

# MaLex – Malawi Lexicon Project

**Pascal Kishindo** University of Malawi • **Dorothee Beermann** NTNU • **Daniel Ridings** NTNU • **Lars Hellan** NTNU

The MaLex project is a cooperation between the Centre of Language Studies (CLS) at the University of Malawi, Zomba, Malawi, and the Department of Language and Communication Studies (ISK) at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway. The project is financed by NUFU. Prof. Lars Hellan, NTNU, is the Norwegian coordinator and Prof. Pascal Kishindo, University of Malawi, is the coordinator in Malawi.



## Main Objectives

The objectives of the Malawi Lexicon project are to fortify language resources in Malawi for applied and theoretical sciences. The “Lex” in MaLex stands for “lexicon”, and dictionary development is one of the objectives of the project. A further goal is the in-depth annotation of representative corpora of the languages of Malawi, starting with Chichewa. An overall objective of the project is to be instrumental in the build-up of a research infrastructure that can further the work that has been done also after this project is completed. The infrastructure consists of academic achievements in the form of the completion of post-graduate degrees and staff training with respect to computational techniques used in language technology.

## Research component

For dictionary development as well as for the annotation of language corpora, MaLex makes use of database technology. Modern digital technology allows for the building of reusable resources so that one lexical database can serve as basis for the development of different types of paper dictionaries. To build a local research infrastructure that supports the use of digital linguistic tools which allow the local linguists to meet the needs of language development

and modern language documentation, is one of the challenges that linguistic research in Africa faces.

## Educational component

The MaLex project supports two graduate students, working on HIV and AIDS related terms and their adaptation to Chichewa and Verbal Extensions in Citumbuka respectively. It further engages into the training of local personnel at the Centre of Language Studies (CLS).

## Best practice

In the past, large archives of paper slips were created in order to supply the lexicographer with the evidence needed to build a dictionary. Large teams of readers were required to go through the literature, to excerpt words That needed to be defined and to document their usage in the form of citations. More generally, language description and language documentation exists up to now mainly in form of academic publications which are often hard to access and mostly unreadable for non-experts. The MaLex project makes use of modern digital tools, in particular database systems to facilitate the work of local linguists and language experts and to support their efforts to develop language resources that represent and promote the

local languages nationally and internationally. The MaLex lexicographers have access to hundreds of thousands of words that can be presented in various ways, from specific contexts to significant collocations, while the project itself and in particular the MaLex graduate students use databases to develop annotated corpora which will be made public to foster further research on the languages of Malawi.

## Capacity building

The advanced techniques of working with empirical data are not easily implemented. The staff at the Centre for Language Studies are continually trained in the background principles and the practicalities involved with implementing and administering the computational infrastructure that lexicography builds upon. Through graduate work in linguistics the project trains experts not only in data management but also in the deeper understanding of language. Both educational initiatives will be instrumental for future work on language preservation and development.

## Relevance (national strategy)

Malawi has become one of the African countries that uses the mother tongue as the language of instruction in the primary schools. The Centre of Language Studies was one of the key players in the process that resulted in this decision.

“The policy was motivated by fact that research had shown that pupils learn better in their own languages, but since there was no research done in Malawi to justify adoption of a new policy, the Centre was engaged to start the process,” says Prof. Al Mtenje, who became the first director of the Centre.

In order for the policy to succeed, educational material in the local languages is needed. Dictionaries provide the backbone of such materials. The University of Malawi already offers courses in the area of syntax and phonology; to strengthen research on local languages is a further important goal.

## Sustainability

The MaLex project strives to enrich the language community in Malawi and to provide primarily schools and universities with tangible results, not only in the form of teaching material but also through the import of modern language technology that will allow access to national language resources not only by language experts. The work of the MaLec project is supported by the Centre for Language Studies.

Four of the Centre’s core mandates coincide with the objectives of the MALEX project:

1. Establish orthographic principles of Malawian languages
2. Develop descriptive grammars of Malawian languages
3. Compile lexicons of Malawian languages
4. Promote research

The goals of the project are integrated seamlessly with existing activities and become a natural part of the working methodology.

