**Preparing an isiNdebele LSP Dictionary: From Idea to design**

Sponono MAHLANGU, isiNdebele National Lexicography Unit,

University of Pretoria, South Africa (sponono.mahlangu@up.ac.za)

Cael MARQUARD, Department of Computer Science, University of Cape Town, Cape Town, South Africa. (cael.marquard@gmail.com)

This paper discusses the plan for the compilation of an online IsiNdebele Dictionary, an isiNdebele Language Special Purposes (LSP) dictionary. IsiNdebele is a lesser resourced language with a limited corpus available. The isiNdebele LSP dictionary will be the first online dictionary that will be compiled by the isiNdebele Dictionary Unit. The software for the proposed online dictionary will be adapted from the isiXhosa.click dictionary. The dictionary will be compiled in phases, with this being the initial phase. The dictionary will be open source and will be tailored to the isiNdebele grammar, bird terminology list, Covid-19 terminology list, and other terminology lists. These terminology lists were chosen because, for example, the bird terminology list will be useful to the users as many language speakers don’t know the names of these birds in isiNdebele. Going forward with the project if possible, pictures of these birds will be added so that the users can be able to differentiate them since most of the birds often look similar. This dictionary will be a useful and a helpful tool to the users. The Covid-19 terminology list has been added because there are new terms that were coined after the eruption of the pandemic and it is imperative that such terms should be captured so that they are known by the language speakers. The online interface of this online dictionary will be in isiNdebele and it will be appropriate for mobile use. However, the users have not been involved yet because this is an initial stage. Once the dictionary development has started, the users will be able to be involved, by suggesting modifications to the word list. The users will have an opportunity to click on the *edit* button to add missing information and to correct errors. Users will also be given an opportunity to make suggestions and submit new words. The online isiNdebele dictionary will be hosted using the IsiXhosa.click software (Marquard 2024). IsiXhosa.click is an open-source, user-friendly, online dictionary software originally built for isiXhosa. It was chosen for the IsiNdebele online dictionary because it is open-source and free to use, supports capturing of grammatical information and examples, and has already been successfully adapted to isiZulu (Buthelezi & Marquard 2024). Since isiZulu and isiNdebele are mutually intelligible, adapting the software to isiNdebele is expected to be straightforward. Additionally, it allows not only the users but also members of the public to suggest edits and contribute new words and example sentences, thereby enabling the dictionary to grow over time through community input.
The pertinent differences between the three languages are the different prefixes used for the noun classes. In order to adapt to isiNdebele, the noun class prefixes were made configurable and then set to the correct ones for isiNdebele. The software is well-suited for these languages, as they all belong to the Nguni language group and share similar grammatical structures, including parts of speech.

 The incorporation of the terminology lists is performed using a set of Python scripts (one for each format). The script extracts the isiNdebele and corresponding English terms from terminology documents (usually PDFs) and reformats them into a format suitable for import into the dictionary.

Once the dictionary is completed it will be made available online through a user-friendly software interface. It will be open source and may be freely adapted for other projects that the dictionary units might think of in future.

**References**:

Buthelezi, M. & Marquard, C. 2024. Using computational tools and a corpus lexicography framework in developing an isiZulu LSP Dictionary. In: *African Association for Lexicography. 28th International Conference.*1 July, 2024. Pretoria, South Africa: 20-22. Available: [https://www.afrilex.co.za/\_files/ugd/bd0c6e\_369dc939341044a6b4671d36d459717c.pd](https://www.afrilex.co.za/_files/ugd/bd0c6e_369dc939341044a6b4671d36d459717c.pdf)f. Accessed on 24/03/2025.

Marquard, C. 2024. IsiXhosa.click: online, open, user-friendly, and searchable isiXhosa-English dictionary software. In: *African Association for Lexicography. 28th International Conference.*1 July, 2024. Pretoria, South Africa: 35-36. Available: <https://www.afrilex.co.za/_files/ugd/bd0c6e_369dc939341044a6b4671d36d459717c.pdf>. Accessed on 24/03/2025.