sample size, the trends are consistent for broader application in Comoros. Addressing these issues requires targeted interventions focused on digital literacy, the creation of content in local dialects, and collaborative efforts among academic, institutional, and community stakeholders to bridge the digital and linguistic divide.

This affirms the need for an interdisciplinary approach and a framework for policy recommendations and implementation plans involving professionals, government, local communities, and researchers.

Recommendations for Shikomori Integration in Comorian ICT

To effectively integrate the Shikomori language into digital tools and address the challenges discussed, a strategic plan is needed. This plan must focus on creating culturally specific solutions, fostering collaboration among diverse stakeholders, and ensuring sustained community involvement to promote lasting digital access. These recommendations are structured to address the identified challenges while leveraging the strong community support and adapting to the dominant mobile landscape.

Sectoral and Technical Focus

Given that **79%** of internet access is via **Smartphone**, all solutions must be lightweight, fast, and user-friendly mobile applications that meet the specific needs of each sector.

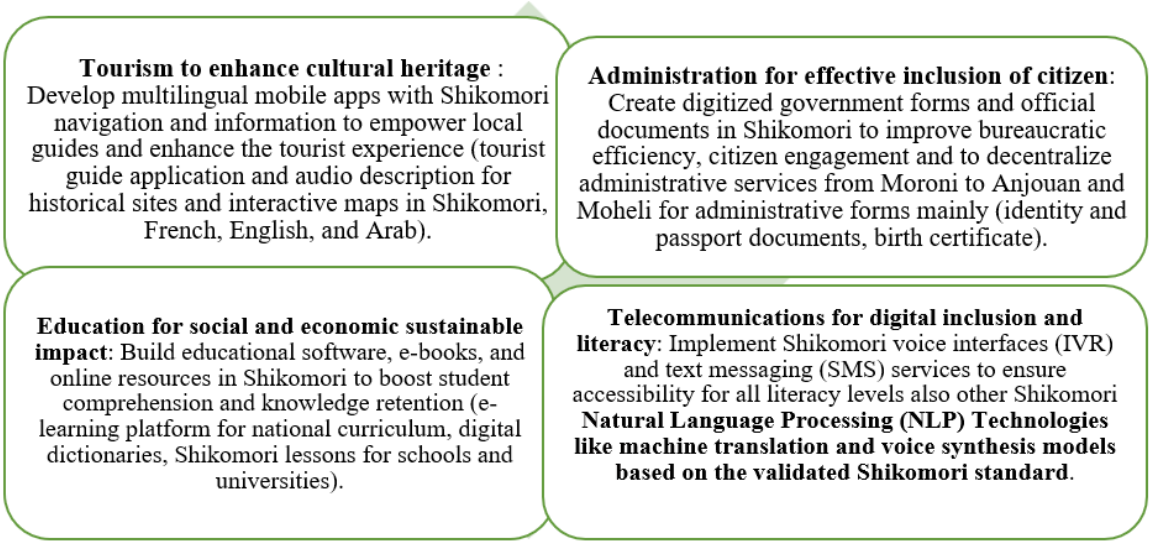

Figure 16: A Roadmap for Inclusive Technology. Source: Author, June 2025

Institutional and Linguistic Leadership

A strategic approach requires three prioritised interventions to overcome structural barriers. The most critical need is institutional support (34 responses), mandating the creation of a National Digital Language Committee and a legal framework to officially recognise and budget for Shikomori in digital public administration. Parallel to this, the lack of standardisation (40 responses) must be solved by having the Committee finalise and deploy a single standardised writing system, a technical prerequisite for developing advanced AI and e-learning tools in both education and telecommunications. Finally, the lack of funding (20 responses) should be addressed by launching an innovation fund for digital inclusion, leveraging public-private partnerships (PPP) to subsidise the development of essential Shikomori utility applications for commerce and public services.

Figure 17: Collaboration between government, Academia, and Local Communities. Source: Author, May 2025

The successful integration of the Shikomori language into the digital world requires a collaborative effort from all stakeholders. The government must take the lead by establishing policies for Shikomori's use in public online services, funding research into language technology, ensuring widespread internet access, and incentivising companies to develop Shikomori-based solutions. Academia plays a crucial role by conducting research on language technologies, creating digital linguistic tools, developing curricula to teach digital skills in Shikomori, and training future experts. Furthermore, local communities (including elders, youth, and cultural experts) must be involved at every stage to ensure that the technology is culturally appropriate and user-friendly. Their participation is vital for providing valuable knowledge, feedback, and advocacy for Shikomori's inclusion in the digital sphere.

Figure 18: Active participation of local communities. Source: Author, May 2025

The success of Shikomori integration into ICT relies on a two-pronged approach: first, resolving the institutional crisis through standardisation and political backing from public administration, and second, building highly targeted mobile solutions for education, tourism, and public services. The enthusiasm and need expressed by the community guarantee that these investments will have an immediate and lasting impact. Furthermore, to improve technology use within Shikomori-speaking communities, it is essential to provide digital literacy training in Shikomori, tailored for all demographics, from young people to elders. It is also crucial to encourage individuals to create their own online content, such as videos and blogs, in their native language, thereby shifting from passive consumption of foreign content to active contribution. By teaching communities how to critically evaluate information and use social media responsibly, it will empower them to fully participate in the digital world in their own language.

Conclusion

The integration of Shikomori into Information and Communication Technologies (ICTs) is not merely a technical option but a fundamental imperative for equitable national development and the empowerment of Comorian communities. The findings of this study reveal a critical paradox: while strong, self-driven demand exists from the diaspora, private sector, and younger users - evidenced by the overwhelming reliance on smartphones and the high perceived value of Shikomori for cultural preservation - this momentum is severely constrained by structural and political barriers.

This study systematically highlights that progress is primarily hampered by techno-linguistic impediments, chiefly the lack of complete spelling standardisation and orthographic inconsistencies across dialects, alongside the critical absence of a national language policy and sufficient political will or funding. Furthermore, the minimal use of computers underlines a digital divide that risks excluding users from engaging in higher-literacy administrative and educational tasks.

To transition from mere linguistic inclusion to genuine digital empowerment, a paradigm shift is required. The path forward demands moving beyond models focused on individual experience (UCD/TAM) toward a strategy that embraces collective socio-cultural dynamics and inclusive co-design practices. Achieving sustainable integration requires a mandatory, multi-stakeholder approach to institutionalise digital resources and implement targeted interventions that boost digital literacy and foster collaboration across academic, governmental, and community sectors.

By proactively addressing these linguistic and structural disparities, the Comoros can leverage its distinctive heritage to strengthen social cohesion, enhance administrative efficiency, and fully unlock its citizens' potential, thereby firmly rooting its digital transformation in culturally relevant and sustainable development goals (SDGs). The success of this transformative endeavour rests upon an unwavering national commitment to endorse Shikomori as a cornerstone of the country's digital future.

References

Ahmed-Chamanga, M. (2010). *Introduction à la grammaire structurale du comorien: Le shiNgazidja*. Moroni, KomÉdit – Ya Mkobe publications.

Ahmed-Chamanga, M. (2011). « La langue comorienne ». In F. Laroussi & F. Liénard (éds.), *Plurilinguisme, politique linguistique et éducation*. Presses universitaires de Rouen et du Havre, 19-35. https://doi.org/10.4000/books.purh.5186

Ahmed Chamanga, M. (2022). « Chapitre 4 Shikomori, la langue bantoue des Comores : statut et perspectives ». In *Manuel de politique linguistique et d'éducation dans les pays de la Communauté de développement de l'Afrique australe* (SADC), 79–98. Leyde, Pays-Bas: Brill. https://doi.org/10.1163/9789004516724_006

Ajibade, P. (2018). 'Technology Acceptance Model Limitations and Criticisms: Exploring the Practical Applications and Use in Technology-related Studies, Mixed-method, and Qualitative Researches', *Library Philosophy and Practice (e-journal), (Paper 1941)*. https://digitalcommons.unl.edu/libphilprac/1941/

Bhattacharya, S., Panyam, S., Deshmukh, G. , Gatala, S., Vemoori, V. & Seth, D. (2024). 'Integrating User Experience and Acceptance in Authentication: A Synthesis of Technology Acceptance Model and User-Centered Design Principles', *International Journal of Computer Trends and Technology*, 72(4), 15-23. https://doi.org/10.14445/22312803/IJCTT-V72I4P102

Borthwick, M., Tomitsch, M., & Gaughwin, M. (2022). 'From human-centred to life-centred design: Considering environmental and ethical concerns in the design of interactive products', *Journal of Responsible Technology*, Vol (10), Article 100032. https://doi.org/10.1016/j.jrt.2022.100032

Chauvet, A. (2015). Statuts des langues et éducation de base aux Comores. *Revue internationale d'éducation de Sèvres. 70*, 77–84. Available at : http://journals.openedition.org/ries/4499. https://doi.org/10.4000/ries.4499. Accessed: November 14, 2025.

Chesire, E., & Kipkebut, A. (2024). 'Current state, challenges and opportunities for natural language processing research and development in Africa: A systematic review', *Africa NLP workshop at ICLR 2024, Kabarak University*, 1-8.

Creswell, J. W., & Creswell, J. D. (2022). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (6th ed.), Thousand Oaks, CA: SAGE Publications.

Daniel, R. S. (2024). Le shikomor pour enseignement / apprentissage du français langue étrangère : issue interculturelle de l'insularité de Ndzuani, *Revue Internationale du Chercheur*. 5(2), 734-757. https://www.revuechercheur.com/index.php/home/article/view/984/818

Davis, F. & Davis, F. (1989). 'Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology', *MIS Quarterly*, 13, 319-339. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). 'User acceptance of computer technology: A comparison of two theoretical models', *Management Science*, 35(8), 982-1003. https://doi.org/10.1287/mnsc.35.8.982

Duran, N. D., Paxton, A., & Fusaroli, R. (2019). 'ALIGN: Analyzing linguistic interactions with generalizable techNiques-A Python library', *Psychological Methods*, 24(4), 419-438. https://doi.org/10.1037/met0000206

Galla, C. (2018). 'Digital Realities of Indigenous Language Revitalization: A Look at Hawaiian Language Technology in the Modern World', *Language and Literacy*, 20, 100-120. https://doi.org/10.20360/langandlit29412

Gay, J. C. (2001). 'La diffusion du tourisme dans l'archipel comorien', *Mappemonde, 64(4), 15–18.* https://doi.org/10.3406/mappe.2001.1699

Goosen, M., Prinsloo, D., & Taljard, E. (2023). 'Creating electronic resources for African languages through digitisation: A technical report', *Journal of the Digital Humanities Association of Southern Africa,* 4(1), 1-7. https://doi.org/10.55492/dhasa.v4i01.4441

Halidi, A. (2018). 'Comores: Colonisation, insularité et imaginaire national', *Comoros: Colonization, insularity and national imagination. Les Cahiers d'Outre-Mer,* 71(278), 403-432. https://doi.org/10.4000/com.9692

Kukutai, T., & Taylor, J. (2016). *Indigenous data sovereignty: Toward an agenda*. Australian National University Press. Canberra. http://press.anu.edu.au/publications/series/centre-aboriginal-economic-policy-research-caepr/indigenous-data-sovereignty

Mugane, J. M. (1997). 'Learning African Languages with Evolving Digital Technologies', *Africa Today,* 44(4), 423-441. http://www.jstor.org/stable/4187190

Naira, A. M., Bahafid., A., & Erraji, Z. (2024). Datasets creation and empirical evaluations of cross-lingual learning on extremely low-resource languages: A focus on Comorian dialects. In S. Henning & M. Stede (Eds.), *Proceedings of The 18th Linguistic Annotation Workshop (LAW-XVIII)*, 140-149. Association for Computational Linguistics.

Naira, A. M., Bahafid., A., Benelallam,I., & Erraji, Z. (2024). 'Harnessing Transfer Learning from Swahili: Advancing Solutions for Comorian Dialects. *6th Deep Learning Indaba Conference (DLI 2024)*. Available at: *https://arxiv.org/abs/2412.12143*. Accessed: June 21, 2025.

Obikudo, E. F. (2022). 'Digitizing African languages: Building a digital lexicon', *Journal for Translation Studies in Nigeria (JTSN), 1(1), 38-52*.

ONTC. (2024). *Catalogue of hotel investment sites in the Comoros*. Moroni: ONTC Publications.

ORELC, *Online dictionary of Shikomori*. Availabe at: https://orelc.ac/academy/ShikomoriWords/?i=kmWords. Accessed: September 24, 2025.

Osborn, D. (2010). *African languages in a digital age: Challenges and opportunities for indigenous language computing*. IDRC-funded African Network for Localization, Cape Town: HSRC Press. http://hdl.handle.net/20.500.11910/4355

Panicker, P. (2020). *Embedding Culture and Grit in the Technology Acceptance Model (TAM) for Higher Education*. Available at: https://arxiv.org/abs/2005.11973. Accessed: June 21, 2025.

Pea, R., (1987). 'User centered system design: New perspectives on human-computer interaction', *Journal of Educational Computing Research, 3*(1), 129-134.

Pinch, T. J., & Bijker, W. E. (1984). 'The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other', *Social Studies of Science*, 14(3), 399-441. https://doi.org/10.1177/030631284014003004

Salawu, A. (2019). *African language digital media and communication*. Routledge Contemporary Africa. https://doi.org/10.4324/9781351120425

Smith, L. T. (2021). *Decolonizing Methodologies: Research and Indigenous Peoples* (3rd ed.). London: Zed Books. https://doi.org/10.5040/9781350225282

Soihili, S. (2011). La langue française dans l'archipel des Comores : statut, usages et pratiques de langue. In F. Laroussi & F. Liénard (eds.), *Plurilinguisme, politique linguistique et éducation*. Mont-Saint-Aignan: Presses universitaires de Rouen et du Havre. 175-185. https://doi.org/10.4000/books.purh.5210

Tahmasebi, F. (2023). 'The Digital Divide: A Qualitative Study of Technology Access in Rural Communities', *AI and Tech in Behavioral and Social Sciences*. 1 (1). 33-39. https://doi.org/10.61838/kman.aitech.1.2.6

Thompson, R. & Berko Gleason, Jean. (2001). *Language death. David Crystal*. Cambridge: Cambridge University Press, 2000. https://doi.org/10.1017/S0142716401212077

UNESCO. (2003). *Language vitality and endangerment*. Unesco Ad Hoc Expert Group on Endangerment Languages. Available at: *https://ich.unesco.org/doc/src/00120-EN.pdf*. Accessed: September 24, 2025.

UNESCO. (2003). *Recommendation concerning the promotion and use of multilingualism and universal access to cyberspace*. Available at: *https://unesdoc.unesco.org/ark:/48223/pf0000153336* Accessed: September 24, 2025.

Van Velsen L, Ludden G, Grünloh C. (2022). 'The Limitations of User- and Human-Centered Design in an eHealth Context and How to Move Beyond Them'. *Journal of Medical Internet Research*. 24(10):e37341. *https://doi.org/10.2196/37341*

Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/6699.001.0001

Wilson, S. (2008). *Research Is Ceremony: Indigenous Research Methods*. Black Point, Nova Scotia: Fernwood Publishing.

World Bank (2021), *Comoros integrated development and competitiveness project: Promoting Technology Adoption to Raise Productivity, Supporting Entrepreneurship and SME Growth*. Available at: *https://documents1.worldbank.org/curated/en/556041518192665327/Comoros-Integrated-Development-and-Competitiveness-Project.docx*. Accessed: September 27, 2025.

Appendix 1: Interview reflection themes on Shikomori integration into ICT solutions

1. Relationship with the Comorian Language and the Socio-linguistics of Shikomori in ICTs -What is your professional background related to Shikomori, and what motivates your choice to work primarily in this language (activism, writing, entrepreneurship)? -How do you perceive the influence of technologies (applications, software) on the learning, preservation, and dissemination of the Comorian language and culture?
2. Applications in Key Sectors - How can Shikomori be used to improve digital services in tourism, culture, education, and public administration? - What specific needs does your sector have regarding the use of Shikomori and its integration into ICTs?
3. Accessibility and Inclusion - What concrete ICT solutions (applications, dictionaries, *open source* resources) are currently available or under development? - What challenges do you face in developing inclusive and accessible solutions for better integrating Shikomori into ICTs?
4. Socio-economic and Political Impact - How could the use of Shikomori in ICTs strengthen linguistic transmission and community identity? - What socio-economic or political factors do you believe influence the adoption of ICTs by communities (political will, funding, stability of public-private partnerships)?
5. Implementation Strategies and Partnerships - What strategies would you recommend to foster the successful integration of Shikomori? - Which partnerships between linguists, public stakeholders, tourism professionals, and digital specialists are necessary to achieve this?

Appendix 2: Survey form on The Role of the Shikomori Language in Information and Communication Technologies (ICT) in the Comoros

This survey is part of a research study on the importance of integrating the Shikomori language into digital technologies (websites, mobile applications, etc.) in the Comoros. Your participation is voluntary and your responses will be kept anonymous and confidential. The goal is to better understand the needs and perspectives of Comorian communities in order to promote better access to information and digital services. Thank you for your time and valuable contribution.

General Information (Please check the appropriate box or fill in the blank)

1. Age: _____
2. Gender: ☐ Male ☐ Female ☐ Other
3. Island of Residence: ☐ Ngazidja (Grande Comore) ☐ Ndzuwani (Anjouan) ☐ Mwali (Mohéli)
4. What is your primary language spoken at home? ☐ Shikomori ☐ French ☐ Arabic ☐ Other: _____
5. Please rate your proficiency in the following languages (1=Beginner, 5=Fluent):
Shikomori: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
French: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5
Arabic: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

ICT Usage and Language

6. Which devices do you use to access the internet? (Select all that apply) ☐ Smartphone ☐ Computer (desktop/laptop) ☐ Tablet ☐ I do not use the internet
7. How often do you use the internet? ☐ Daily ☐ A few times a week ☐ Once a week ☐ Less than once a week
8. When you use the internet or mobile apps, what language do you encounter most often? ☐ French ☐ English ☐ Arabic ☐ Shikomori ☐ Other: _____
9. I sometimes find it difficult to use websites or applications because they are not in a language I fully understand. ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
10. How important is it for you to have access to websites, apps, and digital content in Shikomori? ☐ Very Important ☐ Important ☐ Neutral ☐ Not Very Important ☐ Not Important at All
11. I would use ICT (internet, apps) more often if they were available in Shikomori. ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree

Shikomori Integration in Key Sectors

(For the following questions, please indicate how useful you would find each service)

12. Tourism Sector
A mobile app for tourists with information and audio guides in Shikomori (in addition to French, etc.).
Response: ☐ Very Useful ☐ Useful ☐ Neutral ☐ Not Very Useful ☐ Not Useful at All
13. Administrative Services
Being able to fill out official online forms (e.g., ID card, passport application) using a Shikomori interface.
Response: ☐ Very Useful ☐ Useful ☐ Neutral ☐ Not Very Useful ☐ Not Useful at All
14. Education Sector
Digital educational resources (e-books, interactive apps) for children in Shikomori.
Response: ☐ Very Useful ☐ Useful ☐ Neutral ☐ Not Very Useful ☐ Not Useful at All
15. National Telecommunications
Receiving SMS alerts and transaction confirmations from your mobile operator or mobile money service in Shikomori.
Response: ☐ Very Useful ☐ Useful ☐ Neutral ☐ Not Very Useful ☐ Not Useful at All
16. News and Media
Accessing national news (digital newspapers, Facebook Live broadcasts) in Shikomori.
Response: ☐ Very Useful ☐ Useful ☐ Neutral ☐ Not Very Useful ☐ Not Useful at All

Benefits and Challenges

17. Do you believe that using Shikomori in technology would help preserve the Comorian language and culture? ☐ Yes, a lot ☐ Yes, a little ☐ I'm not sure ☐ No, not really ☐ No, not at all
18. Shikomori has four dialects (Shingazidja, Shindzuwani, Shimaore and Shimwali). Do you think these differences would be a major challenge for creating standardised digital content? ☐ Yes, a major challenge ☐ Yes, a minor challenge ☐ I'm not sure ☐ Not a significant challenge
19. In your opinion, what is the biggest obstacle to developing more digital content in Shikomori? ☐ Lack of funding/resources ☐ Lack of technical experts ☐ Lack of official government support ☐ Lack of a standardized writing system ☐ Other: _____

Comments

20. Do you have any other comments or suggestions regarding the use of the Shikomori language in technology in the Comoros?

Appendix 3: Summary of the findings on Shikomori integration in ICT solutions

Source: Author, September 2025

Opportunities, Benefits, and Challenges of Using Artificial Intelligence in the Teaching and Learning of IsiXhosa Poetry

Lukhanyo Elvis Makhenyane
Department of African Languages
University of Johannesburg
Johannesburg, South Africa
lmakhenyane@uj.ac.za

Nozuko Zukiswa Gxekwa
Department of African Languages
University of Fort Hare
Alice, South Africa
ngxekwa@ufh.ac.za

Abstract

IsiXhosa poetry holds significant potential for cultivating critical thinking, creativity, and problem-solving skills. However, its pedagogical implementation is often hindered by reliance on rote learning and lecture-based instruction, which restricts the higher-order cognitive engagement required for interpreting poetry. Cultural-Historical Activity Theory (CHAT) provides a useful framework for analysing this pedagogical tension by viewing teaching and learning as mediated activity systems involving subjects (students and educators), tools (instructional technologies), rules, community, and division of labour. Guided by CHAT, this study investigates how artificial intelligence (AI) can function as a mediating tool within the isiXhosa poetry learning activity system, identifying its opportunities, benefits, and contradictions. A systematic literature review synthesised global scholarship on AI in education, higher education, and language education to identify AI tools that could be integrated into isiXhosa poetry pedagogy. Studies were analysed through the lens of CHAT to evaluate the transformative potential and systemic tensions arising from AI adoption. The study provides a framework for selecting and integrating AI tools in isiXhosa poetry instruction, and it offers evidence-based recommendations for educators and policymakers to support the responsible adoption of AI in African-language literature education.

Keywords: IsiXhosa poetry, Artificial Intelligence, Cultural-Historical Activity Theory (CHAT), Mediated learning, Systematic Literature Review

Introduction

Scholars agree on the significance of the teaching and learning of poetry, both in basic and higher education. According to Khan (2020: 803), the teaching and learning of poetry "has the tendency to engage and motivate students more profoundly and at the same time open avenues to improve multiple language constructs." This view is corroborated by Moldagali et al. (2024: 169) who note that teaching and learning poetry is key "for teaching linguistic nuances and enriching students' emotions, thoughts, and vocabulary." Jaxa (2024: 119) adds that teaching and learning of poetry "fosters critical thinking, and empowers learners to question, evaluate, and support information effectively." The skills alluded to by these scholars are crucial in the development of human beings as social beings and responsible citizens of the global village, capable of contributing meaningfully to sustainable global development and grounded in ethical responsibility, and social justice. Hence, Jaxa (2024: 119) further notes, "Through poetry analysis, learners develop a nuanced understanding of their social and personal identities and thereby reinforce the integral role of poetry in education." In addition, Makhenyane (2024) alludes to the fact that isiXhosa poetry is abstract in nature, tapping into middle and higher-order cognitive skills, such as analytical thinking, problem-solving, critical thinking, and creativity.

Despite its significant role in improving analytical skills, critical thinking, motivation, language use, creativity, and problem-solving, evaluative skills are often found to be difficult to teach and learn by instructors and students alike. Jaxa (2024) discovered that learners perform poorly in isiXhosa poetry, citing the 2019 to 2022 Basic Education Diagnostic Reports. Jaxa (2024) further argues that this poor performance has ripple effects on the entire academic trajectory of learners. Gxekwa and Satyo (2017) support this view as they discovered that poetry can be used to improve literacy, numeracy, and life skills in learners. Their study focused on how in-service educators can leverage poetry to address difficulties in learning literacy, numeracy, and language skills.

Furthermore, Mbambo and Hlabisa (2024: 262) discovered that some instructors "use approaches that treat poetry as something to be memorised and regurgitated rather than comprehended and appreciated, thus exacerbating learners' anxieties in relation to poetry." Such approaches offer little and yield bitter fruits on the experience of teaching and learning poetry. Participants in their study commented on the difficulty of teaching poetry, due to the language diction and figurative language used by poets, contributing to difficulties in grasping meaning. In addition to language complexities, these participants added difficulties caused by their negative attitudes towards teaching poetry. Mbambo and Hlabisa (2024) argue that formal intervention through professional development is crucial in addressing issues raised in their study about the teaching of poetry. Developing instructors professionally includes keeping up with modern ways of teaching poetry, which include incorporating Artificial Intelligence (AI) in their instruction.

Huang et al. (2023) argue that AI plays a pivotal role in language teaching, which involves the teaching and learning of poetry. Similarly, Semerikov et al. (2021: 1) claim that "AI-assisted language education (AILE) is an emerging interdisciplinary field that integrates AI technology and language education theory and practice." In the Fourth Industrial Revolution (4IR) era, leveraging AI tools in language teaching benefits the language instructor, especially in teaching difficult discipline areas like poetry. Scholars such as Le et al. (2024), Monika (2019), Rahma and Irianti (2024) argue for the use of AI technologies and AI-powered technologies like ChatGPT, Duolingo, and Plotagon in the teaching of language and poetry. AI-powered tools are beneficial to both instructors and students and

are known to improve performance, motivation, pedagogy, as well as bolster problem-solving, critical thinking, creativity, and evaluative skills needed in humanity today.

Although AI has been studied extensively in language teaching, especially its benefits to improve the teaching and learning of poetry, there is a lacuna in what it can do in the context of isiXhosa poetry. It is against such a backdrop that this paper aims to investigate the opportunities, benefits, and challenges of using AI in the teaching and learning of isiXhosa poetry. Therefore, this study endeavours to address this lacuna by situating the integration of AI within the pedagogical context of isiXhosa poetry, thereby contributing to broader discourses on technology-enhanced language education.

Methodology

In this paper, a systematic review, as a research design, was used. According to Page et al. (2021), systematic reviews are critical in synthesising knowledge in a particular field and in addressing key questions in the field that cannot be answered by individual studies. Page et al. (2021: 1) state that the validity and value of the review is determined by "transparent, complete, and accurate account of why the review was done, what they [researchers] did (such as how studies were identified and selected) and what they found (such as characteristics of contributing studies and results of meta-analysis)." To guide the process, Preferred Reporting Items Systematic Reviews and Meta-Analysis (PRISMA) were employed. Purssell and McCrae (2024: 43) note that PRISMA assists "by giving guidance as to what should appear in the report of a systematic review." Therefore, PRISMA principles were used to search, identify, and select material to be included in the study, as well as how to read, extract, and manage data mined from the selected studies (Moher et al., 2015). In using PRISMA, the following methods items, as identified in the PRISMA 2020 statement by Page et al. (2021), were followed:

- Item #5 – Eligibility criteria: specify inclusion and exclusion criteria.
- Item #6 – Information sources: specify all databases consulted to identify studies.
- Item #7 – Search strategy: presentation of full search strategies for all consulted databases.

The methods used are further outlined in the PRISMA flowchart in Figure 1 below.

Eligibility criteria

An electronic search to identify studies for review was conducted. The table below (Table 1) outlines the criterion used to include and exclude material used for review in this study:

Table 1: Inclusion/exclusion criteria

Inclusion	Exclusion
Peer-reviewed articles	Editorials and book reviews
Peer-reviewed book chapters in edited volumes	Conference proceedings
Articles and book chapters written in English	Written in any other official language
Articles and book chapters published from 2021 – 2025	Published prior 2021
Involving basic and higher education	Dissertations and thesis

Inclusion	Exclusion
Discuss opportunities, benefits and challenges of using AI in language education	

Although the timeframe of 2021 to 2025 may seem too restrictive and overlook foundational studies, Tahiru (2021) notes that AI language education has been studied extensively from 2017 to 2020. To avoid reviewing studies that have been reviewed extensively, the focus was on the period between 2021 to 2025. Furthermore, this timeframe was selected for its currency and for revealing new trends in the use of AI in education and language education.

Information sources

Various databases were consulted to identify studies for review in this study. Those are EBSCOhost, Science Direct, ProQuest, and ResearchGate. In consulting the identified databases, the following Boolean search terms were used:

> "Artificial AND intelligence AND in AND education"
>
> "Use AND of AND artificial AND intelligence AND in AND language AND education"
>
> "Artificial AND intelligence AND in AND poetry AND teaching"
>
> "Opportunities AND benefits AND challenges AND using AND artificial AND intelligence AND in AND language AND teaching"

The search yielded a large pool of resources, which were screened using the inclusion/exclusion criteria stated above.

Search strategy

The researchers selected only peer-reviewed journal articles and book chapters in peer-reviewed edited volumes. Crompton and Burke (2023), citing Gough et al., alluded to the fact that peer-reviewed material attests to the quality of the selected data. Furthermore, this ensures validity of data, as all the selected data went through the scrutiny of experts in the field. In collecting data for a systematic review in this study, the search was narrowed to studies published from 2021 to 2025.

Selection process

In this study, only peer-reviewed journal articles and peer-reviewed book chapters in edited volumes were selected to ensure the validity and credibility of data.

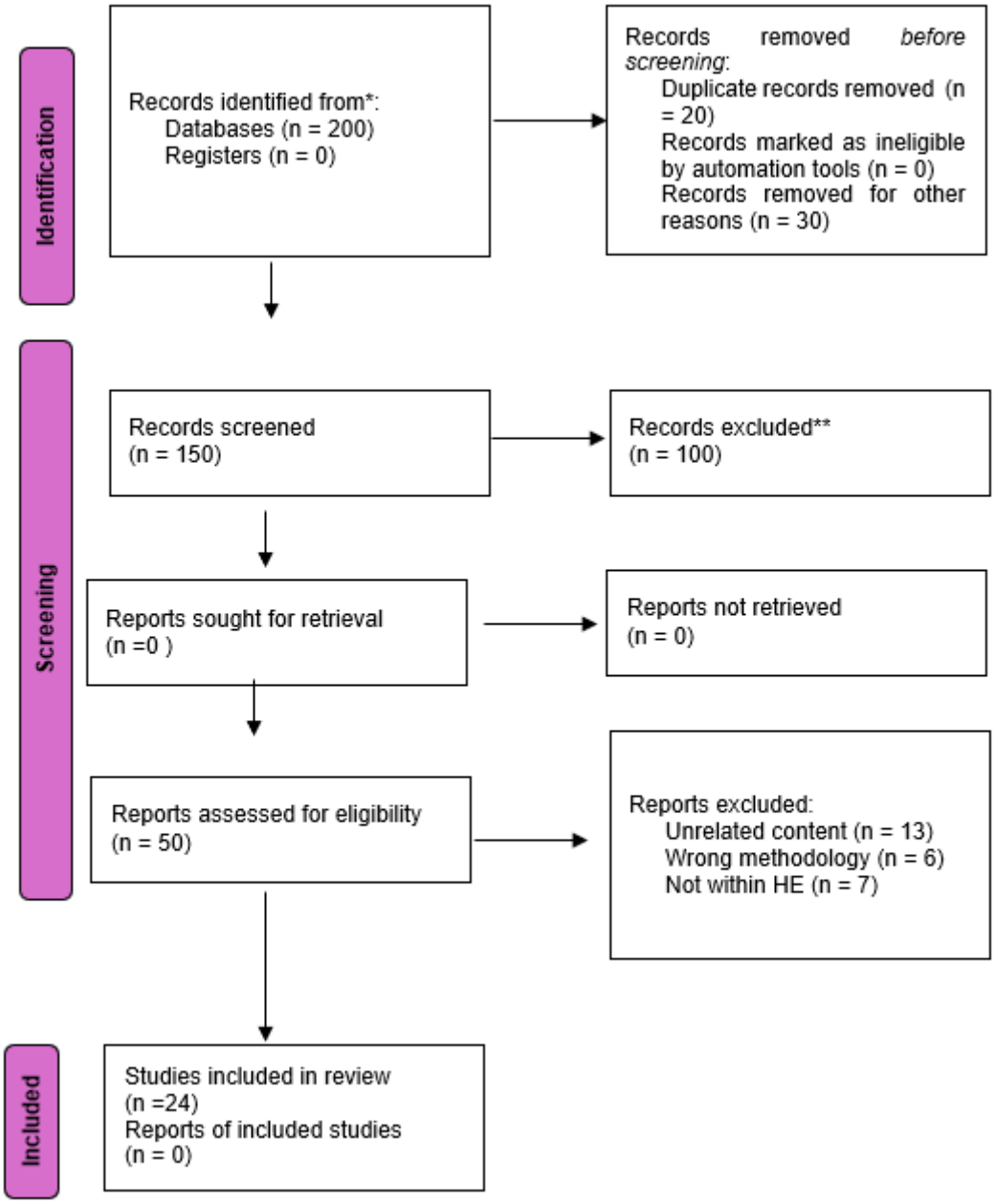

Figure 1: PRISMA flowchart for systematic review which included searches from databases

The theoretical insights that inform the identified design and interpretation of data are discussed below.

Theoretical Framework

The Cultural–Historical Activity Theory (CHAT) was employed as the principal theoretical framework in this paper. This theory is attributed to three scholars or proponents over three stages: Vygotsky at its developmental stage in the 1920s and 1930s, Leontiev, who expanded it in the late 1970s, and, more recently, Engeström, from the late 1980s, further developed the theory (Tkachenko & Ardichvili, 2017). CHAT emphasises that human cognition and learning are not isolated mental processes but are fundamentally socially mediated and

rooted in culture and history (Cong-Lem, 2022; Engeström & Sannino, 2021). Furthermore, CHAT considers learning and teaching as socially mediated activity systems, where artefacts, tools (AI tools), language, rules, community, division of labour, and the historical/cultural context interact to shape cognition and practice (Vygotsky, 1978; Engeström, 1999). Tools (both material and symbolic), language, artefacts, and social interactions play central roles. Within CHAT, an activity system is composed of interacting components: the subject (e.g., lecturers or students), the object (the goal or motive of the activity such as learning isiXhosa poetry), mediating artefacts or tools (AI platforms, language resources), rules, community, division of labour, and outcomes (Engeström, 1987; Daniels et al., 2009).

The key assumption of CHAT is mediation – the notion that technology, language, and culture shape the way individuals engage with learning tasks (Engeström, 1999). Another is historicity, which recognises that activities develop over time are influenced by their cultural and institutional histories (Leontiev, 1978). CHAT also emphasises contradictions within and between elements of the activity system as drivers of innovation and change (Engeström & Sannino, 2010). In educational research, these contradictions often emerge between traditional pedagogical practices and new technological tools, offering opportunities for transformation (Roth & Radford, 2010).

The use of CHAT in this study is justified because the teaching and learning of isiXhosa poetry involve a complex interplay of cultural traditions, pedagogical practices, and emerging digital technologies. Applying CHAT enables the researcher to explore how AI tools can mediate the relationship between lecturers, students, and poetic content within a university context, while also revealing the tensions, opportunities, and challenges that arise from integrating technology into culturally grounded learning (Ballantyne et al., 2022; Batiibwe, 2019). This framework thus provides a holistic lens for analysing the dynamic interaction between culture, technology, and pedagogy in higher education.

Grounded in the principles of CHAT, which emphasises the mediation of learning through tools and social interaction, this framework guided both the research focus and data interpretation. The next section explains how these theoretical assumptions were operationalised through the chosen methodological approach.

Systematic Literature Review

This section reports on the literature that was reviewed to address the aim of this research. There is a growing interest on the use of AI in education, both higher education and basic education.

AI in education

According to Crompton and Burke (2023: 2), it was the work of Turing that began the movement towards AI by describing "the existence of intelligent reasoning and thinking that could go into intelligent machines." Nevertheless, Crompton and Burke (2023), citing Cristianini, reveal that it was McCarthy who coined the term artificial intelligence in 1956, following up on Turing's work of 1937 and 1950. Tahiru (2021: 2) divulges that "In the 1960s, the US Department of Defence took an interest in artificial intelligence and began training computers to mimic basic human reasoning." This early work spread rapidly to various fields such as medicine, entertainment, sci-fi, business, tourism, and in education. Fitria (2021: 134) argues that "In its development, artificial intelligence has also penetrated

the world of education. AI systems allow people to learn with the help of education assistants such as bots." Nevertheless, Tahiru (2021) notes that the pace of adoption of AI in education has been slow in the 20th century and early years of the 21st century. The rise began around 2017. Although there is a rise in the adoption of AI in education, it is still slow in the teaching and learning of isiXhosa.

According to Abdelkader et al. (2024: 103), "AI is a fast-growing topic in education with the potential to significantly expand and improve teaching and learning in higher education". Fitria (2023) is of the same view in stating that AI in education plays the role of expanding intellectual capacity while helping students to effectively and efficiently perform learning activities. In addition, scholars such as Fitira (2023), Tahiru (2021), and Zhai et al. (2021) have identified AI to assist educators and lecturers in facilitating learning. Having identified problem areas in the teaching and learning of isiXhosa poetry in this study, AI places itself in a position to assist in improving the experiences of teachers and learners and their interaction with isiXhosa poetry. As the world shifts into the Fourth Industrial Revolution (4IR) era, isiXhosa poetry needs to be transformed in terms of the tools it uses to mediate teaching and learning, as espoused by CHAT.

Furthermore, the collected data reveals that AI in higher education is used for teaching, learning, assessment, and the management of learning. Azoury and Hajj (2024), focusing on pedagogical enhancement, note that AI is used in HE for learning systems, tutors, and improving curriculum. They argue that "By using AI to customize learning experiences based on students' needs and implementing DT [Digital Transformation] to bridge gaps, the university can cater to a wider range of students, including those from marginalized communities" (Azoury and Hajj, 2024: 156). This thought is predicated on the fact that AI is making education accessible and inclusive. According to Barbie and Milone (1981), Fleming and Mills (1992), and Reid (1987), there are three learning styles - visual, auditory and kinesthetic – and students can either possess one or all of them. Therefore, using AI to mediate activities in isiXhosa poetry classrooms will assist the facilitator in customising learning for each learning style, shaping how students act and learn, and how lecturers instruct their students.

To ensure access to resources needed in any field, scholars such as Crompton and Burke (2023), Fitria (2023), and Javed (2024) allude to the fact that institutions of higher learning use AI-powered Learning Management Systems (LMS). These LMS include Blackboard, Canvas LMS, Google Classroom, iSpring Learn LMS, Moodle, and Ulwazi, to name the notable ones. According to Fitria (2023: 168-169), Blackboard "… is extensively used by professors/lecturers to publish class notes, homework, quizzes, and exams, and it enables students to submit evaluation-related queries and assignments." Qazi et al. (2024) state that before Covid-19, LMSs were limited to supporting roles but during and after 2019-2020, they were thrust into the forefront. Qazi et al. (2024: 3306) posit that these LMSs "enhance the user learning experience through AI-based adaptive learning strategies tailored to individual needs." This view is corroborated by Fitria (2021 & 2023) and Zhai et al. (2021), noting that LMSs create a smart learning environment that is cognizant of students' individual needs.

Furthermore, "AI in the classroom can facilitate this by analysing students' learning styles and providing them with individualized learning aids" (Fitria, 2023: 171). Teaching larger classes makes it difficult to address the individual needs of students; nevertheless, the rise of AI-powered LMSs in HE reaches where the instructor cannot reach, attending to

individual needs of students through teaching, learning, and assessment. In the teaching and learning of isiXhosa poetry, the challenges faced by students range from understanding figurative language, relating the structure of the poem to meaning, analysing the content and the themes of the poem to the effect of literary devices used by the poet. Therefore, isiXhosa poetry instructors can leverage these AI tools to guide students to improve their performance in the challenging areas of teaching and learning isiXhosa poetry, as well as thinking skills by interacting with AI tools to make learning meaningful and effective.

Although LMSs are beneficial for teaching, learning, and assessing, they have their own limitations. With LMSs "it is often difficult to track which student is present and actively participating in the lecture or merely logged into the VCS" (Qazi et al., 2024: 3295). This poses the challenge on learning; while teaching is taking place, there is limited guarantee to how much learning is taking place. These scholars further state that "during the conduct of the exams, it is difficult to detect any means of unfair cheating by the students" (Qazi et al., 2024: 3295). This leads stakeholders to question the validity of assessment results. Furthermore, Barikzai et al. (2024) note that at times, LMSs fall short in supporting more interactive, constructive or active learning pedagogies. The teaching and learning of isiXhosa poetry is particularly affected by this challenge, given its dependence on dynamic exchanges between teachers, learners, and peers. Within a constructivist framework, instructors seek to foster active engagement by prompting students to relate poetic meaning to their own experiences and interpretations. Therefore, in using LMSs, the role of the instructor must be fully established to guide students through their thinking. This is supported by CHAT, as it encourages the accommodation of multiple perspectives, understandings and views.

Literature shows that Intelligent Tutoring Systems (ITS) has been leveraged in HE to lend a hand to the traditional tutoring programmes used by universities to assist students. Crompton and Burke (2023: 16) define ITS as "… adaptive instructional systems that involve the use of AI techniques and educational methods." AI techniques include interactive learning, tailored and immediate feedback, and immediate access to resources, which can be integrated into teaching methods and activities to ensure learning has taken place. According to Tahiru (2021), ITS promotes collaborative learning while improving motivation and learning in students. According to Riaz and Din (2023: 94), "Collaboration is known [as] one of the most important and essential skills of 21st century learning skill set." Therefore, through the use of ITS, isiXhosa poetry students stand to benefit as individual and as a group, in that they can engage with the system about the content they are learning, do activities and get immediate feedback on their learning, while improving their understanding of content. ITS can further be used for preparing for examinations, tests and assignments.

In preparing for exams, besides using ITS, literature reveals that students in HE can use an AI tool called Smart Content, with Apps like Cram101 and JustTheFacts101. According to Fitria (2023: 169), "This AI technology functions to divide and find content material and digital books that are already programmed virtually more easily and quickly." In addition, Tahiru (2021), citing Fagella, argues that this tool is capable of condensing a textbook to manageable content, such as true-or-false questions to prepare for exams. This will make it easy for students to retrieve the needed content (Fitria, 2023), instead of having to page through a textbook on the eve of an exam. These scholars reveal that this tool also incorporates videos in its storage of content. Such a tool requires higher-order thinking skills, as students will need to critically analyse, evaluate, and synthesise (another set of important skills in the 21st century) the content they will need to prepare for exams.

Although the above-stated systems are crucial for teaching and learning in HE, they are not without fault. Davar et al. (2025) argue that AI tools have limitations, such as understanding human emotions and language limitations. In addition, these tools lack the capabilities of understanding the emotions of students and instructors at a particular time. Furthermore, they offer an unfair advantage to English students, while leaving behind users of other languages. The fact that most research is written in English leads to indigenous languages playing catch-up in terms of datasets used to train these AI-powered tools (Makeleni et al., 2023; Simunyu et al., 2021; Wang, 2023). These scholars further reveal that 'The lack of African languages' datasets is discouraging many NLP practitioners to start from scratch and the task will have to be taken up by African researchers because of the low economic interest that our languages represent for top companies driving changes in NLP" (Simunyu et al., 2021: 1). IsiXhosa is among those African languages with low datasets.

Although literature above states benefits of AI on motivation, creativity, and critical thinking, Law (2024) draws attention to contrasting viewpoints, which state that AI has a potential of compromising the very psychological aspects it is known to enhance. Resulting from the literature reviewed in Law's study, this claim resulted from the discovery that some students use AI as a shortcut for written activities, even though it lacks contextual understanding. This researcher further states that AI fails to assess higher order thinking skills. Law (2024: 11) concludes by stating that despite these drawbacks stated in literature, "the potential advantages offered by GenAI programs remain uncontested." Therefore, the advantages of using AI in education outweigh the disadvantages, leaving instructors with opportunities to leverage AI in their teaching.

AI in language education

This subsection focuses on literature on the use of AI in language education, with special reference to poetry teaching and learning.

Law (2024) encourages language educators to solicit the use of Generative AI in their teaching. This view is corroborated by Makeleni et al. (2023: 159) who state that "using AI in language education offers numerous benefits that outweigh any potential or perceived costs". The paucity of empirical research on the use of AI in isiXhosa teaching attests to the low levels of adoption of AI in the teaching and learning of isiXhosa. In such cases, Law (*Op cit*.) encourages professional development, citing it as crucial in ensuring informed decision-making and effective integration of Gen AI tools in the teaching and learning of languages. This view is at the heart of the CHAT which calls for the evolution of educators through development and transformation.

The use of AI in language education has been explored mostly in English teaching and learning. This can be attributed to the increasing popularity of AI in the teaching and learning of English in HE (Javed, 2024) and that "many AI systems are developed primarily in English" (Makeleni et al., 2023: 162). In the study, Javed (2024) examines the development of AI in the teaching and learning of English, considering its uses, advantages, difficulties, and potential. Javed (2024: 276) argues that "to improve teaching and learning, AI technologies are being incorporated into educational environments more and more." In addition, the researcher supports the view that AI technologies offer individualised learning programmes which cater for individual students, and they adjust the level of content in accordance with how the learner is performing. This approach by AI can be compared to scaffolding where an instructor incrementally introduces content in distinct segments

until a student can master new concepts without support. Javed (2024) noted the following as advantages of using AI in teaching and learning English: a tailored learning system, individualised learning outcomes, instant feedback, deeper involvement in the learning process, and machine translation. Although the use of AI has multiple advantages, Javed (2024) notes the challenge it poses in teaching and learning of English, first being the digital divide manifesting itself in disparities in access and proficiency among pupils, as determined by socioeconomic status of students. The researcher adds other challenges, such as ethical challenges. Javed (2024) was generic in his approach, looking at the use of AI in a broad sense in English teaching and learning. Nevertheless, the current study is specific in poetry teaching and learning, as this aspect of language teaching is unique and is used as an independent field in other institutions.

Wang et al. (2025) examined the potential of generative AI to enhance the teaching of writing in Chinese poetry. Their research was a result of discovering challenges in the teaching and learning of poetry. Citing Liao, they state that students struggle with inspiration and are finding difficulties in expressing themselves in their poetry writing. The teaching and learning of isiXhosa poetry in HE includes students creating their own poetry on current issues, about leaders in the society, things they love and hate, etcetera. The challenges faced by Chinese students are similar to those faced by students doing isiXhosa poetry. Wang et al. (2025: 9) discovered that "The integration of generative AI in Chinese poetry instruction had a significant positive impact on students' learning interest and effectiveness." In addition, they learned that AI-generated visuals deepen students' ability to understand poetic imagery and foster creative expression in poetry writing. Furthermore, they posit that AI supports multimodal learning, enhancing students' linguistic creativity and enriching poetic sensibilities. According to these scholars, AI addressed the challenges they identified in teaching Chinese poetry. Since poetry taught in isiXhosa was mostly written before the time of the students, the diction used is difficult to understand and the imagery used is foreign to their imagination.

Le et al. (2024: 51) argue that "The emergence of AI language models, such as ChatGPT, has opened up new possibilities for personalized learning and instructional support." These scholars explored the use of ChatGPT in developing e-portfolios. Their enquiry is premised on the fact that e-portfolios "assist students in improving writing accuracy by utilizing editing, spelling, and grammar-checking features available on the e-portfolio platform" (Le at al., 2024: 55, citing Meshkat and Goli). E-portfolios form part of isiXhosa poetry teaching and learning, as students can be expected to produce a portfolio of their journey in the module, which will include written pieces of poetry and commentary on others poetry. According to Le et al. (2024: 59), "ChatGPT, with its chatbot interface, emerges as a unique and user-friendly e-portfolio platform, catering to the needs of students in a manner that minimizes frustration and enhances their overall experience." ChatGPT has proven to be among leading AI tools to be leveraged in language teaching and learning. In parading the benefits of AI, Le et al. (2024) note its challenges as well. They state that overreliance on AI-generated content can diminish critical thinking and look down upon human-generated content. Such challenges can be averted by educating students and instructors on the use of AI tools in language teaching.

In addition, Davar et al. (2025) explored the benefits and challenges of using AI chatbots. They state that using AI for role playing helps to improve language proficiency through speaking. In addition, Davar et al. (2025) argue that AI chatbots help students overcome obstacles in language learning. In the teaching and learning isiXhosa poetry, students can

face different obstacles, such as understanding literary devices used, dealing with structure used to convey meaning, diction. Therefore, chatbots, such as ChatGPT, are tools that can be used to help students understand poetry.

Mellor (2024: 247) examined the "ways to implement AI-supported paradigm as one of the solutions to mitigate the challenges in Arab higher education, drawing on AI tools, such as ChatGPT." The researcher notes that ChatGPT can be used to provide tailored feedback to each student by interacting with them in a conversation manner. Such conversations allow each student to be met at their point of educational need. According to Mellor (2024: 253), "responses [from ChatGPT] can be used as prompts for complex or difficult questions, enabling students to think more deeply about the topic and develop their critical thinking skills." Students or the instructor can prompt each other about the poem to be studied in isiXhosa poetry class, as a pre-reading activity on ChatGPT, encouraging students not to be mere consumers of information from AI but pushing them to analyse, evaluate, and debate the information. The prompts can be about aspects of poetry, such as themes, style, and recent trends in isiXhosa poetry. Ji et al. (2022) and Li (2023) call for a balance between the use of technological tools and ethical considerations. Academic integrity must be upheld at all times. In allowing students to use ChatGPT, to brainstorm research topics, an isiXhosa instructor must emphasise the importance of avoiding copying, plagiarism, and falsifying data.

In their study, Al-Kadi and Ali (2024) paired ChatGPT with two other large language models (LLM) in the teaching and learning of English, Gemini, and Copilot. They argue that these LLMs improve the command of language and they innovatively support teaching. Furthermore, they noticed that the use of these AI tools improved students' motivation to learn, an aspect that plays a crucial role in student success. Al-Kadi and Ali (2024.) state that the integration of these tools improves students' engagement with learning, which results in learning motivation.

In addition, Al-Kadi and Ali (*Op cit*.) noticed that these LLMs can be adapted to meet each student's language needs. Given the heterogeneous nature of the difficulties students face in engaging with isiXhosa poetry, the integration of intelligent tutoring systems such as ChatGPT, Gemini, and Copilot offers a personalised approach to addressing varied learning styles and instructional requirements. Furthermore, they discovered that these tools provide immediate feedback, minimising frustration for students when stuck in one place. Immediate feedback ensures progression in teaching and learning. Finally, Al-Kadi and Ali (2024: 6) discovered that "LLMs lessen reliance on formal education; learners become less dependent on their teachers, as these technological advances provide language learning opportunities beyond formal and institutionalized language education." According to Le et al. (2024: 58), citing Abdelghani, "using natural language prompts and large language models can help students develop their ability to ask thought-provoking questions." This skill is beneficial in learning isiXhosa poetry, as analysis of poetry requires one to interrogate a poem to unearth deep-lying meanings.

Makeleni et al. (*Op cit*.) focused on the challenges confronting academics on the use of AI in language education in the Global South. Language education in the Global South includes the teaching and learning of isiXhosa. Through literature review methodology, these scholars discovered four categories of challenges faced by academics. These challenges are part of the contradictions that cause tensions between what AI tools promise to bring and the reality faced by languages in the Global South, including isiXhosa. Firstly, they

discovered that AI tools have limited language options, with English, Chinese, Spanish, and French dominating the space. The result is difficulty "for people in the Global South to access digital tools and services in their own languages" (Makeleni et al., 2023: 162).

The second challenge Makeleni et al. (*Op cit*.) alluded to is students using AI to cheat in assignments and examinations. They note that, because African languages like isiXhosa are not well-represented in AI datasets, the accuracy of detecting cheating and plagiarism is limited. The third challenge discovered by these scholars is AI biases that tend to penalise those who do not use standardised languages. The fourth challenge is that AI leads "to laziness among students and lecturers" (2023: 164). This is caused by what Le et al. (*Op cit*,) refer to as overreliance on AI. These four challenges reveal tensions between the promise of decoloniality and the reality of linguistic inequality in the AI tools that are meant to mediate teaching and learning between subjects. Historical linguistic inequalities in South Africa, as a result of colonialism and apartheid, are resurfacing in the digital space, as African languages like isiXhosa continue to be under resourced. Practitioners in isiXhosa need to collaborate with computer scientists to speed up the process of bridging the digital linguistic gap.

Discussion

The opportunities, benefits, and challenges of integrating AI into the teaching and learning of isiXhosa poetry were investigated through the lens of CHAT. This theory conceptualises learning as a mediated activity shaped by the interplay between subjects (students and instructors), tools (AI technologies), community, rules, and division of labour (Engeström, 1987, 2001; Vygotsky, 1978). From this perspective, AI functions as a mediating artefact that reshapes the activity system of the teaching and learning of isiXhosa poetry by altering not only instructional practices but also student agency, community roles, and power relations.

Opportunities and benefits of using AI

Student's performance

The findings of the study demonstrate that AI is leveraged in language teaching to improve students' performance. According to Jaxa (2024), the teaching and learning of isiXhosa poetry is marred by the poor performance of students. AI-powered tools such as intelligent tutoring systems (ITS), speech recognition platforms, automated writing feedback, and adaptive learning applications enable students to receive individualised instruction that responds to their pace, proficiency level, and preferred learning style. Instructors and students in the teaching and learning of isiXhosa poetry can leverage these tools, as they are known for customised learning, which is tailored to students' individual needs, known to engage with students through interactive learning, and known to provide immediate feedback. We argue that instructors who teach isiXhosa poetry must leverage these tools in their teaching and allow students to use them in class to engage in constructive learning. Each student, instead of getting generalised instruction from the instructor, will get instruction that will meet individual needs, while providing immediate feedback. The benefits have been experienced in higher education and in language education, and the teaching and learning of isiXhosa poetry must not lag. The following section outlines

the opportunities and benefits of using AI against the intended outcomes of education in South Africa.

Improving skills, knowledge, attitudes and values

The South African education system is underpinned by the objective of developing learners' skills, knowledge, attitudes, and values that facilitate a seamless transition from the classroom to the professional sphere (Department of Basic Education, 2012). The findings of this study indicate that AI-powered tools significantly enhance students' critical thinking when engaging with isiXhosa poetry in higher education contexts. Interpreted through the lens of CHAT, AI functions as a mediating artefact that reshapes the dynamics of the learning activity system. Traditionally, students (subjects) face contradictions when attempting to interpret poetic language, particularly where idiomatic, metaphorical, or culturally embedded expressions are present. However, AI tools such as ChatGPT expand the mediational means, offering instant access to contextual explanations, translations, and alternative interpretations. This aligns with Mellor's (2024: 253) argument that AI prepares students "for the real world by promoting critical thinking skills," and with Iqbal and Pearl's (2024) findings on AI as a catalyst for analytical reasoning. Within the division of labour, AI shifts the instructor's role from sole knowledge provider to co-mediator, allowing students greater autonomy in inquiry-driven exploration.

Furthermore, the rules governing academic engagement are reconfigured as students increasingly rely on dialogic interaction with technological tools rather than linear consumption of prescribed texts. Consequently, critical thinking should not be viewed merely as an individual cognitive disposition, but as an emergent, socially mediated outcome produced through tool-assisted participation within the broader community of learning. In this way, AI becomes not only an instructional aid but a transformative component of the educational activity system. Existing literature confirms that AI enhances problem-solving and critical analysis, which are essential when engaging with isiXhosa poetry that often addresses current social issues. By breaking down complex datasets, AI supports students in deconstructing the poem's social themes, identifying causes and exploring possible solutions. Within the activity system, AI also extends the mediational means available for creative production, enabling students to compose original poems using higher-order cognitive processes, as required by Bloom's Taxonomy. Moreover, AI has been found to increase learner motivation, further strengthening sustained engagement with poetic texts. This is mediated by AI tools, which are discussed in the following section.

AI tools

The findings of this study indicate that Learning Management Systems (LMSs) and AI-powered tools are among the most prominent technologies currently used in language education and hold substantial potential for the teaching and learning of isiXhosa poetry. Interpreted through CHAT, these technologies function as mediating artefacts within the broader educational activity system (Engeström, 1987). They reshape how learners (subjects) engage with poetic analysis (object) by restructuring the tools, rules, and division of labour within the classroom.

LMS platforms such as Moodle, Blackboard, and Canvas enable instructors to design interactive learning environments, integrating discussion forums, collaborative spaces, and multimedia resources. These affordances align with CHAT's view of mediation, where digital tools enable higher levels of participation and knowledge transformation (Vygotsky, 1978; Ngubane-Mokiwa & Letseka, 2023). Moreover, the ability to automate assessment and deliver immediate and individualised feedback transforms the division of labour, reducing the cognitive and administrative burden on instructors who traditionally had to manually engage with large cohorts holding diverse interpretations of a poem. This form of algorithmic mediation supports differentiated learning pathways, allowing students to progress at their own pace (Mishra, Gupta & Reddy, 2024).

Beyond LMSs, the reviewed literature and study findings highlight the growing integration of AI-powered tools such as ChatGPT, Gemini, and Copilot within language education (Mellor, 2024; Iqbal & Pearl, 2024). These tools serve as dialogic partners, engaging learners in real-time questioning, explanation, and elaboration - activities core to the development of critical and reflective thinking. Within a CHAT framework, AI introduces new mediational means that extend learners' capacity to interpret complex poetic and literary devices, access cultural context, and generate alternative readings of isiXhosa poems. These AI-powered tools, armed with rich datasets, serve as an added resource in the teaching and learning of isiXhosa poetry, supporting knowledge acquisition.

Consequently, AI and LMSs not only support knowledge acquisition but also actively reconfigure the activity system, repositioning students as co-constructors of meaning rather than passive recipients. This aligns with the CHAT view that learning emerges through socially mediated interaction with tools and community rather than through isolated cognitive processing.

Although AI offers opportunities and benefits in relation to the teaching and learning of isiXhosa poetry, it does not come free from challenges, and the section below discusses challenges gleaned from the literature reviewed.

Challenges of using AI

Access

The findings of this study highlight access to AI tools as a critical challenge in HE and language learning. From the perspective of CHAT, this challenge can be conceptualised as a contradiction within the activity system, particularly between the tools (AI-enabled devices and digital platforms) and the subject-community configuration of students and instructors (Engeström, 2001; Vygotsky, 1978). Many students majoring in isiXhosa originates from historically disadvantaged backgrounds, which constrains their access to AI-powered devices and reliable internet connectivity, reflecting broader societal inequalities prevalent in South Africa (Sharonova & Avdeeva, 2024). Although the COVID-19 pandemic prompted universities to provide data and devices to mitigate these disparities, learners in remote or under connected areas continue to face minimal or no network coverage. Within a CHAT framework, this limits the mediating function of digital tools, reducing students' ability to engage fully in the language learning activity system. As a result, learners may struggle to access instructional content, participate in collaborative knowledge construction, or adequately prepare for assessments, thereby exacerbating educational inequalities

(Engeström & Sannino, 2010). Addressing these contradictions necessitates restructuring the activity system, for instance, through the provision of offline AI-enabled resources, the development of community-supported learning hubs, and policy-level interventions to expand digital infrastructure. Such measures ensure that AI functions as an inclusive mediating artefact, supporting equitable participation rather than reinforcing existing barriers to learning.

Ethics

Ethical challenges associated with AI in HE and language education can be understood as contradictions within the activity system, particularly between AI tools, rules (academic integrity norms), and students and instructors (Engeström, 2001; Vygotsky, 1978). Literature documents concerns regarding cheating, plagiarism, and overreliance on AI for writing and research tasks, which raise critical questions about academic integrity (Davar et al., 2025; Makeleni et al., 2023). From a CHAT perspective, these issues reflect tensions in the mediating function of AI: while AI is designed to support learning, its misuse can disrupt knowledge construction and the development of higher-order cognitive skills, resulting in missed learning opportunities. Detecting AI-generated content remains technically challenging, exacerbating the contradiction between assessment practices and tool use (Makeleni et al., 2023).

To address these contradictions, instructors can reconfigure the activity system by promoting ethical, pedagogically aligned AI use. For example, in the teaching and learning of isiXhosa poetry, AI can function as a mediating artefact for pre-reading, brainstorming, and generating reference leads, rather than producing final assignments. Learners should be guided to formulate AI prompts that generate verifiable information, thereby fostering critical engagement and responsible digital literacy. Through such interventions, AI becomes an enabler of expansive learning, supporting both student autonomy and adherence to academic norms, while mitigating ethical risks inherent in technology-mediated language education (Engeström & Sannino, 2010).

Limitations of AI tools

The findings established that AI has limitations in terms of understanding human emotions and the use of indigenous languages. In these AI-powered tools and chatbots, there is limited datasets to be accessed by students who are learning isiXhosa. Nevertheless, there is a growing trend of developing AI-powered tools that use all official languages like SHAKAI. Such tools close the language gaps and reduce inequalities.

Conclusion and Recommendations

This study has highlighted the significant opportunities, benefits, and challenges associated with the use of AI in the teaching and learning of isiXhosa poetry. The findings suggest that AI-powered tools can serve as effective mediating artefacts that enhance problem-solving, critical thinking, analytical, and evaluative skills among students – competencies that are not only essential across the curriculum but also highly relevant to professional contexts in which students will operate. The integration of AI in the teaching and learning

of isiXhosa poetry, when implemented thoughtfully, offers the potential to personalise learning, support self-regulated study, and provide continuous feedback, thereby fostering deeper engagement with isiXhosa poetry and its cultural nuances.

However, the study also underscores the challenges and contradictions that arise from integrating AI. These include digital inequalities, limited access to AI tools among students from historically disadvantaged backgrounds, and ethical concerns such as plagiarism and misuse of AI for assessments. From a pedagogical standpoint, instructors must actively guide students on the responsible and ethical use of AI, including employing AI for pre-reading, idea generation, and reference exploration rather than for completing assignments or assessments. This guidance ensures that AI becomes a supportive educational tool rather than a source of academic misconduct.

To translate these insights into practice, institutions should adopt multi-level strategies:

- Access and infrastructure: Develop initiatives to provide affordable or subsidised AI-enabled devices and reliable internet access, particularly for students in remote regions. Offline or hybrid AI tools can also be implemented to bridge connectivity gaps.
- Capacity-building and pedagogy: Train instructors to integrate AI effectively into curriculum design, emphasising ethical usage, critical engagement, and scaffolding of learning activities in alignment with pedagogical goals.
- Policy and governance: Establish clear institutional policies and guidelines on AI use in teaching, learning, and assessment, including mechanisms to detect and mitigate misuse while promoting digital literacy.
- Cultural and linguistic inclusivity: Prioritise the development of AI tools that support isiXhosa and other under-resourced languages, ensuring that technological innovation aligns with cultural preservation and equitable educational outcomes.
- Monitoring and evaluation: Implement ongoing assessment of AI integration to identify challenges, evaluate impact on learning outcomes, and refine both instructional and policy interventions.

In conclusion, AI holds considerable promise for enhancing the teaching and learning of isiXhosa poetry, but its successful adoption depends on careful alignment with pedagogical objectives, ethical practice, equitable access, and supportive institutional policies. By addressing these areas proactively, higher education institutions can harness AI as a transformative tool that fosters both academic excellence and culturally responsive language education.

Reference list

Abdelkader, A. A. M., Hassan, H., and Abdelkader, M. (2024). The role of Artificial Intelligence in designing higher education courses: Benefits and challenges. In: M. Lytras, A. Alkhaldi, S. Maalik, A. C. erban, and T. Aldosemani (eds). *The Evolution of Artificial Intelligence in Higher Education: Challenges, risks, and ethical consideration*. (102-116). Leeds: Emerald Publishing Limited. https://doi.org/10.1108/978-1-83549-486-820241005

Al-Kadi, A., & Ali, J. K. M. (2024). 'A holistic approach to ChatGPT, Gemini, and Copilot in English learning and teaching'. *Language Teaching Research Quarterly*, 43, 155-166. https://doi.org/10.32038/ltrq.2024.43.09

Azoury, N. and Hajj, C. (2024). Digital transformation in higher education: Best practices and challenges. In: M. Lytras, A. Alkhaldi, S. Maalik, A. C. erban, and T. Aldosemani (eds). *The Evolution of Artificial Intelligence in Higher Education: Challenges, risks, and ethical consideration*. (148–167). Leeds: Emerald Publishing Limited. https://doi.org/10.1108/978-1-83549-486-820241008

Ballantyne, J., Livingston, K., & Garraway, J. (2022). 'A cultural-historical activity theory as a framework for exploring pre-service teachers' use of an intelligent tutoring system for English language proficiency'. *Africa Education Review*, *19*(1), 55–70. https://doi.org/10.1080/18146627.2022.2150245

Barbe, W. B., & Milone, M. N. (1981). 'What we know about modality strengths'. *Educational Leadership*, 38(5), 378–380.

Barikzai, S., Bharathi, V., & Perdana, A. (2024) 'Challenges and strategies in e-learning adoption in emerging economies: a scoping review', *Cogent Education*, 11(1), 1–23. https://doi.org/10.1080/2331186X.2024.2400415

Batiibwe, M. S. (2019). *Using cultural historical activity theory to understand how emerging technologies can mediate teaching and learning in a mathematics classroom: A review of literature. Smart Learning Environments, 6*(1), 1–13. https://doi.org/10.1186/s41039-019-0110-7

Cong-Lem, N. (2022). 'Vygotsky's, Leontiev's and Engeström's Cultural Historical (Activity) Theories: Overview, clarifications and implications'. *Integrative Psychological and Behavioral Science*, (2022)56, 1091–1112. https://doi.org/10.1007/s12124-022-09703-6

Crompton, H. and Burke, D. (2023). 'Artificial intelligence in higher education: the state of the field', *International Journal of Educational Technology in Higher Education*, 20(22), 1–22. https://doi.org/10.1186/s41239-023-00392-8

Daniels, H., Edwards, A., Engeström, Y., Gallagher, T., & Ludvigsen, S. R. (2009). *Activity theory in practice: Promoting learning across boundaries and agencies.* London: Routledge.

Davar, N. F., Dewan, M. A. K., and Zhang, X. (2025). 'AI chatbots in education: Challenges and opportunities'. *Information,* 16(235), 1–25. https://doi/10.3390/info16030235

Department of Basic Education. (2012). National Curriculum Statement Grades R–12 (Policy). Republic of South Africa. Retrieved from https://www.education.gov.za. Accessed: June 20, 2025.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Tokyo: Orienta-Konsultit.

Engeström, Y. (1999). *Perspectives on activity theory.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511812774

Engeström, Y. (2001). 'Expansive learning at work: Toward an activity theoretical reconceptualization', *Journal of Education and Work*, 14(1), 133–156. https://doi.org/10.1080/13639080020028747

Engeström, Y., & Sannino, A. (2010). 'Studies of expansive learning: Foundations, findings and future challenges', *Educational Research Review*, *5*(1), 1–24. https://doi.org/10.1016/j.edurev.2009.12.002

Engeström, Y., & Sannino, A. (2021). 'From mediated actions to heterogenous coalitions: four generations of activity-theoretical studies of work and learning', *Mind Culture and Activity*, 28(1), 4–23. https://doi.org/10.1080/10749039.2020.1806328

Fitria. T. N. (2021). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. Proceeding Seminar Nasional & Call for Paper STIE AAS, Surakarta, 134–147.

Fitria, T. N. (2023). 'The use of artificial intelligence in education (AIED): Can AI replace the teacher's role?', *Epigram*, 20(2), 165–187. https://doi.org/10.32722/epi.v20i2.5711

Fleming, N. D., & Mills, C. (1992). 'Not another inventory, rather a catalyst for reflection', *To Improve the Academy: A Journal of Educational Development*, 11, 137–155. https://doi.org/10.1002/j.2334-4822.1992.tb00213.x

Gxekwa, N. & Satyo, N. (2017). 'The use of isiXhosa children's poetry as a tool to integrate literacy, mathematics and life skills in Foundation Phase: Grade R–3', *South African Journal of Childhood Education,* 7(1), 1–7. https://doi.org/10.4102/sajce.v7i1.530

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). 'Trends, Research Issues and Applications of Artificial Intelligence in Language Education', *Educational Technology & Society*, 26(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).0009

Iqbal, N., & Pearl, J. (2024). *Harnessing AI tools for critical thinking: ChatGPT's impact on teaching strategies and education integration*. Available at: https://doi.org/10.13140/RG.2.2.18034.31682 https://www.researchgate.net/publication/382399833 [Accessed 20 February 2025].

Jaxa, N. P. (2024). 'Pedagogical stylistics: Teaching isiXhosa poetry at further education training phase using text world theory approach', *Journal of Culture and Value\s in Education*, 7(1), 118–133. https://doi.org/10.46303/jcve.2024.8

Javed, F. (2024). The evolution of Artificial Intelligence in teaching and learning of English language in higher education: challenges, risks, and ethical considerations. In: M. Lytras, A. Alkhaldi, S. Maalik, A. C. erban, and T. Aldosemani (eds). *The Evolution of Artificial Intelligence in Higher Education: Challenges, risks, and ethical consideration*. (271–300). Leeds: Emerald Publishing Limited. https://doi.org/10.1108/978-1-83549-486-820241015

Ji, H., Han, I. & Ko, Y. (2022). 'A systematic review of conversational AI in language education: focusing on the collaboration with human teachers', *Journal of Research on Technology in Education*, 55(1), 48–63. https://doi.org/10.1080/15391523.2022.2142873

Khan, S. (2020). 'Why and how to use a poem in ELT classroom', *International Online Journal of Education and Teaching (IOJET)*, 7(3). 803–809. https://iojet.org/index.php/IOJET/article/view/807

Le, A. N., Nguyen, V. N., Nguyen, M. T., & Bo, L. K. (2024). In S. Papadakis (ed), *IoT, AI, and ICT for Educational Applications: Technologies to Enable Education for All*. (51–76). EAI/Springer Innovations in Communication and Computing. Available at: https://doi.org/10.1007/978-3-031-50139-5. Accessed: July 10, 2025.

Leontiev, A. N. (1978). *Activity, consciousness, and personality.* New Jersey: Prentice-Hall.

Li, N. (2023). Ethical considerations in Artificial Intelligence: A comprehensive discussion from the perspective of computer vision. *SHS Web of Conferences*, 179, 04024. https://doi.org/10.1051/shsconf/202317904024

Makeleni, S., Dlamini, T., & Ncube, P. (2023). 'AI in higher education: Plagiarism, learning, and ethical challenges', *Journal of Learning Analytics and Educational Technology*, 8(3), 115–130.

Makeleni, S., Mutongoza, B., & Linake, M. (2023). Language Education and Artificial Intelligence: An Exploration of Challenges Confronting Academics in Global South Universities. *Journal of Culture and Values in Education, 6(2), 158-171.* https://doi.org/10.46303/jcve.2023.14

Makhenyane, L. E. (2024). 'The use of augmented reality in the teaching and learning of isiXhosa poetry', *Journal of the Digital Humanities Association of Southern Africa,* 5(1). https://doi.org/10.55492/dhasa.v5i1.5023

Mbambo, K. & Hlabisa, M.V. (2024). 'South African rural high school teachers' experiences of teaching English poetry', *Journal of Education*, (97), 261–281. https://doi.org/10.17159/2520-9868/i97a13

Mellor, N. (2024). Using AI to develop capabilities in Arab universities. In: A. Al-Marzouqi, A. A. Salloum, M. Al-Saidat, A. Aburayya, & B. Gupta (eds.). Artificial Intelligence in Education: The Power and Dangers of ChatGPT in the Classroom. (247–258). Cham: Springer. https://doi.org/10.1007/978-3-031-52280-2

Mishra, P., Gupta, R., & Reddy, A. (2024). 'Automated formative feedback in higher education: A systematic review of AI-driven assessment tools', *Education and Information Technologies*, 29(2), 2341-2360.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Japanese Pharmacology and Therapeutics,* 1(4), 1-9. https://doi.org/10.1186/2046-4053-4-1

Molgadali, M., Osmanova, Z. & Nurgaziyev, T. (2024). 'A Meta-Analysis of the Impact of Innovative Poetry Teaching Methods on Reading, Writing, and Comprehension Skills', *Journal of Social Studies Education Research*, 15(5), 169-195.

Monika, M. (2019). 'Effective Teaching Of Poetry Through Blended Learning Using Multimedia Technology', *Think India (Quarterly Journal)*, 22(07), 21-25.

Ngubane-Mokiwa, S., & Letseka, M. (2023). 'Digital pedagogies and inclusive participation in South African universities: An activity theory perspective', *South African Journal of Higher Education*, 37(4), 85-99.

Ou, A. W., Stöhr, C. & Malmström, H. (2024). 'Academic communication with AI-powered language tools in higher education: From a post-humanist perspective', *System*, 121(2024), 1-14. https://doi.org/10.1016/j.system.2024.103225

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ,* 372(71), 1-9. https://doi:10.1136/bmj.n71

Purssell, E. and McCrae, N. (2024). *How to perform a systematic literature review: A guide for health and social care researchers, practitioners and students* (2nd ed.). Cham: Springer. https://doi.org/10.1007/978-3-031-71159-6

Qazi, S., Kadri, M. B., Naveed, M., Khawaja, B. A., Khan, S. Z., Alam, M. M., and Su'ud, M. M. (2024). *Computers, Materials & Continua*, 80(2), 3289-3314. https://doi.org/1032604/cmc.2024.048893

Rahma, T., & Irianti, L. (2024). 'Utilizing Quizzes on the Duolingo to Promote Students Vocabulary Mastery', *Journal of English Education Program (JEEP)*, 11(2), 167-174. https://doi.org/10.25157/(jeep).v11i2.15564

Reid, J. M. (1987). 'The learning style preferences of ESL students', *TESOL Quarterly*, 21(1), 87-111. https://doi.org/10.2307/3586356

Riaz, M. & Din, M. (2023). 'Collaboration as 21st century learning skill at undergraduate level', *Sir Syed Journal of Education & Social Research*, 6(1), 93-99. https://doi.org/10.36902/sjesr-vol6-iss1-2023(93-99)

Roth, W.-M., & Radford, L. (2010). *Re/thinking the zone of proximal development (ZPD): A dialectical perspective.Mind,Culture,andActivity,*17(4),299-307.https://doi.org/10.1080/10749030903338509

Semerikov, S. O., Struik, A. M. & Shalatska, H. M. (2021). 'AI-assisted language education: critical review', *Educational Dimension*, 4, 1-7. https://doi.org/10.31812/ed.623

Sharonova, S., & Avdeeva, E. (2024). Smart education: Social risks and challenges. In S. Papadakis (ed), *IoT, AI, and ICT for Educational Applications: Technologies to Enable Education for All*. (99-118). EAI/Springer Innovations in Communication and Computing. https://doi.org/10.1007/978-3-031-50139-5

Simunyu, K., Kalipe, G., Orlic, D., Abbott, J., Marivate, V., Freshia, S., Sibal, P., Neupane, B., Adelani, D. I., Taylor, A., Ali, J. T., Degila, K., Balogoun, M., Diop, T. I., David, D., Fourati, C., Hatem Haddad, H. and Naski, M. (2021). *AI4D - African Language Program*. arXiv:2104.02516v1 [cs. CL]. Available at: https://www.researchgate.net/publication/352787649_AI4D___African_ Language_Program. Accessed: July 10, 2025.

Tahiru, F. (2021). 'AI in Education: A Systematic Literature Review', *Journal of Cases on Information Technology*, 23(1), 1–20. https://doi.org/10.4018/JCIT.2021010101

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Wang, W. (2023). Different natural language, equal importance. *Patterns*, 4(August 11), 1–2. https:// doi.org/10.1016/j.patter.2023.100821

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J., Yuan, J., & Li, Y. (2021). A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021, 1–18. https://doi.org/10.1155/2021/8812542

The Integration of Indigenous Languages in Virtual Assistants

Sanele Khakhu
Department of Sociology
University of Johannesburg
Johannesburg, South Africa
Sanelegeraldkahkhu1404@gmail.com

Abstract

Language's centrality in technology, or the expression of Artificial Intelligence (AI) through language, should be evaluated to consider whether the Fourth Industrial Revolution (4IR) technologies in banking can be harnessed for equity or whether they risk perpetuating digital divides. A subset of Human Language Technologies (HLT) and Virtual Assistants (VAs) are conversational applications that use machine learning – the use of computers to draw patterns and inferences from large sets of data to provide human-like interactions to customers or users of the software on mobile apps or websites. While assertions that 4IR and its proponents (AI, machine learning, robotics) will have positive and radical impacts, it is crucial to question whether the Virtual Assistants (VAs) used in financial technologies (Fintech) will enable financial access by bridging language barriers rather than risking current forms of exclusion, disparity, and digital gaps. The aim of this paper is to evaluate whether the expression of language in Fintech can be harnessed to promote equity or worsen the digital divide. This is foregrounded by a Critical Discourse Analysis theory to highlight how language manifests in Fintech. Employing an interpretive qualitative methodology, data were collected through semi-structured interviews and focus groups to unearth rich, contextual insights. Thematic analysis was used to unravel imaginaries surrounding VAs and to extract meaningful patterns. This paper demonstrates that, while great in theory, accommodating the rich diversity of languages spoken in South Africa often runs counter to the pressures of the market in practice. Therefore, this study looks to catalyse policy development that will ensure that HLTs, like VAs, incorporate indigenous languages. This can be achieved by subsidising the development of Free and Open-Source solutions for use in the financial sector. These policies and solutions can be developed in collaboration with local communities, who can provide valuable feedback on their experiences with VAs.

Keywords: Human Language Technology, Language, Virtual Assistant, Financial Inclusion.

Introduction

Language is essential for honouring and preserving cultural histories. It is crucial in shaping relationships, worldviews, and identities which are deeply implicated in the

exercise of power (Alexander, 2005; Hao, 2022). To a large part, due to the history of colonialism and dispossession, English was entrenched as a dominant language, a language of power, which liberation movements in post-apartheid South Africa feared that the new constitution would not consider the language rights of non-English speaking people, such as Tshivenda, isiZulu, Sesotho, amongst others (Alexander, 2005). This article interrogates whether the new technological frontier, virtual assistants (VAs), can perpetuate linguistic power disparities or enable financial inclusion in South Africa. It critically examines the dominance of English, the marginalisation of indigenous languages, and the unequal access to justice in language representation within Human Language Technologies (HLTs). The assumption that "English is enough" is problematic (Mwikisa, 2016; Alexander, 2005), particularly in a multilingual society where many South Africans speak English as a second language. The exclusion of Indigenous languages across sectors such as education, media, and digital banking risks erasing cultural legacies and reinforcing systemic inequalities. The prioritisation of English in VA development, driven by market forces and global technology standards, may reproduce neocolonial dynamics instead of promoting inclusive digital transformation.

A subset of HLT, VAs are conversational applications, also referred to as chatbots that utilise machine learning – the use of computers to draw patterns and inferences from large datasets to provide human-like interactions with customers or users of the software on mobile apps or websites (IBM Watson Advertising, 2022). Some of the well-known VAs include OpenAI's ChatGPT, Amazon's Alexa, Apple's Siri, Google's Assistant, Microsoft's Cortana, among others.

The prioritisation of English in developing VAs may be due to the perception that intelligent people speak English and that English is essential for economic development, especially in a global knowledge economy (Brooker, 2018). It is also partially because English is the most spoken language in the world. The dominance of English and other Western languages on VAs reinforces cultural hegemony. As long as Western languages predominate, the spread of HLT and frameworks will perpetuate the neo-colonial status quo (Pardue and Bin, 2022: 3).

This suggests that a sense of inertia may be present when companies look at the evolution of their digital technologies. Big Technology companies often save money by defaulting to English or other Western languages when training models (Pardue and Bin, 2022: 7). Developing and using languages that are sometimes spoken by a small audience with limited resources is costly. Other startups reinforce this dominance, as they are incentivised to implement the same models or ideas when developing VAs to ensure they cater to a larger piece of the paying market.

Against the claims made about the beneficial and the transformative nature of the Fourth Industrial Revolution (4IR), the researcher aims to evaluate the language question (the dominance of English, the status of indigenous languages, as well as the unequal access and justice in languages other than English) and examines whether the VAs will facilitate financial inclusion through language integration instead of reproducing existing exclusions, inequalities, and digital divides.

The article was guided by the following question: What are the experiences and perceptions of company executives, developers, engineers, architects, and clients of HLTs in the banking sector in a multilingual society?

To answer these questions, the researcher employs a critical and in-depth approach to explore the 4IR from various angles. First, this entails examining the design choices in the development process - specifically, interrogating the processes of banks and AI companies in integrating indigenous languages into 4IR technologies. Second, the researcher explores social impacts, uncovering the effects that the advancements in digital technologies can have on banking. Lastly, the researcher considers the external environment, including markets and policies to unpack their implications for emerging technologies and multilingualism. To consider these implications, the researcher presents the experiences and perceptions related to developing, implementing, and utilising virtual assistants in banking, as well as the limitations of these systems in addressing the challenges of financial inclusion and linguistic inequality in the digital age.

Methodology

This section outlines the exploratory qualitative research methodology employed to unveil the experiences and perceptions of stakeholders (CEOs, CIOs, and Lead Engineers) and bank customers regarding language diversity in VAs. An interpretivist approach facilitated the discovery of new insights and concepts, verified assumptions about language-related challenges, and evaluated the effectiveness of practices and interventions. This holistic inquiry enabled a deep dive into the language phenomenon, unveiling the perceptions and lived experiences of participants.

A non-probability, purposive sampling method was appropriate for this interpretive qualitative approach. This would allow the sampling and selection of participants who were informed about the aims. Four individuals were interviewed: two from technology startups, one from a private bank, and one from a big commercial bank. Additionally, a focus group of five university students and bank clients was conducted to gather perceptions and experiences with banking VAs. This format enabled rich, interactive discussions and highlighted diverse viewpoints (Schearer, 1981; Morgan, 1996; Bhandari, 2020).

Desktop research helped identify which banks and companies were actively deploying and developing digital services and platforms (VAs), particularly regarding the integration of native South African languages (e.g., isiZulu and Tshivenda, amongst others) spoken by the user base (Bhandari, 2020). The focus group provided a broad spectrum of ideas and illuminated differences among participants, enhancing the depth of the findings (Rabiee, 2004: 656).

The scope of the questions was

Scope of Questions	Research Methods
What are the company policies in regard to language integration in the company's virtual assistant?	Semi-structured interviews
Is there algorithmic bias?	Semi-structured interviews and secondary literature on algorithmic bias in Human Language Technology
How are the Human Language Technology Models Built	Semi-structured interviews

Scope of Questions	Research Methods
Use open source? Giving companies control to make changes to suit their model? For Tshivenda, for example?	Semi-structured interviews
What Banks do you use?	Focus Group
What are your thoughts on using AI for Banking	Focus Group
What does your language mean to you?	Focus Group

Figure 1: Research Site

The research was conducted around the city of Johannesburg, from the streets of Sandton to the Southeast of Sandton, Rosebank, and finally to the inner parts of Johannesburg, in Marshalltown.

Understanding HLT Through a Critical Discourse Lens

To understand the importance of traversing the gap in Indigenous language integration within VAs, Fairclough's (1992) Critical Discourse Analysis (CDA) was employed. This theoretical framework provides a lens for examining how languages shape access, financial

inclusion, and power in FinTech. According to Wodak (2013: 303), CDA demystifies ideologies and power via systemic and retroductive investigations of 'semiotic' data (written, spoken, or visual).

Fairclough (1992: 193) distinguishes between power in discourse and power over discourse. Power in discourse can be understood as how power manifests, how it is negotiated, and how it is resisted through language in various interactions. This is reflected in interactions between bank clients and financial institutions that are shaped by linguistic choices and capabilities to enable access to banking services. Moreover, power over discourse highlights the control Tech companies have over the language and data used to train VAs, reflecting power over discourse. In FinTech, this manifests in the prioritisation of English by technology companies, which influences who can access banking services. The lack of available Indigenous language data to train VAs reinforces exclusion, making language a gatekeeper to financial participation (Weidinger et al., 2021: 9). According to Fairclough (2013: 30), language is a social practice constructed by and shaping structures (institutions) and the forces of social institutions in which people function.

Linguistic practices determine access and participation in "power forums" like banks and universities. This depends on one's ability to speak the language used in these forums. Those unable to speak dominant languages (such as English, Afrikaans, and French, among others) face exclusion, while the undervaluing of indigenous languages perpetuates inequality (Alexander, 1997: 84). The historically constructed financial architecture reproduces social domination through semiotic hegemony, rules of language, interaction, and access that favour certain groups (Wodak, 2013: 305). Institutions wield power by controlling discourse, determining which languages are legitimised and which are marginalised.

Furthermore, Foucault (1982: 18) expands this view by describing technologies of power not as machines, but as systems that shape behaviour. VAs, when developed primarily in English, can subtly govern user interactions, limiting access for speakers of isiZulu, Tshivenda, and other Indigenous languages. Thus, CDA reveals how language choices in HLTs reflect broader power structures. The dominance of English in VAs is not neutral; it is a product of institutional decisions that shape and determine who participates in digital financial systems. Through this framework, the article uncovers how linguistic exclusion is embedded in the design and deployment of VAs (Fairclough, 2013: 25).

Literature Review

Language and Virtual Assistants

VAs have become an integral part of many sectors, particularly the banking sector. Leveraging Natural Language Processing (NLP) to enhance user experiences and interactions, these systems are trained to understand and respond to human language with high accuracy. According to VA advocates (Srinivas et al., 2018; Access Partnerships, 2018), customer experience has been revolutionised, providing real-time, personalised interactions through semantic analysis and continuous learning from big data used to improve their capabilities.

For banks or financial service providers, the integration of digital technologies and VAs can ensure the enhancement of customer services, operational efficiency and service delivery. For bank users, service can be provided with customised, around-the-clock assistance for transaction challenges, complaints, and inquiries, reducing the time spent waiting

for business hours or long queues (particularly for the elderly). Regarding operational efficiency, the automation of routine tasks enables VAs to reduce costs by requiring fewer human staff, thereby lowering operational expenses.

In the integration of technology, the use of NLP, Natural Language Understanding (NLU), and Machine Learning (ML) enables human language to be understood, thereby enhancing the user experience. However, a lack of sufficient language data to train VAs has been cited as one of the hindrances to the development and deployment of bilingual or multilingual VAs. Although general-purpose chatbots (e.g., OpenAI's ChatGPT and Google's Bard) or VAs (e.g., Apple's Siri, Amazon Echo, and Google Assistant) are widely available to internet users (Reyes et al., 2019), these HLTs are only developed for a small percentage of linguistic communities (about twenty per cent). According to Alexander (1997: 84), African languages are not valued in the commercial marketplace, as speaking outside of one's relevant 'speech community' is not socially or materially beneficial. Furthermore, environments also act as places where culture is embedded and identities are formed. Nevertheless, if developed equitably, VAs can be used to revitalise indigenous languages instead of reinforcing dominant ones (Meighan, 2021).

Language and Data

If virtual assistants are to become linguistically diverse, developers need to devote sufficient resources to developing models for all languages. Myers-Scotton, as cited in Dyers and Abongdia (2014: 6), argues that there are pseudo-differences in language histories, referring to the perception that people who are most closely linked to speaking one language are more likely to do so than those who are not. This perception is exploited in the political and economic arenas. An example of this is seen in the branching out of Nguni languages into separate languages due to political and historical developments, rather than 'fundamental linguistic' differences. This is further exacerbated by the predominant use of English to train virtual assistants. While isiZulu has some data available, it still fares poorly compared to English, and Tshivenda, in turn, is even more underrepresented than isiZulu. For example, isiZulu has newspapers – *UMAfrica* (The African), *Isolezwe* (The Eye of the Nation), *Ilanga* (The Sun), and *Inkazimulo* (The Glory). Salawu, as cited by Tshabangu and Salawu (2022: 37), discovered that digital journalism practices by African-language newspapers, Alaraye (Nigeria) and Isolezwe (South Africa), have a digital presence. There are television news channels (SABC) that use South African languages and data (material) from these channels can be used to train the chatbots.

Although data may be available to train VAs, Koenecke et al. (2020: 7684) highlight the issues of pronunciation that prevent these systems from functioning effectively. In their study, pronunciation disparities were found in some of the most popular applications of VA (Apple, Amazon, Google, and IBM, amongst others) between White Americans and African Americans. This was due to differences in pronunciation, including rhythm, accent, and pitch (Koenecke et al., 2020: 7685). For language to be fully utilised through technology, a vast amount of textual and audio data from other languages must be used to train VAs.

AI in Banking

Digital technology has been widely adopted in various sectors of society, with the banking sector being one of the fastest-growing users of technology, particularly artificial intelligence, to maintain a competitive edge over other banks. Banking in modern society is a crucial factor that can aid in poverty alleviation in African countries.

According to Access Partnership (2014), a company's ability to exploit and use AI to its fullest potential can help open a door to new possibilities, making AI a potentially "powerful tool for development" and poverty alleviation. In addition, in the financial industry (banks), AI can be used to ensure that the unbanked in Africa get access to banking, providing them with financial inclusion (Access Partnership, 2014). This will showcase the possibilities AI can present to banks, uncovering new ways to access financial services, save money, and obtain insurance.

Access Partnership (2014) revealed that African banking markets are the most exciting as they present untapped avenues and are growing rapidly, making them a "hotbed" for innovation with leading players finding contemporary ways to develop business models to respond to challenges Africa faces (low levels of banking penetration, heavy use of notes, and a 'limited credit bureau). Moreover, research conducted by Srinivas et al. (2018) in the Deloitte banking outlook found that approximately 40% of banking customers opt to use digital methods for making transactions – including making payments, checking balances, and receiving alerts about suspicious activities involving their bank cards.

Digitisation and the development of new digital technologies can help banks grow by tailoring their products to meet their customers' requirements and needs, thereby giving them a competitive edge over their competitors. This suggests that banks can leverage AI analytics and big data to establish effective strategies that assist in cutting costs – automating laborious tasks, improving the customer experience, and potentially acquiring new clients.

Although the banking industry is thriving with expected long-term growth, the proliferating claims made by authors (Srinivas et al., 2018; Access Partnerships, 2014) about the potential technologies in banking may have issues of financial exclusion, imperialism, and extractivism that have not been adequately explored. Although technology companies claim that their products can support multiple languages, they often fail to accommodate all non-Western languages, and those that are supported tend to perform less well than the dominant languages. This is where hegemonic dynamics rooted in capitalist production and distribution shape unequal technology outcomes (Pardue and Bin, 2022: 7).

Although Fintech may be cited as an impact approach to help previously disadvantaged communities access financial services, such as mobile money, microloans, among others (Kampani, 2024; Matias, 2023; Lin and Lee, 2025), various scholars argue that Fintech may be a false messiah as access to banking services may create new challenges that arise due to predatory inclusion. Taylor (2019) provides an examination of a predatory inclusion case involving African Americans in homeownership. Taylor (2019) discovered that homeownership (for individuals who were previously excluded), sponsored by private lenders and the Federal Housing Administration, enabled lenders and sellers to sell subpar and unsafe houses to homebuyers. Instead of protecting homebuyers, Black families were exposed to risky loans and poor housing quality. Consequently, homebuyers found themselves in debt.

Raji (2020) explores another example of predatory inclusion through predatory lending practices. Although under project M-Pesa, financial inclusion for previously unbanked Kenyans increased, with digital loans taken out amounting to fourteen per cent for adults, this tokenistic approach was a "perfected form of digital mining". This allowed the extraction of "small tributes" from every transaction made by the poor, through microloans, money

transfers, grant disbursements, and credit card usage (Raji, 2020; Bateman et al., 2019; Achieng, 2023; Bateman et al., 2019; Kimani, 2020).

On the other hand, in South Africa, the unbanked were assisted by the South African Social Security Agency (SASSA) in collaboration with Net1's subsidiary, Cash Paymaster Services (CPS), which provided financial services to the previously unbanked. A 2020 study by Black SASH uncovered predatory inclusion by Net1. Net1 was given sole power to disburse grants and further advance the financial inclusion agenda. NET1, through its subsidiaries (uManje Money, Smartlife, and Moneylife), made the purchase and access to electricity, insurance, and credit easy to obtain (James et al., 2020). However, control over biometric information and financial data also gave Net1 control over recipients' monies, as debit orders for insurance, for instance, were deducted before grantees had access to their monies.

The above cases highlight that while financial inclusion or other forms of inclusion (such as homeownership) may increase, such gains can serve as a double-edged sword; either enabling access for previously excluded communities or functioning as a predatory instrument that exploits their vulnerability. This suggests that while the inclusion of indigenous languages may increase access to banking services, this may also result in the decrease of the financial health of bank clients, as it did with South Africans, Kenyans, and African Americans. While digital technology may improve banking and financial transactions, various harms must be considered in the deployment and development VAs, especially in the chase for profits over ethical and social good.

Harms

Language and Cultural Imperialism

According to Ngũgĩ wa Thiong'o (1987), the use of African languages maintains the richness of literature, culture, philosophy, and other treasures that African languages possess. The choice of language used in banking, education, and media is, therefore, central to how people define themselves and preserve their identities. However, the definition of African people and who they are is under the continuous control of the West due to its influence and control over Africa's economy, politics, and cultures. This stronghold is perpetuated by the neo-colonial design of AI, as evidenced by the lack of diversity in the digital space, which leads to the underrepresentation of minority languages (Wa Thiong'o, 1987; Kupfer and Muyumba, 2022).

Indigenous societies are challenged to preserve their cultures and represent their people. Africans define themselves in terms of the languages of the colonial era and then imperialist imposition. Wa Thiong'o (1987) stated that we should interrogate these due to the abandonment of indigenous languages in favour of European languages (Kupfer and Muyumba, 2022). Neo-colonial cultural impositions force Africans to distance themselves while identifying with something "furthest" from themselves when using languages other than their own. Alexander (1997: 840) argues that the fatality of losing one's self-esteem, confidence, and dignified self-image is an "important aspect of the syndrome of a colonised mind." The lack of integration of indigenous languages means English, French, and Portuguese retain social control and represent the exertion of neo-colonial power over Africans, their cultures, and languages. This is evident in the linguistic barrier created using "industry-standard language," i.e., English and other Western languages, to program virtual assistants (Dahal et al., 2022). By virtue of owning the hardware, software, and data

used to build VAs, Big Tech corporations in the Global North are digitally exercising a new form of colonialism.

Language and Digital Colonialism

Considering the above challenges, harms, and opportunities related to technological advancements and changes, technology should be scrutinised and viewed as more than just changes and advancements. According to Sutherland (2019: 2), "4IR is not the result of careful historical analysis; it is a flag to rally and a rhetorical device for those trying to create particular economic and commercial futures, hoping to ride waves of Schumpeterian economic disruption caused by 'extreme automation and extreme connectivity.'" In addition, he critiques the slow and inadequate drafting of legal and policy frameworks surrounding cybersecurity, data protection, education, infrastructure, and skills, all of which are essential for the "4IR". In addition, a warning in the context of South Africa is placed on the gap this provides, pushing South Africa to rely on privately owned and international organisations such as the World Economic Forum (WEF), which blurs the distinctions between the impacts and benefits of "4IR" (Sutherland, 2019: 2).

Although Africa has the largest deposits and producers of raw materials for cobalt, dysprosium, terbium, lithium, and graphite (Baskaran, 2023), which are mostly used in 4IR technologies, the economic benefits are disproportionately captured by Big Tech corporations, resulting in low and unevenly distributed benefits for African nations. This extractive process has taken a new form, digitally. According to Kwet (2019), the ownership and control of digital technologies by Big Tech corporations, which enables them to exercise power and control over societies, constitutes a new form of colonialism, known as digital colonialism. This power and control are achieved through the means of computation (software, hardware, and network connectivity) and knowledge (intellectual property and data) owned and controlled mainly by countries in the West, led by the United States of America.

Various authors (Kwet, 2019; Sutherland, 2019; Schelenz & Schopp, 2018) explore the effects of foreign players on African countries, highlighting the extractivist nature of Big Tech corporations. Foreign countries covet data to develop facial recognition systems, credit check systems, surveillance systems, and LLMs leading to heavy investments from their financial sector. While China has a notable global presence, it remains far behind the US in the Global South. As Kwet (2019) observes, many core functions of the digital ecosystem are dominated by transnational corporations.

Digital colonialism is not limited to the infrastructure used and data collected, but also encompasses the ideas and frameworks employed in the development of digital technology or systems. An example is South Africa's attempt to replicate and replace Silicon Valley products and services via Showmax (a Netflix clone), Mr D (akin to Uber Eats), and inDrive (South Africa's "Uber"). One can similarly argue that the frameworks and LLMs designed by US corporations have hegemonic features which may lead to biases in the algorithms used in virtual assistants.

Digital Divide

The continuous development of technologies presents new opportunities for businesses and their clientele. However, the focus should be extended to social inequality in terms of access to resources, participation, and various forms of capital (van Dijk, 2018), as well as

the use of technologies. This must be done while exploring attitudes and motivations in relation to adopting new technologies and digital and information literacy.

Van Dijk (2018) argues that the digital divide should not be limited to physical access (hardware and software), but also encompasses other considerations, such as skills, various ways of using technology, and the "complexity of access," redefined as the conclusive use of technology beyond physical access. In addition, digital inequality should not be viewed as a matter of individual experiences but as categorical disparities between people or communities; workers and executives, males and females, Black people, and White people – whereby certain groups seek to attain and use technology first, harnessing opportunities that are present with the use, thus reinforcing power disparities and positions in relation to other groups or communities.

Furthermore, these categorical differences may lead to the unequal distribution of resources, which in turn results in unequal access to emerging technologies (Dijk, 2018). This suggests that the differences that reinforce power dynamics result in unequal participation, consequently leading to reinforced inequalities and the distribution of resources, which mostly affect marginalised communities or people who have not advanced beyond the initial stage, such as physical access.

Moreover, the literature on the digital divide (van Dijk, 2018; Aksoy et al., 2020; Hechanova & Dioquino Jr., 2004; Nimrod, 2018; Faloye et al., 2022) presents an essential factor that is less commonly discussed, technophobia. Van Dijk (2018) and Faloye et al. (2022) define technophobia as the avoidance of using any form of technology, as well as the perceived distrust users have in the beneficial effects or potential threats. The complexities that users perceive in emerging technologies make them anxious and apprehensive about appropriating them (Aksoy et al., 2020; Hechanova & Dioquino Jr., 2004; Nimrod, 2018; Faloye et al., 2022), thus leading to technology avoidance, resulting in a more widening gap of an existing digital divide, which various literature (Cox et al., 2018; Nyahodza & Higgs, 2017; Robb, 2020) found is deepening in South Africa. This gap can be bridged through localisation of digital technologies, thus putting language at the heart of technology development.

VA Localisation

Dr Ngubane (former Minister of Arts and Culture, Science and Technology), through the Language Plan Task Group (LANGTAG), discovered that the management of language diversity in post-apartheid South Africa was made problematic by the lack of a clear and defined linguistic policy, consequently leading to "undue reliance on the utilisation of non-indigenous languages as the dominant, official languages of the state" – where English and Afrikaans were used as the most dominant languages in socio-economic and political spheres of society (National Language Framework Policy, 2003: 10).

Alexander (1997: 85) posits that to resolve the language issue, institutions and the government must develop policies that ensure the standardisation of community dialects by promoting them through media, technology, and education. Heeding this message can help ensure that a wide diversity of languages and dialects are used in virtual assistants. This should also be applied to schools by teaching science and technology-related subjects in students' native language. Through the phenomenon of localisation, using Free and Open-Source Software, the integration of Indigenous languages in VAs may be possible.

The use of Free and Open-Source Software is a good way to ensure that language communities play an active role in using, changing, and adapting localised software to their benefit. Issues of digital colonialism were developed by policymakers in the early 2000s. South Africa was the first leading supporter of FOSS for use in the public sector on the African continent (Kwet, 2019: 180). Early studies of policy exposed the threat of proprietary software to local interests, with Microsoft at the centre. Proprietary software serves the profits of (primarily foreign) Big Tech corporations at the expense of marginalised communities, which cannot afford to pay for proprietary software. Moreover, proprietary ownership prevents local communities from controlling and customising their own computer-mediated experiences. On the other hand, FOSS grants individuals and communities control over their software, giving them the power to modify it for their own purposes and benefits. It can also be used and shared free of charge, making it more accessible to the poor.

Nichols et al. (2005: 140) argue that the FOSS license encourages the participatory design of software with a community-based end-user interface for localising various languages (of which there are many). Because language is the driver of thought, communication, and cultural identity, the active participation of FOSS systems in VAs for other languages can assist in the preservation and dissemination of language, knowledge, and culture to speakers of different languages (Jimerson and Prud'hammeaux, 2018: 4161; Nichols et al., 2005: 139).

According to Nichols et al. (2005: 145), most software or technologies owned by Big Tech have restrictions due to the unavailability of source code. Consequently, language localisation is hindered. However, through the implementation of the FOSS policy, preference can help fuel the development of open VAs. FOSS is affordable, which enables less affluent organisations and developer communities to make suitable changes to software that will cater to the needs of a given community.

In addition to implementing FOSS in the public sector, the government could adopt a comprehensive, intersectional approach to language and technology policy. This may be achieved through localised technologies – where products are adaptable and customised for specific languages or markets. Osborn and Osborn (2010: 19) propose the term "localisation ecology" to identify the factors affecting localisation and the specific efforts to localise software. The localisation of software depends on sufficient standardisation in terms of dictionaries, terminology, and orthographies. Efforts to develop and enrich standards can benefit from government language policies or institutional programs in healthcare, education, and financial systems, as well as VAs developed in local languages. Educational institutions also play a vital role in the process of localisation ecology, where the installation of localised software or programmes (VAs) on computers in various languages can offer multiple ways of studying and interacting with technology, thus providing people access to various services (Osborn and Osborn, 2010: 125).

For the financial sector, a combination of policies and governmental initiatives that encourage the use of FOSS, localisation, and multilingualism can enable the use of diverse languages in VAs, allowing various language communities to benefit from using them in their own languages. As language exclusion can manifest into digital inequality, creating measures to bridge the financial digital divide is essential. Giving language communities control over how to manage their languages and for what purposes they are used is essential to ensuring that language communities play an active role in maintaining the integrity of their languages and culture, bridging the digital gap, and thus enabling access to financial

services. This will ensure every language and linguistic community is respected, and robust policies are created. Therefore, this review addresses a significant gap in the inclusion and accessibility of Fintech-linked language.

Findings and Discussion

As humanity has entered a new era, with a surge in AI use in various sectors, the use of VAs – a human's new best friend, able to assist with anything, is now a fingertip away. However, the slow pace at which South Africa has absorbed new technologies through its parts, such as language, has added another element to the already existing division, the 'digital divide' (Mallikarjun, 2004). Banks need to consider a few aspects in relation to the development of virtual assistants for the benefit of everyone, not just privileged clients.

Challenges with VA Development

There are various challenges that come with the development of multilingual VAs. Participant 2 disputes the idea that the lack of diverse language use in VAs is an issue, highlighting that "*first, we need to close the digital divide. Get people comfortable with technology, and once they're comfortable with it, it becomes easier to start. So, I don't think it is a language barrier thing, it's a trust thing*."

This is reiterated by Faloye et al. (2022: 4) who state that technophobia arises when users view new technology as a threat due to the complexity, leading to fear, anxiety, and avoidance (Faloye et al., 2022; Aksoy et al., 2020; Hechanova & Dioquino Jr., 2004; Nimrod, 2018). Faloye et al.'s (2022: 8) study also discovered that users with no former employment have high levels of technophobia, while young adults exhibit low- to mid-levels. However, beyond access and anxiety, the researcher argues that language is a crucial factor. Using native languages in digital technologies like VAs can ease apprehension and foster comfort and adoption.

When asked how language would be a challenge or major issue for bank clients when using the bank's virtual assistant, Abby, Participant 1 responded: *"In terms of language, 100% [there are challenges], because if you don't know English, then you are stuck." They explained that the VA are "looking for those trigger words like airtime, and may not understand alternative language like "please recharge for me," even though the intent is the same*.

The above highlights that challenges still arise even within the English language. At the bare minimum, without sufficient training data, the VA will fail. This sentiment is echoed by Participant 2 who emphasises that without sufficient "intelligence", VAs struggle to interpret user requests, leading to access barriers. When asked about challenges in using VAs, focus group participants highlighted challenges with pronunciation. Participant 5 explained, "*Our accents are borrowed. They are not our own, sir. If I want to send money to Boikanyo, and it would send to Boikano, you see that's gonna [going to] be a problem…It's Boikanyo and not Boikano*.

Participant 6 agreed, noting that "*there are certain pronunciation [differences] in our pronunciations that can make using these systems difficult*." These concerns are echoed by Koenecke et al. (2020: 7685) who found in their study that pronunciation variations between white and black communities affect VA, as models are trained on "standard English" often

spoken by White users. The biases in chatbot pronunciation perpetuate racial disparities (Weidinger et al., 2021: 17).

Supporting studies (Keniston, 2003; Mallikarjun, 2004; Warschauer, 2002; Baasanjav, 2014) emphasise that language influences access to the digital ecosystem, an often-overlooked aspect of the digital divide. Therefore, language integration should be considered holistically, including how the structural features of capitalism shape the development and accessibility of VA in banking.

Ownership and the Market as the Dictator

Western imposition has shaped polices and technological frameworks of various countries, including South Africa, often aligning them with neoliberal ideas. According to Heugh, as cited by Phaahla (2015: 185), this is based on linguicism, privileging those fluent in English in systems of power. Consequently, Western languages like English have a better chance of being used to train Large Language Models (LLMs), thereby sidelining South Africa's indigenous languages due to global market pressures.

When asked about the possibilities of integrating isiZulu and Tshivenda, Participant 3 acknowledged the difficulty:

> "We're definitely struggling with other languages. We sort of only focusing on English on for now, just simply because other languages aren't on the scale of what it is. It's super hard, especially for that switch-code in the South African context – where we can switch what language you speak quite a lot. We don't have experience with it that much because our things are for more English products at the moment."

They emphasised that this challenge extends beyond individual companies. Given English's global dominance and extensive digital footprint, startups like Participant 3's, naturally gravitate toward it for VA development. Participant 3 reinforced this by highlighting that English's widespread use and digital footprint make it a default choice, even if it means excluding other languages.

This suggests that, considering all factors, the development of virtual assistants is market-driven, with language choices influenced by commercial viability. However, the researcher argues that this is also due to the dominance of Western ideologies embedded in the design of LLMs. With the West leading the AI race, languages used to train LLMs may give a particular tech giant a competitive edge due to their profitability and digital saturation.

As Africa becomes more digitised, the process warrants critical scrutiny. The growing dominance and influence of US-based Big Tech corporations and other foreign players bring not only their technologies but also their values, models, perspectives, and ideas. These are deeply embedded in the software, hardware, and AI systems sold to African nations, reinforcing Western paradigms and limiting integration of local linguistic and cultural contexts.

When asked about products used for their company's core products, Participant 3 explained that:

> "Our real products is our own LLMs with that system around it but we do heavily use OpenAI and ChatGPT to test our things and even do our small little things in our entire pipeline. But not extensively".

This reflects a hybrid approach that leverages Western frameworks while maintaining some independence. The researcher argues that while it is potentially acceptable to use established models to assess practicality and feasibility, the context in which the products are developed and deployed is crucial. Companies like OpenAI design systems primarily for languages (English, French, Chinese, German) with extensive digital footprints, making their frameworks effective for those markets. However, in South Africa, deploying VAs requires consideration of the country's nine indigenous languages. Tailoring these models to reflect South African's linguistic diversity demands collaboration with language communities and research initiatives such as Masakhane, ensuring VAs are developed by and for South Africans.

When asked about data storage for business operations, Participants 1 and 3 revealed their reliance on Western cloud infrastructure. Participant 2 stated, "*…its [Microsoft] Azure*". They further explained "*there's no services which are free, specifically if you've got an enterprise agreement…whether it's Amazon, or whether it's Microsoft, you have enterprise agreements that you would pay for*."

For their company, Participant 4 stated that:

> "We use [Amazon] AWS. A lot of our stuff is done and stored on there. We try to store all of our data in encrypted forms of data. We really don't have sensitive data. There's no personal data that we use."

These responses underscore Western imposition and dominance in the development of and deployment of VAs. The reliance on paid cloud services like Azure and AWS fills the pockets of Western tech giants and reinforces their control over the technology market in South Africa. This reflects a broader pattern of Western impositions where technological platforms, frameworks and data ecosystems are shaped by foreign interests, limiting local independence and innovation.

Toward a Localised Virtual Assistant

As Africa becomes increasingly exposed to the digital ecosystem through digitisation, the process of digitisation should be scrutinised, given the growing dominance of US-based Big Tech corporations and other foreign players. Additionally, the US-based Big Tech corporations' dominance is accompanied by their values, models, perspectives, and ideas, which are deeply embedded in the technologies imported to African countries – including software, hardware, and artificial intelligence sold to these countries. The participants were asked if an immense amount of data is required for VAs to work efficiently, and their responses were that:

> "The more data you have, the better the engine gets. 'Cause you would filter it if you created a model for a language with 5 people, compared to making a language model that is getting inputs not from 5 people, but from a million people." (Participant 3).

> "Remember, chatbots and AI require intelligence; you need to be feeding it intelligence on a continuous basis, so that it can pre-empt what Sanele's next question is going to be or what his requirements are going to be." (Participant 2).

Although the above responses demonstrate the idea that more and more data are required to have the 'best' language model, a Western-based idea of building LLMs, Participant 4 provides differing views from those presented by Participants 2 and 3 – that language models require astronomical amounts of data to function effectively.

> No, I really don't believe that's the future. I have my own AGI (artificial general intelligence) beliefs and there I can see scale and a lot of data can get you very far but then in terms of product building then definitely not that amount of data.

If Participant 4 is correct, enormous amounts of data are not necessary for chatbots to function effectively. The reliance on Western-based tech companies further illustrates the neocolonial status quo, as foreign tech giants develop core features of the digital ecosystem and extract rent in the form of intellectual property, access to infrastructure, and licenses to use proprietary software (Coleman, 2021: 21; Kwet, 2019: 88).

Moreover, Kwet (2019: 79) states that FOSS gives the users control over how software is used or adapted to benefit them. The researcher argues that Western influences are prevalent in the development of technologies such as virtual assistants. This is highlighted by the systems used to store data, where these cloud computing infrastructures are paid for monthly, filling the pockets of Western tech giants while giving them control over the technology market in South Africa. This echoes the dominance of technology and software, thus preventing, for example, language communities that may use LLMs deployed by tech companies from collaborating with them to adapt the LLMs to meet their ever-changing requirements and needs.

Participant 4 provided a view on the use of FOSS in LLMs, when asked about their perceptions on using FOSS in VAs:

> "I support it a lot. I really do believe that language models can be open source. The good outweighs the bad…and in this case for and just the progress that has been made and sort of the good that can come. There will come products that will help millions of people in ways we've not seen and it's purely because it was open source, and the technology was driven forward miles than if it was just constantly being kept behind walls."

Therefore, to add to the argument against the use of Western ideologies to develop South African LLMs, the use of FOSS and working with people on the ground (language communities) and linguists will assist in handing control over to relevant stakeholders and communities to ensure that the people develop LLMs for the people. This is supported by banking clients who argue that:

> Why are we not included like other languages as well? Because mina [me] I'm a Zulu native, I think it will be easier for me to access the banking app using my own language because, in sports there are sports commentators, there are Afrikaans

> commentators, there are English commentators … there's that inclusiveness for everyone. So why not apply that to banking as well? (Participant 8).

All participants shared the same view that using their native language is essential. Participant 6 explained that "*Venda should also be there…let's say when you include another language, English should also be there or let's say there's an option that says you can change*." Similarly, Participant 9 stated, "Yes, *they just need to accommodate everyone*." The participants' responses highlight the importance of having language diversity in banking applications. Although some of them may opt to use English, it is still crucial to include all languages to accommodate non-English speakers.

Conclusion

In agreement with Fairclough's (1992) critical discourse analysis, stakeholders may perpetuate access issues by prioritising the more profitable English language in their product lines rather than developing technologies that meet the needs of a diverse population. This practice risks excluding financially marginalised excluded (unbanked) from accessing online banking services such as virtual assistants (VAs). CDA reveals how language can both enable and inhibit participation in social institutions, reinforcing its role as a social practice essential for equitable access.

To make the Fourth Industrial Revolution (4IR) inclusive, the integration of indigenous languages must be central to technological development. In South Africa, the digital divide is not only infrastructural, marked by unreliable internet, loadshedding, and high data costs, but also linguistic. The lack of language diversity in digital services disproportionately affects low-income, non-English-speaking communities in townships and rural areas, as highlighted by focus group participants. Policies and practices must be revised to more accurately, appropriately, and respectfully represent Indigenous peoples and their languages, removing barriers to banking access (Farnel et al., 2018: 15).

This article proposes that beyond technological innovation, VAs and the concept of 4IR more broadly should be critically examined through the lens of socioeconomic justice. Indigenous language integration in banking LLMs should be taken seriously. If English continues to predominate, hegemonic dynamics that marginalise African communities will be further perpetuated through the spread of VAs. Using CDA, this article reveals the interplay between HLT, power, and language in banking, challenging the surface meanings of such communication technologies (Wodak, 2013: 305).

Localisation through Free and Open-Source Software (FOSS) offers a pathway to linguistic representation. As illustrated by the findings, the "free" market dictates how and for whom LLMs are developed, favouring those with economic power. The capitalist bias places profit above people. While integrating Indigenous languages into VAs is promising in theory, accommodating the rich diversity of languages spoken in South Africa runs counter to the pressures of the market. This often hinders practical implementation.

To counter these, policies are required to mandate the inclusion of Indigenous languages in HLTs. Government support, such as subsidising FOSS development for the public sector, can help. These efforts should be co-designed with local communities, universities, and linguists, incorporating their feedback to ensure relevance and accessibility. Ultimately,

technology should be viewed through a lens that explores how the creation of a seemingly benign technology, such as virtual assistants, can act as a keyhole to explore and unveil the entire background set of considerations and motivations of why and how certain languages and communities are included in the fruits of digital society, and why others are not.

References

Access Partnership. (2014). Artificial Intelligence for Africa: An Opportunity for Growth, Development, and Democratisation (University of Pretoria). Available at: https://accesspartnership.com/reports/artificial-intelligence-for-africa-an-opportunity-for-growth-development-and-democratisation. Accessed: 20 July 2025.

Achieng, P. G. (2023). The Effect of Foreign Direct Investment on Economic Growth in Kenya. Unpublished Doctoral dissertation. University of Nairobi: Nairobi.

Aksoy, N. C., Alan, A. K., Kabadayi, E. T., & Aksoy, A. (2020). 'Individuals' intention to use sports wearables: the moderating role of technophobia.' *International Journal of Sports Marketing and Sponsorship*, 21(2), 225-245. https://doi.org/10.1108/IJSMS-08-2019-0083

Alexander, N. (2005). *Language, class and power in post-apartheid South Africa. Harold Wolpe Memorial Trust open dialogue event*. Cape Town, 27 October. Available at: https://wolpetrust.org.za/dialogue2005/CT102005alexander_transcript.pdf. Accessed: 20 July 2025.

Baskaran, G. (2023, October 25). *Prospects for U.S Minerals with Africa*. Available at: https://www.csis.org/analysis/prospects-us-minerals-engagement-africa. Accessed: 20 July 2025.

Bateman, M. (2019). South Africa's post-apartheid microcredit experiment: moving from state-enforced to market-enforced exploitation, *Forum for Social Economics*, 48(1): 69-97. https://doi.org/10.1080/07360932.2015.1056202

Bender, E.M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021). March. *On the dangers of stochastic parrots: Can language models be too big?*. In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency (610-623). https://doi.org/10.1145/3442188.3445922

Bhandari, P. (2020). *Data Collection: Definition, Methods and Examples*. Scibbr. Available at: https://www.scribbr.com/methodology/data-collection/. Accessed: 20 May 2024.

Brooker, D. (2018). *English for development: A review of research into the role of English as a driver for socio-economic development*. Cambridge Assessment English. Available at: https://www.cambridgeenglish.org/Images/539681-perspectives-english-for-development.pdf. Accessed: 20 October 2024.

Coleman, M. C. (2021). Leveraging the rhetorical energies of machines: COVID-19, misinformation, and persuasive labor. *Human-Machine Communication*, 3, 11-26. https://doi.org/10.30658/hmc.3.2

Cox, M., Cheng, L., & Forbes, A. (2018). *Bridging the digital divide: a South African perspective*. Paper presented at the Photonic Networks and Devices Available at: https://doi.org/10.1364/NETWORKS.2018.NeW2F.3. Accessed: 26 May 2023.

Dahal, A., Khadka, A., Kharal, B. and Shah, A. (2022). 'Effectiveness of native language for conversational bots'. *Research Square*. Available at: https://doi.org/10.21203/rs.3.rs-2183870/v1

Dyers, C., & Abongdia, J. F. (2014). Ideology, policy and implementation: Comparative perspectives from two African universities. *Stellenbosch Papers in Linguistics*, 43(1), 1-21. https://doi.org/10.5774/43-0-157

Fairclough, N. (1992). 'Discourse and text: linguistic and intertextual analysis within discourse analysis'. *Discourse and Society*, 3(2): 193-217. https://doi.org/10.1177/0957926592003002004

Fairclough, N. (2013). *Language and power*. London: Routledge. https://doi.org/10.4324/9781315838250

Faloye, B. (2022). Sustainable digital pedagogy in language teacher education: Perception of teachers in Ekiti state government colleges. *International Journal of English Language Teaching*, 10(1), 13-22. https://doi.org/10.37745/ijelt.13/vol10.no1pp.13-22

Farnel, S., Koufogiannakis, D., Laroque, S., Bigelow, I., Carr-Wiggin, A., Feisst, D., & Lar-Son, K. (2018). Rethinking Representation. *The International Journal of Information, Diversity, & Inclusion*, 2(3), 9-25. https://doi.org/10.33137/ijidi.v2i3.32190

Filologia, Q. (2006). 'What is Critical Discourse Analysis?'. *Estudis Lingüístics*, 11, 9-34.

Foucault, M. (1982). *Technologies of the self: Lectures at University of Vermoint in October 1982. Foucault*. info. Available at: https://foucault.info/documents/foucault.technologiesOfSelf.en/. Accessed: 20 July 2023.

Geniusee. (2021, September 2023). *Virtual Assistants and Digital Custody for Banking.* https://www.geniusee.com/single-blog/virtual-assistants-and Digital-Custordy-for-Banking/.

Hao, K. (2022, May 10). *Artificial Intelligence: A New Vision of Artificial Intelligence for the People*, Avalable at: https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people/. Accessed: 20 June 2024.

Hechanova, R. M., & Dioquino Jr, M. C. (2004). 'Technophobia and the Filipino Worker.' Philippine *Journal of Psychology*, 37(2), 1-1.

IBM Watson Advertising. (2022). *The Ultimate Guide to Machine-Learning Chatbots and Conversational AI*. Availabe at: https://www.ibm.com/watson-advertising/thought-leadership/machine-learning-chatbot. Accessed 01 January 2023.

IBM. (2023, October 24). 'What is Interactive Voice Response?' Available at: https://www.ibm.com/topics/interactive-voice-response. Accessed: 19 March 2024.

Jimerson, R. and Prud'hommeaux, E. (2018). 'ASR for documenting acutely under-resourced indigenous languages.' *In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Miyakazi, Japan. European Langauge Resources Association (ELRA). Available at: https://aclanthology.org/L18-1657/#. Accessed: 04 April 2022.

Keevy, J., Rajab, R., Arnesen, J., Ngeleza, B., Beukes, C.J., Freeth, R., Akpan, A., Ntuli, S., Laughton, P., Marais, M.A. and Vannini, S. (2022). 'Reclaiming self-sovereignty for all South African citizens data by 2030'. *Johannesburg: JET Education Services and merSETA*. Available at: https://psetcloud.org.za/news/mzansi-citizensdatamatter/mzansi-citizensdatamatter_d2.pdf. Accessed: 30 March 2022.

Kimani, S. M. (2020). *Financing preferences of micro, small and medium enterprises in Kenya: A discrete choice experiment*. Available at: https://su-plus.strathmore.edu/items/ea842434-05b4-4a80-82ef-3a4fe7827a4d. Accessed 24 April 2025.

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the national academy of sciences*, 117(14): 7684-7689. https://doi.org/10.1073/pnas.1915768117

Kupfer, M. and Muyamba, J. (2022). 'Language and Coloniality: Non-Dominant Languages in the Digital Landscape'. *Pollicy*. Available at: https://pollicy.org/wp-content/uploads/2022/08/Languages-Coloniality-Report.pdf. Accessed: 28 May 2022.

Kwet, M. (2019). Digital Colonialism: South Africa's Education Transformation in the Shadow of Silicon Valley. Unpublished PhD thesis. Rhodes University: Grahamstown. https://doi.org/10.2139/ssrn.3496049

Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, 60(4): 3-26. https://doi.org/10.1177/0306396818823172

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *The lancet digital health*, 5(3): e105–e106. https://doi.org/10.1016/S2589-7500(23)00019-5

Meighan, P.J. (2021). "Decolonizing the digital landscape: The role of technology in Indigenous language revitalization. AlterNative: *An International Journal of Indigenous Peoples*, 17(3): 397–405. https://doi.org/10.1177/11771801211037672

Mpunzana, T.D. (2012). A Critical Analysis of Communication and Langauge Practices Prevalent Between ABSA Bank (Empangeni) and their Clients. Unpublished Master's Dissertation. University of Zululand: Richards Bay.

Mthombeni, Z. M., & Ogunnubi, O. (2020). An appraisal of bilingual language policy implementation in South African higher education. *South African Journal of African Languages*, 40(2), 186–197. https://doi.org/10.1080/02572117.2020.1804278

Mwikisa, P. (2016). 'Interviews with Neville Alexander: The power of languages against the language of power.' *African Historical Review*, 48, Available at: https://doi.org/10.1080/17532523.2016.1275299

Myers-Scotton, C. (2006). Natural codeswitching knocks on the laboratory door. *Bilingualism: Language and cognition*, 9(2), 203–212. https://doi.org/10.1017/S1366728906002549

Nichols, D.M., Witten, I.H., Keegan, T.T., Bainbridge, D. and Dewsnip, M. (2005). 'Digital libraries and minority languages'. *New Review of Hypermedia and Multimedia*, 11(2), 139–155. https://doi.org/10.1080/13614560500351071

Nimrod, G. (2018). Technophobia among older Internet users. *Educational Gerontology*, 44(2–3), 148–162. https://doi.org/10.1080/03601277.2018.1428145

ONLIM. (2022, September 09). *The History of Chatbots from ELIZA to ChatGPT*. Available at: https://www.onlim.com/en/the-history-of-chatbots/. Accessed: 20 August 2024.

Osborn, D., & Osborn, D. Z. (2010). *African languages in a digital age: Challenges and opportunities for indigenous language computing.* Cape Town: HSRC Press.

Pardue, L. and Bin, S.A. (2022). *The Other Hegemony: Effects of software development culture on music software, and what we can do about it*. In NIME 2022. PubPub. Available at: https://nime.org/proc/nime22_30/. Accessed: 04 March 2023.

Raji, R. (2020). *Digital financial inclusion in Africa: an analytical assessment of Kenya & Nigeria*. Available at: https://dr.ntu.edu.sg/entities/publication/c4f6f82d-29b2-437c-a86f-a27418c6c600. Accessed: 03 May 2024.

Reyes, R., Garza, D., Garrido, L., De la Cueva, V. and Ramirez, J. (2019). *Methodology for the implementation of virtual assistants for education using Google Dialogflow. In Advances in Soft Computing: 18th Mexican International Conference on Artificial Intelligence*, MICAI 2019, Xalapa, Mexico, October 27–November 2, 2019, Proceedings 18. Springer International Publishing.

Robb, G. (2020). *Bridging the digital divide in South Africa: Case studies in incumbency versus competition*. Cape Town: HSRC Press.

Ruane, E., Birhane, A. and Ventresque, A. (2019). Conversational AI: Social and Ethical Considerations. *Irish Conference on Artificial Intelligence and Cognitive Science*. Available at: https://www.semanticscholar.org/paper/Conversational-AI%3A-Social-and-Ethical-Ruane-Birhane/02246dd3867dbbd8b2c19b9ace4c6168787dd04e. Accessed: 15 May 2024.

Salawu, A. (2019). *Alaroye, Isolezwe and the adoption of digital technologies. In African language digital media and communication*, edited by A. Salawu, 1–8. London: Routledge. https://doi.org/10.4324/9781351120425-1

Schelenz, L. and Schopp, K. (2018). Digitalization in Africa: Interdisciplinary perspectives on technology, development, and justice. *International Journal of Digital Society*, 9(4), 1412–1420. https://doi.org/10.20533/ijds.2040.2570.2018.0175

Srinivas, V., Fromhart, S., Goradia, U. and Wadhwani, R., 2018. 'Banking Outlook: Accelerating the transformation.' Deloitte Center for Financial Services. Deloitte. Available at: https://www.deloitte.com/us/en/insights/industry/financial-services/digital-transformation-in-banking-global-customer-survey.html. Accessed: 03 March 2022.

Sutherland, E. 2020. "The Fourth Industrial Revolution – The Case of South Africa". *Politikon*, 47(2), 233–252. https://doi.org/10.1080/02589346.2019.1696003

Taylor, K.Y., 2019. *Race for profit: How banks and the real estate industry undermined black homeownership*. Chapel Hill, North Carolina: UNC Press Books. https://doi.org/10.5149/northcarolina/9781469653662.001.0001

Tiroshi, O. (2023, August 22). *South African languages: Exploring the languages of South Africa*. Available at: https://www.tomedes.com/translator-hub/south-africa-languages. Accessed: 19 February 2023.

Tshabangu, T., & Salawu, A. (2022). Technology innovation and digital journalism practice by indigenous African-language newspapers: The case of uMthunywa in Zimbabwe. *African Journalism Studies*, 43(2), 37–52. https://doi.org/10.1080/23743670.2022.2049618

Tuli, F. (2010). 'The basis of distinction between qualitative and quantitative research in social science: Reflection on ontological, epistemological and methodological perspectives'. *Ethiopian Journal of Education and Sciences*, 6(1): 98–108. https://doi.org/10.4314/ejesc.v6i1.65384

Van Dijk, J.A. (2017). *Digital divide: Impact of access. The international encyclopedia of media effects*, pp.1–11. https://doi.org/10.1002/9781118783764.wbieme0043

Wa Thiong'o, N. (1987). *Decolonising the Mind*. London: James Currey.

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A. and Kenton, Z. (2021). *Ethical and social risks of harm from language models*. Available at: https://arxiv.org/abs/2112.04359. Accessed: 15 November 2024.

Wodak, R. (2013). *Critical Discourse Analysis – Challenges and Perspectives*. Critical Discourse, eds Wodak, R, Sage, London. https://doi.org/10.4135/9781446286289

Zemčík, M.T. (2019). *A Brief History of Chatbots. DEStech Transactions on Computer Science and Engineering, 10*. https://doi.org/10.12783/dtcse/aicae2019/31439

Zhou, J., Müller, H., Holzinger, A. and Chen, F. (2023). *Ethical ChatGPT: Concerns, challenges, and commandments*. Available at: https://arxiv.org/abs/2305.10646. Accessed: 20 October 2023.

Linguistic Exclusion in South African E-Governance

A Qualitative Study of Xitsonga-Speakers' Access to Digital Public Services

Janine Shout Ramothwa
Department of Political Science
University of South Africa
Pretoria, South Africa
ramothwajanine@gmail.com

Abstract

Digital platforms offer new avenues for public service access, yet indigenous communities like Xitsonga speakers in South Africa face exclusion due to language barriers in e-government and ICT tools. This qualitative study investigates how integrating Xitsonga into digital spaces can enhance access to essential services, revealing that the current systems marginalise indigenous language users, limiting their government engagement. Framed through Decolonial Theory, Language Rights, and Ubuntu philosophy, the findings reveal that the current systems marginalise indigenous language users, limiting their government engagement. The findings underscore the need for inclusive digital policies and platform design to bridge the digital divide. The study contributes to debates on language equity, digital access, and equitable service delivery, offering actionable recommendations for policymakers.

Keywords: Indigenous Languages, Digital Governance, Xitsonga, Digital Divide, E-Government.

Introduction

The Fourth Industrial Revolution (4IR) has fundamentally transformed public service delivery, shifting many essential government services from in-person consultations to digital platforms. This global trend towards civic technology, digital tools designed to promote citizen engagement, transparency, and participation, is reflected in South Africa's growing digital landscape. The proliferation of apps for tracking services, filing applications, and reporting corruption aims to make governance more efficient and accessible (Ohamadike, 2025; Lesame & Malatji, 2022). Nationally, digital participation is significant with over 50.8 million internet users (78.9% of the population) and 124 million cellular connections, indicating widespread, multi-device usage (HelloYesMarketing, 2025).

However, this digital forward leap has a profound exclusionary flaw: the infrastructure is monolingual. In a nation constitutionally committed to twelve official languages (including sign language), South Africa's civic technology is overwhelmingly delivered in only one language, English. This turns a tool of empowerment into a mechanism of marginalisation, creating what can be termed a governance failure and a technical oversight (Ohamadike, 2025). The consequence is not merely a technological limitation but a daily, real-life barrier for millions, restricting their access to essential services like job applications, healthcare information, and social grants.

While this is a systemic issue affecting many indigenous languages, its impact is acutely felt by speakers of Xitsonga, a language spoken by 4.7% of the population. For Xitsonga-speaking communities, the promise of digital governance is a hollow one. They are caught between a constitutional guarantee of linguistic equity (Constitution, 1996, Section 6) and the reality of English-only digital platforms. This disconnect is exacerbated by the failure of key legislation, such as the Use of Official Languages Act (2012), which mandates multilingual government services but remains unenforced in the digital realm. As Artificial Intelligence (AI) continues to shape economic development, this linguistic exclusion risks locking out not only individuals, but the entire language communities from the future digital economy (Primus, 2025).

Therefore, this study moves from the broad problem of digital language exclusion to a focused investigation of the Xitsonga-speaking experience. This research investigates the specific barriers Xitsonga speakers face in accessing South African e-government services and explores their proposed solutions for more equitable digital governance. Furthermore, it analyses the critical policy gaps, specifically the failure of the Use of Official Languages Act (2012) that enable and perpetuate this linguistic marginalisation.

Research Problem

Despite constitutional guarantees of linguistic equity, Xitsonga speakers face systematic exclusion from essential e-government services due to the English-only design of digital platforms. This language barrier actively disenfranchises citizens, violating their rights, and eroding trust in the state, while existing policy, such as the Use of Official Languages Act (2012), fails to provide an enforceable remedy in the digital sphere. This study investigates this gap between policy and practice by focusing on the lived experiences of the Xitsonga speech community.

Significance of the Study

This research is significant for four key reasons:

1. **Policy Relevance**: It generates empirical evidence that can compel policymakers to amend and proactively enforce the Use of Official Languages Act within digital services, directly advancing the implementation of 'Batho Pele' (People First) principles.
2. **Theoretical Contribution**: It extends decolonial and linguistic human rights theories by applying them to the digital realm, introducing the concept of "digital coloniality" to explain how colonial-era linguistic hierarchies are reproduced in state architecture.
3. **Social Justice**: It centres the voices of a marginalised speech community, advocating

for linguistic justice by framing access to government services in one's mother tongue as a fundamental human right, consistent with the Ubuntu ethic of collective dignity.

4. **Practical Impact**: The community-proposed solutions identified in this study provide a practical, ground-truth blueprint for government departments and platform developers to create more accessible and inclusive e-governance tools.

Research Questions

This study is guided by the following primary research question and its sub-questions:

1. How do Xitsonga speakers experience and navigate linguistic exclusion on South African e-government platforms?
2. What are the specific barriers and socio-emotional impacts of English-only digital platforms on Xitsonga speakers?
3. What coping strategies, such as intergenerational dependency or translation tools, do they employ?
4. What community-driven solutions do participants propose for more inclusive digital governance?

Methodology

This research is located within the qualitative research paradigm to facilitate an in-depth exploration of participants' experiences, emotions, and perspectives within a digital exclusion context (Rapanyane, 2021; Tenny et al., 2022). This approach prioritises detailed, context-rich insights and allows for an open-ended investigation while minimising researcher bias by cantering participant voices (Mwita, 2022).

Research Design and Site Selection

An exploratory case study design was employed to examine Xitsonga-speaking communities' experiences with e-government platforms (Yin, 2003). This design was selected for its ability to analyse real-world phenomena within their natural context, making it ideal for investigating how language barriers impact digital service accessibility. The study was conducted in Chiawelo, Soweto, Johannesburg, in the Gauteng Province. This area was purposively selected for its significant Xitsonga-speaking population, providing a rich context to explore the intersection of linguistic identity and access to essential digital public services.

Data Collection and Sources

Data were collected from primary and secondary sources to ensure a comprehensive analysis. Primary data, semi-structured interviews were conducted with five Xitsonga-speaking participants. Interviews focused on their experiences with three key national e-government platforms:

1. The SASSA online portal for social grants.
2. The Department of Home Affairs e-Services (e.g., smart ID applications).
3. The central government services portal (www.gov.za).

For secondary data, a critical policy analysis was conducted using the Constitution of the Republic of South Africa, 1996 (Chapter 1, Section 6) and the Use of Official Languages Act, No. 12 of 2012. These documents, accessed via the South African Government Gazette (www.gov.za), were analysed alongside primary findings to assess the gap between legislative intent and real-world implementation.

Participant Sampling and Demographics

A non-probability, purposive sampling technique was used to recruit participants who were Xitsonga speakers and had experience with the target e-government services (Vehovar et al., 2016). The initial target was 8 - 10 participants. However, recruitment challenges, including participant reluctance due to concerns about scams and data privacy, resulted in a final sample of five individuals. While small, this sample size is consistent with in-depth qualitative inquiry where the priority is thematic saturation over statistical generalisability (Braun & Clarke, 2021). Thematic saturation was confirmed as the final interviews yielded redundant information, and no new themes emerged.

The participant cohort was diverse, ensuring a wide spectrum of experiences. Demographic details are summarised in the table below:

Table 1: Participant Demographics

Pseudonym	Age	Gender	Occupation	Education Level	Key E-Gov Services Used
P1	38	Female	Nurse	University Degree	Home Affairs, Gov.za
P2	27	Male	Business Owner	Matric (Grade 12)	Gov.za, SASSA
P3	59	Male	Domestic Worker	Primary School	SASSA
P4	56	Male	Officer	Diploma	Home Affairs, Gov.za
P5	26	Female	Unemployed	Matric (Grade 12)	SASSA, Gov.za

Data Analysis and Ethical Considerations

Thematic analysis was used to interpret the data, following a structured six-phase process (Braun & Clarke, 2022; Dawadi, 2020). The process involved: 1) familiarisation with the data through transcription and translation from Xitsonga to English while preserving participants' original phrasing; 2) generating initial codes; 3) searching for themes; 4) reviewing potential themes; 5) defining and naming themes; and 6) producing the report. Analysis was conducted using Microsoft Word for clarity and organisation.

The study adhered to strict ethical standards in compliance with the Protection of Personal Information (POPI) Act. All participants provided verbal consent, and their anonymity was protected through the use of pseudonyms (P1–P5) throughout the data collection, analysis, and reporting processes.

Theoretical Framework

This study employs a multi-layered theoretical approach to examine the exclusion of Xitsonga speakers from digital spaces. Decolonial Theory is placed as the foundational lens, used to diagnose the root of the problem. This perspective is then given moral force by the Language Rights Framework, which articulates the exclusion as a rights violation and is guided toward solutions by the Ubuntu Philosophy, which provides an African-centred ethical blueprint for inclusion. Together, these frameworks move from critiquing the colonial past to demanding justice in the present, and envisioning an equitable future.

Decolonial Theory: Diagnosing the Roots of Digital Exclusion

Decolonial Theory argues that political independence does not dismantle the deep-seated structures of power, knowledge, and established by colonialism, a condition termed "coloniality" (Quijano, 2000; Ndlovu-Gatsheni, 2013). Ndlovu-Gatsheni's (2013) concept of "coloniality of power" provides a foundational lens for understanding how these colonial hierarchies persist in post-apartheid South Africa's digital spaces. In this context, this manifests through what Kroeze (2024) termed "digital coloniality", the reproduction of colonial linguistic hierarchies through technological platforms that privilege European languages while marginalising indigenous knowledge systems. This theory allows for the English-domination of e-government to be framed not as a simple oversight, but as an active continuation of colonial oppression and epistemicide (the killing of knowledge systems).

Language Rights Framework: Naming the Injustice

Drawn from the field of linguistic human rights, this framework posits that language rights are fundamental human rights (Skutnabb-Kangas, 2017; Phillipson, 1992). Skutnabb-Kangas's (2017) linguistic human rights theory positions access to government services in indigenous languages as a fundamental right, not just a convenience. It distinguishes between tolerance-oriented rights (allowing indigenous languages to exist) and promotion-oriented rights (actively supporting their development and use). This study argues that by failing to integrate Xitsonga into digital platforms, the South African government operates at a mere tolerance level, thereby violating its duty to uphold promotion-oriented rights. This framework transforms participants' frustration into a legitimate claim for justice and holds the state accountable to its constitutional and international obligations.

Ubuntu Philosophy: Charting an Ethical Path Forward

'Ubuntu' is an African humanist philosophy centred on communal interdependence and mutual care, often summarised by the phrase "*umuntu ngumuntu ngabantu*" (a person is a person because of other people) (Shutte, 2009; Ewuoso & Hall, 2019). Ubuntu (Ramose, 2015), encapsulated in the Xitsonga phrase "*munhu I munhu hi van'wana*", emphasises collective responsibility and interconnectedness over individuality. It therefore challenges individualistic approaches to digital literacy that place the burden on users to adopt English-dominated systems. Instead, Ubuntu posits that digital inclusion is a collective responsibility requiring systemic changes to accommodate all community members. This philosophy provides the ethical grounds to critique the shame and dependency caused by exclusion and to validate the community-driven solutions proposed by participants, framing inclusion as a necessary condition for collective dignity. The interplay between the three core theories guiding this analysis - Decolonial Theory, Language Rights, and

Ubuntu philosophy – is illustrated in Figure 1, demonstrating their complementary roles in diagnosing the problem, articulating the injustice, and envisioning solutions.

An Integrated Analytical Chain

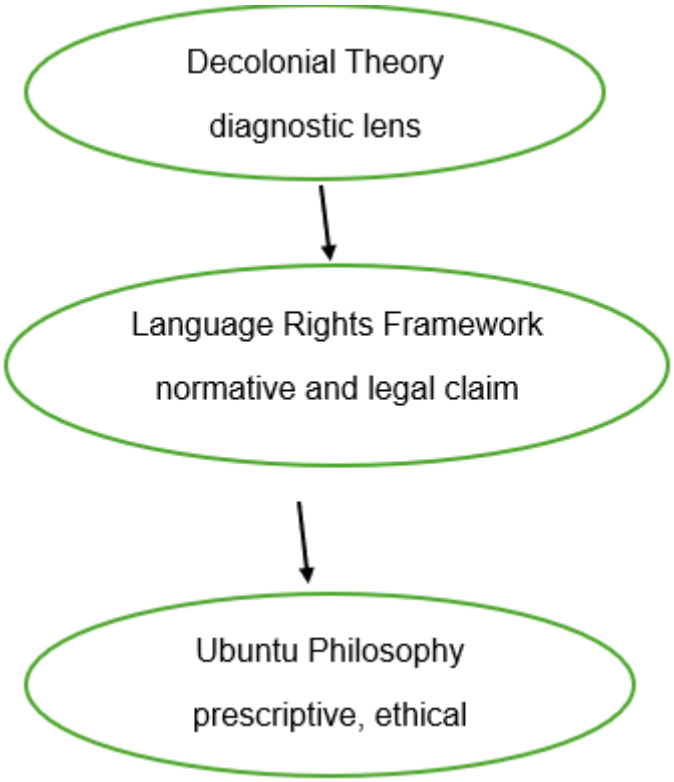

Figure 1: The cohesive analytical chain of decolonial theory, language rights, and Ubuntu philosophy.

Used in concert, this integrated framework allows for an analysis that is historically deep, legally sound, and ethically guided toward an inclusive future.

Literature Review

The digitisation of government services or e-governance promises greater efficiency, transparency, and accessibility for citizens. In a multilingual society like South Africa, this promise is intrinsically linked to the question of language. This literature review critically examines the existing scholarship on the intersection of indigenous languages, digital technology, and public service delivery in South Africa. It argues that the historical marginalisation of indigenous languages, established during the colonial and apartheid eras, is being systematically reproduced in the digital realm. By synthesising research on historical language policy, the current state of indigenous languages and the problematics of digital technology, this review will demonstrate a critical research gap: a lack of focused inquiry into the lived experiences of specific linguistic communities, such as Xitsonga speakers, as they navigate an increasingly digital state. This gap underscores the urgent need for the present study.

The Historical Roots of Linguistic Marginalisation

South Africa's democratic Constitution of 1996 promised a decisive break from apartheid's linguistic imperialism, explicitly granting eleven languages "parity of esteem" and equitable treatment. This was a radical departure from a past where colonial powers systematically undermined African languages by imposing their own languages, a practice that devalued indigenous knowledge systems and instilled lasting linguistic hierarchies (Azubuike & Aji, 2021; Adedokun & Zulu, 2022). However, the transition to genuine multilingualism has been incomplete. Scholars argue that the post-colonial landscape is often characterised by a "lack of political will," resulting in language regulations that are forgiving of former colonial languages and poorly implemented for indigenous ones (Adedokun & Zulu, 2022). The consequence, as Balfour (2019) contends, is that millions of citizens cannot fully participate in national life, this includes the judiciary, economy, and public administration when it is conducted in an "imported" language they do not comprehend. This failure to realise constitutional promises has created a persistent gap between policy and practice, setting the stage for old inequities to be reproduced in new digital contexts.

The Digital Shift: A New Frontier for Old Exclusions

The rapid digitisation of public services, part of the broader Fourth Industrial Revolution (4IR), has transformed citizen-state interaction. Yet, this shift has become a critical new site where historical linguistic inequalities are being intensified. While technology is often hailed as a potential solution for language revitalisation, offering new tools for communication and access (Olaitan, Issah, & Wayi, 2021; Ndzendze, 2020), the reality for many indigenous language speakers is quite the opposite.

Research indicates that speakers of South Africa's indigenous languages often feel their languages are excluded from formal and technological communication (Gumbi, 2019). This sentiment is compounded by significant structural barriers: a severe lack of digital resources for indigenous languages, unique structural complexities that require specialised approaches, and occasional orthographic problems (SADiLaR, 2021; Malatji & Lesame, 2019). Consequently, the development of South Africa's e-government landscape has largely followed colonial-era patterns. As Malatji and Lesame (2019) note, the use of African languages in ICT remains minimal, with European languages continuing to dominate digital platforms. This creates a form of "digital neo-colonialism" where technologies that should empower instead serve as vectors for globalised content and hegemonic languages, potentially undermining local traditions and knowledge systems (Sundani, 2023). The result is that digital government risks becoming, by default, English-dominated government, effectively locking out a significant portion of the population.

Bridging the Gap: Centring the Xitsonga-Speaking Experience

The existing literature successfully establishes the broad nexus between colonial history, marginalised indigenous languages, and the digital divide. Scholars have convincingly documented the macro-level challenges, including a lack of political will (Adedokun & Zulu, 2022) and the technical barriers hindering the development of digital resources for indigenous languages (SADiLaR, 2021; Malatji & Lesame, 2019). However, a critical gap remains. The research tends to focus on systemic overviews and policy analysis, often failing to drill down into the lived, qualitative experiences of specific linguistic communities. For instance, while Malatji and Lesame (2019) highlight the minimal use

of African languages in ICT, their work does not detail the socio-emotional impact or the practical coping strategies citizens employ. Similarly, broader discussions of the digital divide (e.g., Ndzendze, 2020) and linguistic marginalisation (e.g., Gumbi, 2019; Makgopa, 2022) often remain conceptual, lacking deep, empirical investigation into how these forces converge in the daily lives of individuals trying to access essential services.

The Xitsonga-speaking community, constituting 4.7% of the South African population, epitomises this gap (Madlome, 2022). The general challenges outlined in the literature converge acutely for this group, yet their specific plight remains largely invisible in academic research. There is a scarcity of studies that centre the voices of Xitsonga speakers themselves to understand how they navigate, interpret, and resist linguistic exclusion on e-government platforms. Therefore, this study addresses this gap by asking: How do Xitsonga speakers experience and navigate linguistic exclusion on South African e-government platforms, and what do their experiences reveal about the failures of current language policy in the digital age? By answering these questions, this research moves beyond describing the systemic problem to amplifying the grounded, human experiences that define it, providing a nuanced evidence base for more inclusive digital governance.

In summary, the literature reveals a clear and troubling trajectory: the linguistic hierarchies of South Africa's past are not being dismantled in its digital future but are being digitally reinforced. The constitutional promise of multilingualism remains unfulfilled, and the advent of e-government has created a new, high-stakes domain of exclusion. While scholars have effectively mapped the historical roots and broad structural dimensions of this problem, the human experience at its core remains underexplored. The gap in the literature is not that the problem is unknown, but that the voices of those navigating this digital marginalisation, like Xitsonga speakers, are largely absent. It is this critical gap that the present research aims to fill, using qualitative inquiry to ground the macro-level problem of digital language policy in the micro-level realities of citizen experience. By doing so, it seeks to contribute a vital, human-centred perspective to the discourse on digital inclusion and linguistic justice in South Africa.

Findings

Theme 1: Language Barriers as Exclusion

The study confirms that participants are willing to partake in digital platforms; however, they have identified language barriers as a primary obstacle towards their participation in digital spaces. English-dominated e-government platforms allow Xitsonga speakers to be excluded from accessing essential services. Several participants have pointed out how minor misunderstandings of a word error led to significant misunderstandings, undermining their ability to complete applications accurately:

as Participant 1 explained, "One English word misunderstood changes the answer and makes everything incorrect".

This aligns with Hietaranta's (2014: 3-4) cognitive theory of translation errors: misinterpretations force the brain to accept plausible yet flawed solutions; therefore, users overlook inconsistencies in their translated content. For Xitsonga speakers, this creates a double vulnerability, a linguistic marginalisation compounded by cognitive bias where even a minor error distorts access to jobs or grants, reinforcing digital exclusion.

Critically, translation is not neutral nor purely a mechanical act, it is deeply embedded within the context of culture, words carry connotations, values, and historical baggage that differ across linguistic boundaries (Mahammadjonova, 2025: 2). When a Xitsonga speaker tries to translate the e-government information to their home language, they apply their language's cultural context to understand the information according to the culture. For example, Participant 3 struggles with "misinterpreted words changing the meaning of the paragraph", which shows how the technical jargon, such as public tender, lacks a conceptual framework in Xitsonga. These oversights compound exclusion, and minor misunderstandings lead to "missed opportunities" (Participant 3) for grants and jobs.

This exclusion is aggravated when Xitsonga speakers become cautious of machine-generated translations, therefore withdraw entirely, an issue explored in theme 2.

Theme 2: Distrust in Translation

This study found that translation tools exist to convert between Xitsonga and English, yet Xitsonga speakers remain deeply sceptical about automated translation systems. Their distrust stems from recognising that these tools lack the cultural competency of native speakers who understand the language's contextual nuances and cultural roots. As Participant 1 explained, "Some people try to translate languages without proper understanding of the indigenous language, but if someone who knows Xitsonga deeply were to translate it, I would believe their work".

This scepticism reflects what Helm et al. (2025: 8-9) termed "techno-linguistic bias", the phenomenon whereby technology is built to work better for English while marginalising Xitsonga. This distrust is rooted in participants' direct experiences with translation tools. Both participants 1 and 3 expressed frustration with automated systems, with one stating, "it is (machine translation) wrong half the time" and emphasising the need for human linguists to preserve meaning.

Translation accuracy concerns

The table below illustrates critical translation errors discovered when testing Google Translate with Xitsonga terms commonly used in government contexts.

Table 1: Tested on Google Translate (June 2024). Results may vary with updates. (Author's own table, 2025)

Xitsonga Word	Google Translation	Correct Translation	Risk in E-Gov
Nhluvuko	Hambana	Development	Distorts policy information
Mbvhoma wa mathlarhi ya masocha	Drone ya masocha	Military drone	Fuel distrust in government tech.

These errors exemplify Giunchiglia et al. (2023: 2) described as "lexical untranslability", the inherent difficulty of translating certain concepts across linguistic boundaries. The mistranslation of "nhluvuko" (development) as "Hambana" (differences) demonstrates how these errors used daily could mislead Xitsonga speakers when accessing policy documents, potentially resulting in a misunderstanding of government programs or services.

Theme 3: Intergenerational Dependency

The study confirmed that Xitsonga speakers depend on younger tech-literate family members to access monolingual digital spaces and e-government websites because of the language barrier, which shows the digital literacy gap. Depending on younger family members created shame and disappointment, according to participants, as it is believed that the elders should be the ones helping the younger generation, according to the culture, this brought power imbalances in Xitsonga-speaking households.

Participant 3: "I had to call someone to translate and understand some of the words on the website since they are more educated than I am."

Participant 4: "I had to ask someone to apply for me since I did not go to school. The challenge is understanding what they mean in English."

The participants highlighted the emotional strain of relying on others to access the internet for government services and resources meant for all South African citizens. The need to call for assistance in accessing their government underscores a perceived loss of dignity and autonomy. Thus, those without formal education face compound exclusion.

Theme 4: Community-Proposed Solutions

The study confirms that Xitsonga speakers acknowledge the challenges they face when accessing e-government websites and apps; thus, participants have proposed community-driven and technological solutions to bridge the language gap in e-government platforms, emphasising accessibility and cultural relevance.

Participant 5: "They (the government) should add videos that explain the process in Xitsonga."

Participant 3: "Let us record enquiries in Xitsonga and get responses in Xitsonga."

These solutions align with the oral traditions that prioritises auditory learning. Oladele and McCall (2024: 6-8) state that auditory learning would allow Xitsonga speakers to advance their comprehension and information retention to advance their understanding of the content presented, further, this would allow Xitsonga speakers to create a stronger mental association and solidify their understanding of the material.

Participant 2: "The government should create offices with Xitsonga helpers for free application and remove the need for data to access their websites and apps."

Discussion

The main objective of this study was to examine the exclusion of Xitsonga speakers from e-government platforms due to a language barrier. The study highlights factors beyond individual citizens' control, such as English-language dominance, resulting in dependency and distrust. The findings of this research are consistent with prior studies conducted by Ranchordas (2021), Saeed and Masters (2021), and Mlambo and Matfunjwa (2024), which also demonstrate the exclusivity of English-dominated ICT platforms.

Coloniality of Power and Epistemic Violence in Digital Governance

These findings expose the gap in South African language policy (2012), an advocate for multilingualism, yet fails to enforce multilingualism in digital governance, reinforcing exclusion, what Participant 5 called "undermining Xitsonga". This mirrors the broader critiques of decolonisation of digital spaces, as Eurocentricity continues to dominate spaces of pluralistic societies (Langmia & Sani, 2025: 2).

The constitution recognition of indigenous languages (section 6, 1996) marked a significant rupture from apartheid's linguistic imperialism. However, Participant 1's testimony stating "I have never come across a Xitsonga government website. Everything is in English" reveals a promise that only exists in the state of becoming, a deferred reality that mirrors what Ndlovu-Gatsheni (2013: 18) noted as "coloniality of power" in post-colonial democracies. The absence of a language spoken by 4.7% of the South African population from government services exposes neoliberal democracy as a replication of apartheid-era marginalisation.

The exclusion of Xitsonga from e-government platforms is a bureaucratic failure and an epistemic violence against Xitsonga speakers (Brunner, 2023). This results in the erosion of trust in the state, reducing Xitsonga speakers to second-class digital citizens instead of first-class digital citizens due to the government's failure to include Xitsonga in their digital government service providers (Schous & Hjelholt, 2018; Magro, 2012).

Policy Failure and the Performance of Democracy

The Use of Official Language Act was enacted to monitor and regulate the use of official languages by the national government, while maintaining inclusivity in government information and services. However, according to the testimonies of the participants, they access information in English; therefore, Xitsonga is excluded, resulting in Xitsonga speakers' oppression.

Section 4(2) of the Official Languages Act (2012) mandates bilingual government services; however, it fails to recognise digital platforms as sites for service delivery. The Act's silence on digital spaces marks what Ding (2020: 24) termed "performative democracy", where good governance policies are portrayed, yet their objectives are not achieved. The Act's (2012: 4) permission for departments to select languages based on practicality creates a critical loophole that institutionalises colonial language preferences. This discretionary clause enables departments to justify English-only platforms under the guise of operational efficiency, perpetuating linguistic hierarchies despite constitutional promises.

Therefore, the absence of indigenous languages in digital technology exemplifies Ndlovu-Gatsheni's (2013: 18) 'coloniality of power' whereby the neoliberal democracy only hides the colonial legacies. This is explained by Brunner (2023) as an act of epistemic violence whereby a missed opportunity because of a language barrier violates Xitsonga speakers' rights to access to jobs, as stipulated by the Employment Equity Act No. 55 (1998).

Lessons from Global Counterparts

The challenges Xitsonga speakers face reflect the global patterns of digital linguistic exclusion. However, global states have successfully implemented interventions demonstrating that

multilingual e-governance is achievable. Kenya's Huduma Kenya Portals offering services in both Swahili and English (eCitizen, 2025), and India's UMANG App, which incorporates 23 languages shows how policy mandates can reshape colonial language hierarchies.

Table 2 compares how South Africa, Kenya, India, and New Zealand have adopted interventions that enable indigenous communities to be included in their e-government services.

Table 2: Comparative Analysis of Digital Language Inclusion Policies of South Africa, Kenya, India, and New Zealand. (Authors own table, 2025)

State	Policy	Key provision	Implementation	Lesson for South Africa
South Africa	Use of Official Language Act (2012).	Bilingual services (unenforced digitally)	English/Afrikaans dominance	Amend the Act to include digital mandates
India	Digital India Act (2015).	Multilingual government service platform	UMANG app offers 23 languages through AI	Invest in Xitsonga NLP tools with human oversight
Kenya	Digital Master Plan (2022).	Bilingual interface for assessing government services	Hudama Kenya portals in Swahili and English	Amend the Use of Official Languages Act to acknowledge digital platforms
New Zealand	Māori Language Act (2016, No.17).	Government agencies are required to provide key services in te reo Māori	Co-design with Iwi (tribal councils	Partner with Xitsonga linguists for translations

Unlike Kenya and India, the South African policy gap reflects Skutnabb-Kangas's (2017) linguistic human rights theory framework of tolerance rights, which only enables Xitsonga to exist as an indigenous language, yet fails to adopt models that enable Xitsonga to be used in digital spaces. Participant 5's demand for 'video tutorials in Xitsonga' mirrors Kenya's community-centric approach, suggesting a co-design as a framework to achieve promotion-oriented rights. Additionally, Participant 1's call for 'someone who knows Xitsonga deeply to translate' mirrors New Zealand's implementation strategy of working with Iwi (Tribal council) to ensure accuracy.

Ubuntu and the Path to Co-Designed Inclusion

The dependency described in Theme 3 clashes with Ubuntu's ethos of communal support where collective care fosters personal and societal growth (Ewuosa & Hall, 2019: 5). However, while Ubuntu idealises reciprocity, the government's failure to provide Xitsonga language services forces speakers into a one-sided reliance, framing this reliance as a burden rather than mutual aid.

Participants articulated a desire for meaningful inclusion, proposing solutions such as government-employed assistants, the need for video inquiries, and easy access tutorials to

reclaim their role as digital citizens. This call for assistance and the need for unrestricted access towards government resources should be the guiding principles of the Batho Pele framework. The government fails to rationally put the people first regarding ensuring inclusivity in South Africa. The solutions urge the government to act on the Batho Pele principles.

Recommendations

Based on the findings and discussion, the following recommendations are proposed:

1. **Legislative Reform**: The government's "efforts" to ensure inclusivity are to be seen at the marginalised communities' levels by revising the Use of Official Language Act to update the act to acknowledge digital platforms as part of service delivery. The Use of Official Languages Act amendment should mandate Xitsonga in high-impact e-government platforms such as SASSA, Labour, and Home Affairs.
2. **Co-Design with Communities**: All government tenders or contracts to design ICT for government departments should mandate the incorporation of Xitsonga as one of the indigenous languages, co-designed with native Xitsonga speakers to ensure cultural roots are preserved.
3. **Implement Multi-Modal Solutions**: The government should invest towards indigenous languages. To enhance communication, there is a need for video enquiries and video tutorials that explain a step-by-step process in Xitsonga. The existing government libraries in communities can be utilised by dedicating a section of the library towards "digital language hubs" with Xitsonga mediators, as Participants 5 and 2 suggested.

Conclusion

The study's findings revealed how the exclusion of Xitsonga from e-governance platforms perpetuates colonial hierarchies under the appearance of neoliberal democracy, violating the constitutional promise of linguistic equity. Participants' experience about requiring assistance to apply for a job and their distrust of machine translation exposes how the language barrier re-emerges from apartheid-era marginalisation. Despite the constitutional guarantees, the state's failure to enforce the Use of Official Languages Act in digital spaces highlights epistemic violence.

However, Xitsonga speakers demand co-designed solutions, such as video tutorials and community hubs that offer a blueprint for decolonial digital inclusion. The findings underscore Ndlovu-Gatsheni's (2013) 'coloniality of power', where indigenous language erasure in e-governance mirrors broader patterns of neoliberal exclusion. This necessitates reimagining digital citizenship through Ubuntu's ethic of collective dignity (*Munhu I Munhu hi Van'wana*). Until Xitsonga speakers can access e-governments without shame or mediation, South Africa's democracy remains an unfulfilled promise.

Strengths, Limitations and Future Research

The study has several strengths; the first strength lies in its ability to adopt a comprehensive search strategy, encompassing Xitsonga participants' voices, academic databases, acts, and the constitutions of South Africa. Secondly, the researcher is a Xitsonga speaker;

therefore, all the raw data were accurately translated and maintained the participants' voice and emotions.

Nevertheless, the study contains limitations. Firstly, the small sample size (n=5) limits generalisability, though this was appropriate for the exploratory qualitative design. The difficulty recruiting participants due to scam concerns reflects broader trust issues with digital technologies that warrant a separate investigation. Secondly, the research was conducted within the urban Xitsonga community in Gauteng, rural experiences may differ. Future research should explore the intersection of class and education. Thirdly, the research did not include perspectives from government officials or platform developers, limiting understanding of implementation challenges from the supply side. Future studies should incorporate multi-stakeholder perspectives.

Reference list

Abiero, D., & Kipkoech, D. A. (2023, September 27). *Kenya's Digital Deserts - Centre for Intellectual Property and Information Technology Law. Centre for Intellectual Property and Information Technology Law*. Available at:https://cipit.org/kenyas-digital-deserts/#:~:text=Crafting%20local%20content%20in%20diverse%20languages%20serves,thereby%20enhancing%20the%20accessibility%20of%20digital%20resources.&text=By%20leveraging%20successful%20global%20initiatives%2C%20tailoring%20solutions,vulnerable%20communities%20across%20Kenya%20and%20Sub%2DSaharan%20Africa. Accessed: May 11, 2025.

Adedokun, T. A., & Zulu, S. P. (2022). 'Towards digital inclusion in South Africa: The role of public libraries and the way forward'. *Interdisciplinary Journal of Economics and Business Law*,11(4), 127-154.

Ahn, E., & Kang, H. (2018). 'Introduction to systematic review and meta-analysis'. *Korean Journal of Anesthesiology*, 71(2), 103-112. https://doi.org/10.4097/kjae.2018.71.2.103

Anstey, G. (2022, November 18). *SA's indigenous languages must all be fully developed and featured in digital resources, a UP professor says*. Universities South Africa. Available at: https://usaf.ac.za/sas-indigenous-languages-must-all-be-fully-developed-and-featured-in-digital-resources-a-up-professor-says/. Accessed: May 23, 2025.

Asrini, H. W. (2019, October). Language semiotics of campaign banners and billboards of legislative candidates in Malang. In *Proceedings of the 6th International Conference on Community Development (ICCD 2019)* (pp. 217-220). Atlantis Press. https://doi.org/10.2991/iccd-19.2019.58

Azubuike, F. C., & Aji, T. C. (2021). 'Role of information and communication technologies (ICTs) in repackaging indigenous knowledge: A 21st century perspective'. *Library Philosophy and Practice* (e-journal), 1-14. Retrieved November 18, 2025, from https://digitalcommons.unl.edu/libphilprac/5226

Balfour, R. (2019). *The impact of the fourth industrial revolution on indigenous languages as priorities for teaching and learning in higher education in South Africa*. Available at: https://news.nwu.ac.za/sites/news.nwu.ac.za/files/files/Robert.Balfour/Address%202019_10_18.UNISA.Impact-of-4IR-on-indigenous-languages.doc. Accessed: May 23, 2025.

Braun, V., & Clarke, V. (2021). 'To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales'. *Qualitative Research in Sport, Exercise and Health*, 13(2), 201-216. https://doi.org/10.1080/2159676X.2019.1704846

Brunner, C. (2023). *Epistemic violence. In Virtual Encyclopedia on Post- and Decolonial Peace and Conflict Studies*, Available at: https://postcolonial-peace.org/encyclopedia/epistemic-violence Accessed: June 1, 2025.

Chandel, A., Lee, R., Newton, C. E. L., & van Gameren, D. (2025). 'Advancing Spatial Justice: A Framework Integrating Spatial Justice and Spatial Capital'. In R. Rocco, C. Newton, & J. Gonçalves (Eds.). *A Manifesto for the Just City* (Vol. 4, pp. 38–51). TU Delft OPEN Publishing.

Dawadi, S. (2020). 'Thematic analysis approach: A step-by-step guide for ELT research practitioners'. *Journal of NELTA*, 25(1–2), 62–71. https://doi.org/10.3126/nelta.v25i1-2.49731.

Ding, I. (2020). Performative Governance. *World Politics*, 72(4), 525–556. https://doi.org/10.1017/S0043887120000131

eCitizen. (2025). *eCitizen.go.ke - Government of Kenya services simplified*. Available at: www.ecitizen.go.ke. https://accounts.ecitizen.go.ke/en. Accessed: May 28, 2025.Ewuoso, C., & Hall, S. (2019). 'Core aspects of Ubuntu: A systematic review'. *South African Journal of Bioethics and Law*, 12(2), 93–103. https://doi.org/10.7196/SAJBL.2019.v12i2.679

Giunchiglia, F., Bella, G., Nair, N. C., Chi, Y., & Xu, H. (2023). 'Representing interlingual meaning in lexical databases'. *Artificial Intelligence Review*, 56(10), 11053–11069. https://doi.org/10.1007/s10462-023-10562-9

Gumbi, P. (2019). 'IsiZulu is an indigenous African language, and technology in the basic education sector in KwaZulu-Natal'. *South African Journal of African Languages*, 39(2), 204–210. https://doi.org/10.1080/02572117.2019.1618026

Helm, P., Bella, G., Koch, G., & Giunchiglia, F. (2023). *Diversity and language technology: How techno-linguistic bias can cause epistemic injustice*. https://doi.org/10.1007/s10676-023-09742-6

Hietaranta, P. (2014). Understanding misunderstanding: Towards a cognitive view of misunderstandings in translation. In P. Penda, T. Bijelić, & Ž. Babić (Eds.), *Going against the grain: Essays on literary and language studies* (pp. 215–223). Faculty of Philology, University of Banja Luka.

Laskowska, A., Skowron, Ł., Kozłowski, E., & Laskowski, J. (2025). 'Bridging the digital divide: The role of proxy users in online purchasing and social sustainability'. *Sustainability*, 17(3), 112. https://doi.org/10.3390/su17030112

Lesame, C., & Malatji, E. (2022). 'The use of South African languages by youth on social media: The case of Limpopo Province'. *Communicare: Journal for Communication Studies in Africa*, 38(1), 76–95. https://doi.org/10.36615/jcsa.v38i1.1545Journals.co.za+2UJ Press Journals+2UJ Press Journals+2

Lirola, M. M. (2016). 'Multimodal analysis of a sample of political posters in Ireland during and after the Celtic Tiger'. *Revista Signos*, 49(91), 245-267. https://doi.org/10.4067/S0718-09342016000200005

Madlome, S. K. (2022). Revitalization of Tsonga identity through indigenous language media outlets in Southern Africa. In R. Tunhuma & T. Chimuanya (Eds.), *Indigenous language for development communication in the Global South* (pp. 157–174). IGI Global. https://doi.org/10.5040/9781666997071.ch-9

Magro, M. J. (2012). 'A review of social media use in e-government'. *Administrative Sciences*, 2(2), 148-161. https://doi.org/10.3390/admsci2020148

Makananise, F. O. (2023). 'Reimagining South African political campaigns through indigenous language posters in the 4IR: A political communication perspective'. *Communicare: Journal for Communication Studies in Africa*, 42(1), 52–63. https://doi.org/10.36615/jcsa.v42i1.1404

Makgopa, M. (2022). *Implications of the fourth industrial revolution (4IR) on the development of indigenous languages of South Africa: Challenges and opportunities*. https://doi.org/10.4018/978-1-7998-7492-8.ch009

Malatji, E., & Lesame. C. (2019). 'The use of South African languages by youth on social media: The case of Limpopo Province'. *Communicare*, 38(1), 76–95. https://doi.org/10.36615/jcsa.v38i1.1545

Manda, M., & Backhouse. J. (2018). Inclusive digital transformation in South Africa: an institutional perspective. *In Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, Galway, Ireland, April 2018. https://doi.org/10.1145/3209415.3209486

Masenya, M. T. (2022). Digital preservation of indigenous knowledge in South African rural communities. In R. Tshifhumulo & T. J. Makhanikhe (Eds.), *Handbook of research on protecting and managing global indigenous knowledge systems* (pp. 317–336). Hershey, PA: IGI Global. https://doi.org/10.4018/978-1-7998-7492-8.ch017

Mlambo, R., & Matfunjwa, M. (2024). 'The use of technology to preserve the indigenous languages of South Africa'. *Literator*, 45(1), a2007. https://doi.org/10.4102/lit.v45i1.2007

Muhammadjonova, M. (2025). 'Intercultural communication in translation: The consequences of incorrect translations'. *Modern Science and Research*, 4(4), 1382-1390.

Ndlovu-Gatsheni, S. J. (2013). 'Coloniality of power in postcolonial Africa: Myths of decolonization. Dakar, Senegal'. *Council for the Development of Social Science Research in Africa* (CODESRIA).

Ndzendze, B. (2020). *The fourth industrial revolution and its implications for African languages*. Retrieved from: https://www.usaf.ac.za/wp-content/uploads/2020/12/Keynote-Address-by-Dr-BhasoNdzendze.pdf. Accessed: May 29, 2025.

Ngcobo, S., & Mvuyana, B. Y. C. (2022). 'Communication and governance in a linguistically diverse human settlement in South Africa'. *Journal of Local Government Research and Innovation*, 10(2), 83–95. https://doi.org/10.4102/jolgri.v10i2.83Jolgri

Ohamadike, N. (2025, July 6). Digital technology must speak African languages. *The Mail & Guardian*. Available at: https://mg.co.za/thought-leader/opinion/2025-07-06-digital-technology-must-speak-african-languages/. Accessed: July 13, 2025.

Oladele, O. K., & McCall, A. (2024). 'Auditory learning: The importance of verbal instruction and discussion'. *Journal of Educational Psychology*, 116(2), 45–60. https://doi.org/10.1037/edu0000456

Olaitan, O. O., Issah, M., & Wayi, N. (2021). 'A framework to test South Africa's readiness for the fourth industrial revolution'. *South African Journal of Information Management*, 23(1), 1–10. https://doi.org/10.4102/sajim.v23i1.1284

Oparaugo, B. (2021). 'Media and politics: Political communication in the digital age'. *International Journal of Multidisciplinary Research and Explorer*, 1(2), 14-23. https://doi.org/10.2139/ssrn.3780554

Pan South African Language Board (PanSALB). (2022). *The survival of African languages in a digital era*. Retrieved from https://www.pansalb.org/the-survival-of-african-languages-in-a-digital-era/. Accessed June 1, 2025.

Penney, J. (2017). 'Social media and citizen participation in "official" and "unofficial" electoral promotion: A structural analysis of the 2016 Bernie Sanders digital campaign'. *Journal of Communication*, 67(3), 402-423. https://doi.org/10.1111/jcom.12300

Phillipson, R. (1992). *Linguistic imperialism*. Oxford: Oxford university press.

Primus, M. (2025, February 12). Opinion: Breaking the digital divide: Why Africa's future depends on embracing its languages in AI. University of Johannesburg News. Available at: https://news.uj.ac.za/news/opinion-breaking-the-digital-divide-why-africas-future-depends-on-embracing-its-languages-in-ai/. Accessed: May 11, 2025.

Quijano, A. (2000). 'Coloniality of power and Eurocentrism in Latin America'. *International Sociology*, 15(2), 215-232. https://doi.org/10.1177/0268580900015002005

Ramose, M. (2015). Ecology through ubuntu. In R. Meinhold (Ed.), *Environmental values: Emerging from cultures and religions of the ASEAN region* (pp. 69–76). Bangkok, Thailand: Konrad Adenauer Stiftung & Assumption University.

Ranchordás, S. (2021). Connected but still excluded? Digital exclusion beyond internet access. In M. Hildebrandt & K. O'Hara (Eds.), *The Cambridge handbook of life sciences, information technology and human rights* (pp. 244–258). Cambridge, UK: *Cambridge University Press*. https://doi.org/10.1017/9781108775038.020

Saeed, S. A., & Masters, R. M. (2021). 'Disparities in health care and the digital divide'. *Current Psychiatry Reports*, 23(9), 61. https://doi.org/10.1007/s11920-021-01274-4

Shutte, A. (2009). Ubuntu as the African ethical vision. In M. F. Murove (Ed.), African ethics: *An anthology for comparative and applied ethics* (pp. 85–99). Scottsville, South Africa: *University of KwaZulu-Natal Press*.

Skutnabb-Kangas, T. (2017). Language rights. In O. García, A. M. Y. Lin, & S. May (Eds.), *The handbook of bilingual and multilingual education* (pp. 185–202). Malden, MA: Wiley Blackwell. https://doi.org/10.1002/9781118533406.ch11

South African Centre for Digital Language Resources (SADiLaR). (2021). *2nd workshop on resources for African indigenous languages*. Retrieved from https://www.sadilar.org/index.php/en/2-general/286-rail-2021. Accessed: June 1, 2025.

South Africa. (2022). *Employment Equity Act 55 of 1998*. South African Government. Available at: www.gov.za. https://www.gov.za/documents/employment-equity-act. Accessed: May 29, 2025.

Sundani, Ndivhuwo. (2023). 'South African Indigenous Languages and Digital Technologies: Access, Promotion and Preservation'. *International Journal of Social Science Research and Review*, 6, 41–51. https://doi.org/10.47814/ijssrr.v6i8.1385

UMANG. (2019). *One app, many government services*. Available at: https://web.umang.gov.in/landing/. Access date: May 28, 2025.

Use of the Official Languages Act. (2012). *Government Gazette REPUBLIC OF SOUTH AFRICA*. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/35742gon8010.pdf. Accessed: May 28. 2025.

Zia, S. (2025). 'Digital Colonialism: Reimagining Power, Identity, and Resistance by Decolonizing AI'. *Yayasan Drestanta Pelita Indonesia*, 129-147.