

# The use of technology to preserve indigenous languages of South Africa

**Authors:**

Respect Mlambo<sup>1</sup>   
Muzi Matfunjwa<sup>1</sup> 

**Affiliations:**

<sup>1</sup>South African Centre for Digital Language Resources, Faculty of Humanities, North-West University, Potchefstroom, South Africa

**Corresponding author:**

Respect Mlambo,  
Respect.Mlambo@nwu.ac.za

**Dates:**

Received: 06 June 2023  
Accepted: 15 Nov. 2023  
Published: 22 Feb. 2024

**How to cite this article:**

Mlambo, R. & Matfunjwa, M., 2024, 'The use of technology to preserve indigenous languages of South Africa', *Literator* 45(1), a2007. <https://doi.org/10.4102/lit.v45i1.2007>

**Copyright:**

© 2024. The Authors.  
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Indigenous languages in South Africa must be preserved to ensure that they do not lose their identity and become extinct. The four indigenous languages with the fewest speakers among South Africa's 12 official languages are: Xitsonga, Siswati, Tshivenda and isiNdebele. The preservation of these languages in South Africa has been a long-standing challenge because of various social and economic factors. With the advancement of technology, opportunities have arisen to preserve and promote the use of these languages. Therefore, this study explores various technological strategies that can be used to preserve the South African indigenous languages. These languages can be preserved by making them widely accessible to users through various strategies such as localisation of daily used technology, translation through crowdsourcing, digitisation and archiving. 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. This study demonstrates the significance of technology in preserving indigenous languages and promoting their use around the world.

**Contribution:** This study fills the practical gap in the use of technology 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. Therefore, the study provides practical technological strategies that need to be implemented to preserve the indigenous minority languages. The insight of this study into the use of technology to preserve South African languages fits well within the scope of *Literator*, which is to publish studies in linguistics and literature with a special focus on South African languages. This publication will bring solutions to how minority languages could be preserved in the context of South Africa.

**Keywords:** indigenous languages; South Africa; technology; technological strategies; preserve; digital learning tools; localisation; digitalisation.

## Introduction

South Africa is a multilingual country with 12 official languages namely, English, Afrikaans, Sepedi, Setswana, Sesotho, isiZulu, isiXhosa, Tshivenda, Xitsonga, Siswati, isiNdebele and the South African Sign language. These languages are supposed to enjoy the same status of officiality. However, practically, they do not receive equal attention in development as some languages are more resourced than others. Despite that, the South African Constitution (1996) section 6(5) gave the Pan South African Language Board a mandate to 'promote and create conditions for the development of all official languages' and all other typically used non-official languages in South African communities. Indigenous languages face limited digital accessibility because of resource scarcity; they, however, must be technologically preserved so that they can be used in various disciplines without losing their identity. Mercuri (2012) posits that language is perceived as a fundamental component of one's cultural identity and its loss is equivalent to the collapse of an entire culture. UNESCO (2023) also affirms that:

[W]hen an indigenous language is lost, not only does the knowledge accumulated by the community of its speakers fade away, but also the world's cultural and biological diversity is jeopardised. (n.p.)

Despite the importance of language in society, many of the world's languages have become extinct, whereas others are endangered to varying degrees (Etim 2016). Therefore, the preservation of languages is an imperative undertaking, and all stakeholders must create useful strategies to make successful efforts.

In the context of South Africa, Tshivenda, Xitsonga, Siswati and isiNdebele are indigenous languages that have fewer language speakers and language resources (Luvhengo 2012). Statistics

**Read online:**

Scan this QR code with your smart phone or mobile device to read online.

South Africa's 2022 data show that Xitsonga has 4.7% language speakers, followed by Siswati 2.8%, Tshivenda 2.5% and isiNdebele 1.7% ([https://census.statssa.gov.za/assets/documents/2022/Census\\_2022\\_SG\\_Presentation\\_10102023.pdf](https://census.statssa.gov.za/assets/documents/2022/Census_2022_SG_Presentation_10102023.pdf)). The preservation of these languages digitally in South Africa has been a long-standing challenge because of various social and economic factors. With the advancement of technology, opportunities have arisen to preserve and promote the use of these languages. In South Africa, employing technology to address indigenous language preservation challenges has the potential to support economic development, cultural diversity and social cohesion of these languages (Maja 2007). This study focusses on four South African indigenous languages, namely, Tshivenda, Xitsonga, Siswati and isiNdebele. The aim is to explore various technological strategies that can be used to preserve these languages. This study is systematically divided into five sections. In section 2 related works are discussed. Section 3 describes the methodology and section 4 deals with the analysis and discussion of strategies. Lastly, a conclusion is drawn in section 5.

## Related work

Osborn (2006) noted that there is minimal use of African languages in information and communication technologies (ICTs) such as computer applications and the Internet, even though these languages are vital in ways of generating and expressing useful knowledge. The author conceded that if African languages are utilised in other fields, their inclusion in technology becomes vitally important because the reliance on European languages such as English and French to transmit contemporary knowledge and information has been detrimental for societies and people who cannot understand these foreign languages. Osborn (2006) also observed that the implementation of African languages in ICTs is hampered by several factors such as insufficient resources to incorporate African languages in ICT, non-existence of standardised orthographies for indigenous languages with fewer speakers, and use of specialised alphabets in many languages that need specialised fonts. The lack of collaboration between ICT developers and African linguists has also contributed to the minimal existence of these languages in ICTs. Despite these challenges, the author argued that information must be made available in African languages through localisation of the ICTs such as computer software and websites.

Riza (2008) studied indigenous languages in Indonesia and determined that there are languages that are facing extinction on a huge and quick scale. He predicted that languages and cultures of communities with very few speakers might not survive beyond the end of the current century. This problem is because of linguistic and cultural assimilation with the majority group, migration to cities and a lack of social and economic support from the government. Riza identified three important tasks that need to be done to preserve an endangered indigenous language of Indonesia. These tasks are computational linguistics techniques that allow a multidimensional view of language resources, increasing the

orientation of regional research centres towards resource creation, and using statistical or empirical models of language, particularly if the language is near extinction. The author also observed that there is a growing understanding of the relevance of corpus resources in Indonesia, which has assisted in the development of resources that can preserve endangered languages.

Galla (2009) described areas in which technology plays a significant role in language and culture revitalisation. The study discussed efforts made by indigenous communities to preserve, maintain and revitalise their languages using technology. An overview was provided about the significance of language preservation for indigenous communities and the challenges that languages face because of globalisation and cultural assimilation. The author explained that languages can be preserved using technology through creation of language learning applications, online language resources and use of social media and other digital platforms to connect language speakers. The preservation and revitalisation of indigenous languages can also be done by the wax cylinder recording of digital audio recordings, email to chat, video recording of interactive audio videoconferencing and surfing the Internet to play interactive computer games. Some of the challenges and limitations of using technology to preserve languages, such as the need for resources, expertise and the need to balance the use of technology with more traditional language learning methods were also highlighted.

Osborn (2010) posited that the unavailability of ICTs in indigenous languages decreases the chances of creating and having content in local languages online. This affects the culture and information that could be made available to scholars interested in studying them. The author proposed that ICTs should be localised through translating and culturally adapting applications and software graphical user interfaces to indigenous languages. Localisation can also be achieved by developing online content in various languages and translating the content from a dominant language to other target languages. The author also elucidated that localisation will assist many African language speakers to easily access ICTs and information in their local languages. The target group for localisation would be Africans based in Africa, those who live abroad as well as non-Africans. However, there are several challenges that delay the use of ICT in African languages such as expenses associated with translating materials into these languages as well as the fact that the African languages are very diverse. Negative attitude from educated African language users is another issue that contributes to the limited use of languages in ICT. Moreover, African languages are marginalised by educational policies in several countries, resulting in a negative effect on their use in ICTs.

Olaiya (2014) examined the role of libraries in language preservation and development and noted that libraries can play a significant role in language preservation and development, especially for indigenous and endangered languages. The author discussed the importance of languages

and the challenges they face in the modern world. The challenges are language loss, language shift or even language death because many languages are still not properly documented. To address these challenges, Olaifa explored the role of libraries in language preservation by highlighting the various ways in which libraries can support language documentation, revitalisation and promotion. These include the gathering and preservation of language materials, language learning resource provision and language event and activity hosting. The study also discussed some of the issues that libraries face in terms of language preservation, such as a lack of resources and expertise, as well as the need for collaboration with language communities. Olaifa contended that libraries may overcome these problems and effectively preserve languages by having a community-centred approach and engaging closely with language speakers and other stakeholders.

Mirza and Sundaram (2016) investigated how harnessing and leveraging collective intelligence techniques are used to support the preservation and learning of endangered languages. Mirza and Sundaram (2016) defined collective intelligence as a group's ability to perceive and solve challenges that cannot be solved individually. The authors argued that using teamwork, cooperation, coordination and cognition within communities can assist in overcoming some of the issues faced by language revitalisation projects, such as a lack of resources and expertise to preserve endangered languages. A variety of ways in which collective intelligence techniques can be employed in language preservation initiatives were outlined. These included the creation of online platforms and crowdsourcing tools to help in language documentation and revitalisation. The utilisation of online communities and peer-to-peer learning networks to enhance language acquisition was also viewed as an advantage of collective intelligence techniques. Some of the challenges and limitations that might arise from collective intelligence approaches in language preservation were noted. These challenges include the requirement for participants' trust and cooperation as well as the significance of assuring that these approaches are respectful and appropriate for different cultures.

Martín-Mor (2017) explored the use of technology in preserving endangered languages. The author focussed on the minority languages of Sardinia, an Italian island with linguistic diversity. He contended that Sardinian languages are in a state of diglossia because official national language policies do not appear to be capable of reversing the severe situation with these languages, leaving their preservation mostly to individual devotion. Martín-Mor outlined three levels of technological measures that may be utilised to preserve Sardinian languages as endangered languages: translation, localisation and the establishment of language tools and resources. In this study, it was emphasised that it is vital to involve communities in the preservation of languages as each technological measure requires linguistic competencies and technical skills from language speakers.

The conclusion was that language preservation initiatives must be culturally sensitive and take into consideration each community's specific demands.

Moodley (2020) explored the significance of localising online computer software interfaces to include Setswana for increased accessibility and inclusivity for South African teachers. The study found that teachers' experiences with dual English-Setswana interface educational software significantly impact their attitudes towards African languages in ICT. The author also noted that despite teachers' satisfaction with the Setswana software, they expressed concerns about the non-translated words and neglect of dialect accommodation. Teachers also felt that the software could be improved by including more Setswana cultural content and providing more support for teachers in using the software in the classroom. The study concluded that the localisation of online computer software into indigenous African languages is essential in ensuring that all South African teachers have access to the ICT they need to teach effectively.

Sundani (2023) examined the relationship between South African indigenous languages and digital technologies, highlighting the challenges and opportunities in access, promotion and preservation. The author posited that digital technologies can significantly aid in the promotion and preservation of South African indigenous languages through their multimedia capabilities, storage capacity and communication tools. However, limited access to digital technologies supporting these languages negatively impacts their promotion and preservation. The access to digital technologies for indigenous languages in South Africa is hindered by barriers such as lack of expertise, collaboration, equitable digital services and clear advocacy. The author recommended that the South African government, the Pan South African Language Board and other stakeholders, including language experts, researchers and ICT companies, should implement digital technology strategies to prevent these barriers to effective access, promotion and preservation of South African indigenous languages.

The literature consulted demonstrates that the use of technology to preserve indigenous South African languages has not been widely studied to the best knowledge of the authors, especially Xitsonga, Tshivenda, Siswati and isiNdebele, which are languages with fewer speakers in South Africa. The study fills this gap as it explores the use of technology to preserve these indigenous languages. It describes various technological strategies, which could be employed practically to preserve these languages.

## Methodology

This is qualitative research that looks at how technology can be used to preserve indigenous languages of South Africa with fewer speakers. This approach seeks to offer in-depth insights and knowledge of real-world situations (Moser & Korstjens 2017). It requires researchers to employ a systematic

and rigorous approach to gather, interpret and analyse non-numerical data. This study utilised systematic self-observation (SSO) as a valuable data collection method. This method involves individuals systematically recording their observations and experiences, often through written or digital means, about a specific phenomenon of interest (Lumma & Weger 2021; Rydberg 2023). It emerges from an intention to observe, describe and clarify the patterns, events and phenomena of everyday life (Mick, Spiller & & Baglioni 2012). This study employed SSO to determine technological strategies and describe how they can be implemented for promoting and preserving indigenous languages in South Africa with fewer speakers and digital resources, akin to their global application.

## Discussion of strategies

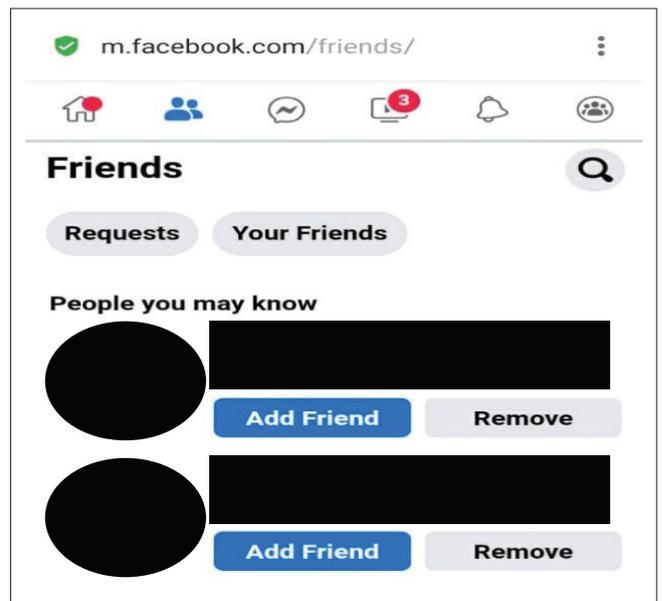
Technologies can be used to preserve indigenous and endangered languages (Martín-Mor 2017). For South African indigenous languages, the technological strategies: localisation of daily used technology, translation through crowdsourcing, digitisation and archiving, digital learning tools, machine translation (MT) and creating online dictionaries that can preserve these languages are discussed in the next subsections.

### Localisation of daily used technologies and/or application software

The term localisation can be explained as 'a communicative, technological, textual and cognitive process by which interactive digital texts are modified for use by audiences around the world than those originally targeted' (Jiménez-Crespo 2013:1). Sandrini (2008) categorised localisation into software application and websites. Both categories of localisation involve linguistic and cultural adaptation of a specific product for a particular group (Esselink 2000). Sandrini (2008) emphasised that the purpose of localisation is to ensure that individuals from specific areas and settings can easily and effectively use software and websites in their home language.

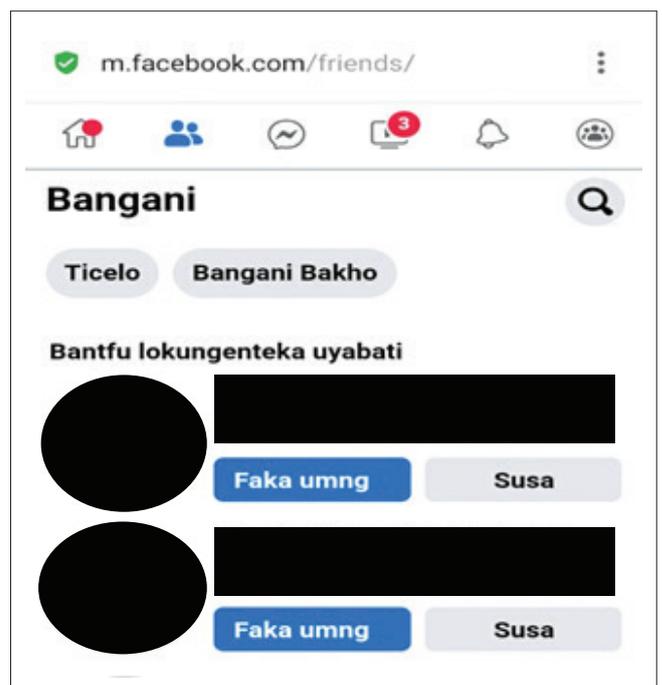
In the context of South Africa, popular web browsers, digital tools and application software interfaces such as Google, Mozilla, Facebook, Twitter, Slack and Skype can be localised into South African indigenous languages to promote the preservation and use of these languages. This can be done by linguists and speakers of the languages to ensure that the interfaces of these applications are in their indigenous languages. For instance, the localisation of the Facebook interface from English to Siswati has been done as shown in Figure 1a and Figure 1b.

From Figure 1a, we observe that the Facebook interface is in English, and in Figure 1b, the Facebook interface has been localised into Siswati. In Figure 1a, the English words 'Friends', 'Requests', 'Your friends' and 'People you may know' were localised in Figure 1b to *Bangani*, *Ticelo*, *Bangani bakho* and *Bantfu lokungenteka uyabati*, respectively.



Source: Matfunjwa, M., 2023, *Friends*, viewed n.d., from <https://m.facebook.com/friends/>

FIGURE 1a: English Facebook interface.



Source: Matfunjwa, M., 2023, *Friends*, viewed n.d., from <https://m.facebook.com/friends/>

FIGURE 1b: Siswati Facebook interface.

A similar approach can be applied to Tshivenda, Xitsonga and isiNdebele, respectively. Motsa (2023) quoted Msibi who stated that:

Facebook has become a very powerful tool for preserving languages. Most of the youth use this social media platform and if you want to conserve knowledge, you need to do it through the youth. If you remove the language from the youth, you have eradicated it. (n.p.)

Therefore, the localisation of social media applications such as Facebook is a necessity for all indigenous languages. The

usefulness of the localisation of web applications is that it makes digitised documents accessible to many sociolinguistic groups worldwide (Jiménez-Crespo 2013). Martín-Mor (2017) also affirmed that the purpose of localisation is to ensure that certain products are available in the languages of minority groups. Hence, in the case of South Africa, localisation ensures that technologies and application software are available in indigenous languages and enhance their use and preservation.

## Translation through crowdsourcing

South African indigenous languages can be preserved using translation through crowdsourcing. Brabham (2013:xix) defined crowdsourcing as 'an online, distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organisational goals'. Kalinin and Savchenko (2014) highlighted that crowdsourcing is a quick approach to obtaining high-quality translations in which community members and native speakers of the languages participate willingly. In South Africa, speakers of Xitsonga, Tshivenda, Siswati and isiNdebele can voluntarily and collaboratively translate articles on the website, specifically on Wikipedia which permits the translation of English articles to a specific target language. Wikipedia data shows that there is a huge gap in the number of articles written in indigenous languages of South Africa with fewer speakers as shown in Table 1.

From Table 1, we observe that the South African indigenous languages with fewer speakers have the least number of articles found on Wikipedia: Tshivenda has 808, Xitsonga 741 and Siswati 656. Disappointingly, isiNdebele has no articles available. These statistics show that more work and effort must be geared towards the translation of articles into Tshivenda, Xitsonga, Siswati and isiNdebele, with the latter language in dire need. Therefore, translation through crowdsourcing will improve the number of articles and preserve these indigenous languages.

The online translation exercise will make information accessible in all disciplines and provide corpora that are useful in developing language resources and tools for under-resourced indigenous languages. The translation must also

**TABLE 1:** Number of articles in South African official languages in Wikipedia.

South African official languages	Wikipedia language code	Number of articles available
English	En	6 657 289
Afrikaans	Af	108 020
isiZulu	Zu	10 940
isiXhosa	Xh	1316
Setswana	tn	1050
Sesotho sa Leboa or Sepedi	Nso	8548
Sesotho	St	914
Tshivenda	Ve	808
Xitsonga	Ts	741
Siswati	Ss	656
isiNdebele	-	-

Source: Wikimedia, 2023. All Wikipedias ordered by number of articles, viewed 18 May 2023, from [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias)

be verified for quality. Translation quality assurance is a challenge in many translated materials (House 2013). The benefit of using crowdsourcing in translation is to have language speakers with a variety of expertise contribute to the translation process to improve the quality of the translated materials. Regardless of the challenge associated with translation quality, having web publications translated from English to indigenous languages is a beneficial step toward ensuring that knowledge is open, promoted and conserved, as well as used in a range of fields.

## Digitisation and archiving

Digitisation and archiving can play an important role in the preservation of South African indigenous languages. Digitisation is a method of transforming numerous kinds of information into electronic forms, such as written material, sound, picture and speech (Khan, Khan & Aftab 2015). Technological tools such as scanners and voice recorders can be used to digitise language documents, voice and audio so that they can be stored easily in electronic formats. Optical Character Recognition software is suitable for the digitisation of documents as it converts hardcopy documents and images into digital formats (Hocking & Puttkammer 2016). Finlayson and Madiba (2002) noted that all South African indigenous languages have written materials in the form of dictionaries, literary works and terminology lists. These materials exist in hard copies that makes them not easily accessible to every user. Written materials that contain the multifaceted use of indigenous languages including those that are out of print can be digitised to preserve them and make information accessible to everyone. One of the major merits of digitisation is that it increases access to data, information, and resources because many people can reach them without difficulty and there is no need to go to places where printed materials are kept (Khan et al. 2015).

Archiving is also essential in preserving minority and endangered languages (Joshua 2014). Archiving can be done through language documentation in which documents in Xitsonga, Tshivenda, Siswati, and isiNdebele are deposited into digital archives or repositories for different purposes such as language preservation, conducting research, orthography, and spelling verification. The advantage of using digital archives is that documents are stored in a digital form which makes them immune to physical deterioration (Henke & Berez-Kroeker 2016). Joshua (2014) also noted that archiving in the digital format ensures that information is easy to retrieve as it is accessible to every user through the Internet. Therefore, language materials must be archived digitally to preserve them so that future language users can access them freely.

## Digital learning tools

Traditional modes and techniques of learning are being replaced by digital learning, which is a new renaissance in education (Maria et al. 2019). Digital learning involves the use of technology-based resources or platforms to facilitate

teaching and learning. Ahmadi (2018) postulated that digital learning tools have abilities to maximise successful language learning outcomes and increase student engagement. Digital learning tools such as MT tools, online dictionaries, audio, visual multimedia content and social media platforms can all facilitate faster and more comprehensive language acquisition. Language preservation may also be accomplished with the use of these tools. For this study, MT tools and online dictionaries will be examined as digital learning tools to show how they might be utilised to preserve indigenous languages in South Africa.

### Machine translation tools

El-Banna and Naeem (2016) defined MT as a branch of artificial intelligence (AI) that entails the development of specialised computer systems that can translate text between different human languages. Machine translation goes beyond merely translating words from one language to another by applying sophisticated linguistic analysis to the text and selecting the most likely words and sentence structures from corpora that were previously translated texts (Korošec 2011). This simply means that MT is a computer software that automatically translates texts from one natural language into another without the involvement of a human (Baker & Saldanha 2009). Machine translation systems are categorised into two approaches: rule based and statistical. Rule-based MT is dependent on explicit linguistic data, including grammar, morphological dictionaries, bilingual dictionaries and structural transfer rules (Forcada et al. 2011). The statistical MT tools, on the other hand, rely on sizable parallel corpora of human-engineered translations that are used to automatically infer a statistical model of translation (Korošec 2011).

In the context of South Africa, there are two MT tools that support some of the official languages: Google Translate and Autshumato Machine Translation Web Service. From these tools, only Xitsonga as an indigenous language with fewer speakers in South Africa was added. Other indigenous languages with the least number of speakers such as Tshivenda, Siswati and isiNdebele were neglected. Machine translation tools are used to translate documents, articles and educational materials into various languages. Having other indigenous languages in these tools will not only help for learning purposes but will also play a vital role in their preservation. Bird and Chiang (2012:126) also affirmed that 'when source texts are translated into a major world language, using machine translation, we guarantee that the language documentation will be interpretable even after the language has fallen out of use'. Furthermore, Mager et al. (2018) indicated that MT is ideally suited for indigenous languages as it might facilitate communication with other commonly spoken languages as part of a preservation effort. MT tools can be used to preserve the indigenous languages of South Africa by documenting and translating texts from these languages into languages that are widely used. This can help to increase awareness and understanding of the languages and potentially help in their preservation. Therefore, the development of MT tools for the indigenous languages of South Africa can ensure

that they continue to be spoken and used in South Africa and around the world.

### Online dictionaries

Alberts (2017) described dictionaries as:

[I]nformation resources that reflect human knowledge of language and the world. They serve as authoritative reference works on spelling, pronunciations, meaning and usage as well as on the origin of words, and supply translation equivalents in other languages. (p. 32)

There are two categories of dictionaries: printed dictionaries and online dictionaries. Lew (2010) stated that printed dictionaries employ a variety of techniques to provide word meanings in a paper format. Online dictionaries, as described by Hilary and Warwick (2000:839), are any sources of information that are electronically stored and provide details on the pronunciation, definition or use of words. When compared to a hard-copy dictionary, an online dictionary is far more innovative because of the retrieval technology rather than the information content (Hilary & Warwick 2000). Dictionaries are helpful instruments for language documentation and standardisation as they cover and record the vocabulary of a language (Klein 2009).

Few printed dictionaries that were produced by National Lexicographic Units (NLU) and private companies exist in the context of South African indigenous languages. According to an announcement made by the Pan South African Language Board at the beginning of 2005, bilingual dictionaries for Tshivenda, Xitsonga, Siswati and isiNdebele were created (Brand South Africa 2005). The printed dictionaries have partially preserved these languages like other printed materials. The same is true for online dictionary services in siSwati, Xitsonga and isiNdebele. These sites are not adequate to preserve these languages because of various limitations such as lack of specialised vocabulary. To preserve these indigenous languages concisely, language speakers, the NLU and private companies must develop online comprehensive dictionaries that can be easily accessed and updated. These dictionaries should address the current challenges that indigenous languages face, such as globalisation and cultural assimilation. One of the biggest benefits of using an online dictionary is that there are almost no space limitations, allowing it to have more entries and examples than a printed dictionary (Klein 2009). Online dictionaries offer an important avenue for the preservation of languages by providing accessible and comprehensive resources for language users. Such dictionaries can also help to document and preserve the South African indigenous languages, thus making them more accessible to a wider audience.

### Conclusion

This study demonstrated how technology can be used to preserve the indigenous languages of South Africa with fewer speakers. The technological strategies: localisation of daily used technology and/or application software, translation through crowdsourcing, digitisation and archiving, digital

learning tools such as MT and creating online dictionaries provide viable means to preserve the languages. The localisation of daily used technology and/or application software will ensure that language users access tools in their mother tongue. Translation through crowdsourcing makes information open and accessible in indigenous languages on the Internet and across various disciplines. Digitisation and archiving preserve language information and enhance accessibility. Machine translation tools and online dictionaries can also be used to preserve languages by documenting and translating texts from these languages into languages that are widely used. This will render them more accessible to a wider audience. If these technological strategies can be implemented adequately, the promotion of South African indigenous languages will be facilitated, and their preservation could be maintained.

## Acknowledgements

We thank Prof Kosch for contributing to this article.

## Competing interests

The authors have declared that no competing interest exists.

## Authors' contributions

R.M. and M.M. contributed equally to the research and writing of this article with the support of Prof. Kosch with editing.

## Ethical considerations

This article followed all ethical standards for research without direct contact with human participation.

## Funding information

This publication was made possible with support from the South African Centre for Digital Language Resources (SADiLaR). The Centre is a research infrastructure established by the Department of Science and Innovation of the South African government as part of the South African Research Infrastructure Roadmap (SARIR).

## Data availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

## Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors and the publisher.

## References

Ahmadi, R.M., 2018, 'The use of technology in English language learning: A literature review', *International Journal of Research in English Education* 3(2), 115–125. <https://doi.org/10.29252/ijree.3.2.115>

- Alberts, M., 2017, *Terminology and terminography principles and practice: A South African perspective*, McGillivray Linnegar Associates, Cape Town.
- Baker, M. & Saldanha, G., 2009, *Routledge Encyclopedia of translation studies*, 2nd edn., Routledge, London.
- Bird, S. & Chiang, D., 2012, 'Machine translation for language preservation', in M. Ray & C. Boitet (eds.), *Proceedings of the 24th International Conference on Computational Linguistics*, Mumbai, India, December 8–15, 2012, pp. 125–134.
- Brabham, D., 2013, *Crowdsourcing*, MIT Press, Cambridge.
- Brand South Africa, 2005, *Dictionaries for all 11 languages*, viewed 05 March 2023, from <https://www.brandsouthafrica.com/governance/services/rights/dictionaries080605>.
- El-Banna, A.I. & Naeem, M.A., 2016, *Machine translation as a model for overcoming some common errors in English-into-Arabic translation among EFL University Freshmen*, viewed 15 May 2023, from <https://files.eric.ed.gov/fulltext/ED580942.pdf>.
- Esselink, B., 2000, *A practical guide to localization*, John Benjamins Publishing, Amsterdam.
- Etim, A.F., 2016, 'Morphosyntactic structures of Ibibio personal names', *Journal of Science, Education & Humanities* 6(1), 1–16.
- Finlayson, R. & Madiba, M., 2002, 'The intellectualisation of the indigenous languages of South Africa: Challenges and prospects', *Current Issues in Language Planning* 3(1), 40–61. <https://doi.org/10.1080/14664200208668036>
- Forcada, M.L., Ginestí-Rosell, M., Nordfalk, J., O'Regan, J., Ortiz-Rojas, S., Pérez-Ortiz, J.A. et al., 2011, 'Apertium: A free/open-source platform for rule-based machine translation', *Machine Translation* 25, 127–144. <https://doi.org/10.1007/s10590-011-9090-0>
- Galla, C., 2009, 'Indigenous language revitalization and technology from traditional to contemporary domains', in J. Reyhner & L. Lockard (eds.), *Indigenous language revitalization: Encouragement, guidance & lessons learned*, pp. 167–182, Northern Arizona University, Flagstaff, AZ.
- Glosbe.com, 2023, *Swati-English Dictionary*, viewed 15 May 2023, from <https://glosbe.com/ss/en/siSwati>.
- Henke, R.E. & Berez-Kroeker, A.L., 2016, 'A brief history of archiving in language documentation, with an annotated bibliography', *Language Documentation and Conservation* 10, 411–457.
- Hilary, N.E.S.I. & Warwick, U.K., 2000, 'Electronic dictionaries in second language vocabulary comprehension and acquisition: The state of the art', *Paper presented at the 9th European Association for Lexicography International Conference*, Universität Stuttgart, Stuttgart, 8–12th August.
- Hocking, J. & Puttkammer, M., 2016, 'Optical character recognition for South African languages', *Paper presented at the pattern recognition Association of South Africa and Robotics and Mechatronics International Conference*, Stellenbosch University, Stellenbosch, 30th November–2nd December.
- House, J., 2013, 'Quality in translation studies', in C. Millan & F. Batrina (eds.), *Routledge handbook of translation studies*, pp. 534–547, Routledge, London.
- Jiménez-Crespo, M.A., 2013, *Translation and web localization*, Routledge, New York, NY.
- Joshua, W., 2014, 'Archiving for the community: Engaging local archives in language documentation projects', in D. Nathan & P.K. Austin (eds.), *Language documentation and description*, pp. 85–102, SOAS, London.
- Kalinin, A. & Savchenko, G., 2014, 'Using crowdsourcing system for creating site-specific statistical machine translation engine', *Asian Journal of Information Technology* 13(1), 18–20. <https://doi.org/10.36478/ajit.2014.18.20>
- Khan, S., Khan, S. & Aftab, M., 2015, 'Digitization and its impact on economy', *International Journal of Digital Library Services* 5(2), 138–149. <https://doi.org/10.4018/IJDL.S.2015070101>
- Klein, J., 2009, 'South Africa's new African language dictionaries and their use for the African speech communities', *Lexikos* 19, 423–435. <https://doi.org/10.4314/lex.v19i1.49138>
- Korošec, M.K., 2011, 'Applicability and challenges of using machine translation in translator training', *ELOPE: English Language Overseas Perspectives & Enquiries* 8(2), 7–18. <https://doi.org/10.4312/elope.8.2.7-18>
- Lew, R., 2010, 'New ways of indicating in electronic dictionaries; Hope or hype', in *International symposium on lexicography*, viewed 29 May 2023, from <http://www.lexicographystaff.amu.edu.pl>.
- Lumma, A.L. & Weger, U., 2021, 'Looking from within: Comparing first-person approaches to studying experience', *Current Psychology* 42, 10437–10453. <https://doi.org/10.1007/s12144-021-02277-3>
- Luvhengo, N., 2012, 'Linguistic minorities in the South African multilingual context: The case of Tshivenda', MA dissertation, Rhodes University.
- Mager, M., Gutierrez-Vasques, X., Sierra, G. & Meza, I., 2018, 'Challenges of language technologies for the indigenous languages of the Americas', in E.M. Bender, L. Derczynski & P. Isabelle (eds.), *Proceedings of the 27th International Conference on Computational Linguistics*, Santa Fe, New Mexico, USA, August 20–26, 2018, pp. 55–69.
- Maja, I., 2007, 'Towards the protection of minority languages in Africa', PhD thesis, University of Pretoria.
- Maria, I.J., Thirupathi, D., Rajendran, R. & Velumani, B., 2019, 'Technologies, challenges and tools for digital learning', Paper presented at the 2019 IEEE Tenth International Conference on Technology for Education, Goa University, Goa, 9–11th December.
- Martin-Mor, A., 2017, 'Technologies for endangered languages: The languages of Sardinia as a case in point', *MTm* 9, 365–386.

- Matfunjwa, M., 2023, *Friends*, viewed n.d., from <https://m.facebook.com/friends/>.
- Mercuri, S.P., 2012, 'Understanding the interconnectedness between language choices, cultural identity construction and school practices in the life of a Latina educator', *Gist Education and Learning Research Journal* 6, 12–43.
- Mick, D.G., Spiller, S.A. & Baglioni, A.J., 2012, 'A systematic self-observation study of consumers' conceptions of practical wisdom in everyday purchase events', *Journal of Business Research* 65(7), 1051–1059. <https://doi.org/10.1016/j.jbusres.2011.02.033>
- Mirza, A. & Sundaram, D., 2016, 'Harnessing collective intelligence to preserve and learn endangered languages', in P.C. Vinh & L. Barolli (eds), *Proceedings of 2nd EAI International Conference on Nature of Computation and Communication*, Rach Gia, Vietnam, March 17–18, 2016, pp. 273–245.
- Moodley, M., 2020, 'Localising an online computer software to include Setswana, an indigenous African language, for the South African teacher', PhD thesis, University of the Witwatersrand.
- Moser, A. & Korstjens, I., 2017, 'Series: Practical guidance to qualitative research. Part 1: Introduction', *European Journal of General Practice* 23(1), 271–273. <https://doi.org/10.1080/13814788.2017.1375093>
- Motsa, F., 2023, *Siswati language added to Facebook languages*, viewed 08 April 2023, from <https://eswatinipositivenews.com/2023/01/11/siswati-language-added-to-facebooklanguages/>.
- Ndebele, 2014, *Ndebele-English online talking dictionary*, viewed 18 May 2023, from <http://talkingdictionary.swarthmore.edu/ndebele/>.
- Olaifa, T.P., 2014, 'Language preservation and development: The role of the library', *Journal of Library and Information Sciences* 2(1), 23–28.
- Osborn, D., 2010, *African languages in a digital age: Challenges and opportunities for Indigenous languages computing*, Human Sciences Research Council, Cape Town.
- Osborn, D.Z., 2006, 'African languages and information and communication technologies: Literacy, access, and the future', in J. Mugane, J.P. Hutchison & D.A. Worman (eds.), *Selected Proceedings of the 35th Annual Conference on African Linguistics*, pp. 86–93, Cascadia Proceedings Project, Somerville.
- Riza, H., 2008, 'Indigenous languages of Indonesia: Creating language resources for language preservation', Paper presented at the IJCNLP-08 workshop on NLP for less privileged languages, Language Technologies Research Center (LTRC), Hyderabad, 7–12th January.
- Rydberg, A., 2023, 'Self-observational life in eighteenth-century Germany', *Intellectual History Review* 1–22. <https://doi.org/10.1080/17496977.2022.2154995>
- Sandrini, P., 2008, 'Localization and translation', *MuTra Journal* 2(1), 167–191.
- Statistics South Africa, 2022, *Census 2022 Results*, viewed 27 October 2023, from [https://census.statssa.gov.za/assets/documents/2022/Census\\_2022\\_SG\\_Presentation\\_10102023.pdf](https://census.statssa.gov.za/assets/documents/2022/Census_2022_SG_Presentation_10102023.pdf).
- 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>
- The South African Constitution, 1996, Constitution of the Republic of South Africa, viewed 16 May 2023, from <https://www.justice.gov.za/legislation/constitution/>.
- UNESCO, 2023, *Cutting edge | Indigenous languages: Gateways to the world's cultural diversity*, viewed 08 April 2023, from <https://www.unesco.org/en/articles/cutting-edge-indigenous-languages-gateways-worlds-cultural-diversity>.
- Wikimedia, 2023, *All Wikipedias ordered by number of articles*, viewed 18 May 2023, from [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias).
- Xitsonga.org, 2023, *Xitsonga online dictionary*, discontinued 2024, viewed 09 May 2023, from <https://www.xitsonga.org/dictionary>.