

## Introduction

The Bamanankan [also known as Bamana or Bambara] is the last language on our list of lexicons in the context of the Mandenkan Lexicon Project at the LDC. It is the most important Mandenkan in terms of size and number of speakers, with those speakers mostly located in West Africa. According to Ethnologue, 15 million people speak Bamanankan [4 million native speakers in Mali, and 10 million persons who are second-language users]. The actual evaluation of the speakers of the different dialects of the Mandenkan stands at 30 to 40 million people mainly in Mali, Burkina Faso, Senegal, Guinea Bissau, Guinea, Liberia, Ivory Coast, Sierra Leone and the Gambia. Bamanankan happens to be also the most studied among the Mandenkan, certainly because it is spoken as first or second or language by at least 80% of the population of Mali, and is vastly used as second or/and third language in most of the West African countries. Therefore, this Bambara lexicon is one that can be the most helpful in reaching the main goal of the Mandenkan Lexicon Project.

There exist a certain number of lexicons of the Bambara, with few available online as is the SIL Lexique Pro Bamanankan – French, Bamanankan – English. However, as stated in the introductions of the Mawukakan and Maninkakan lexicons, the transcription system adopted for this project is the most suitable for the development of technologies for the Mandenkan dialects. The reason given by some researchers for the omission of certain tones in the Mandenkan textbooks is to make the learning easier. If that makes perfectly sense in the context of literacy where the learners are native speakers of the language, it is simply not helpful in all the other contexts of the study of that language. Our view is that given the role of tone in Mandenkan, the omission of certain tones in the transcription for reasons of simplification does not necessarily make the learning easier. For example, it's not an accident that true tonal systems exclude complex syllable structures, allowing only open syllables and consonant clusters where one component is a glide /y/ or /w/. The omission of a linguistic constituent [a segment or a tone] in the transcription should only be considered if the public the transcript is destined to is capable of restituting the missing element when necessary. It is essential for the learner to have a clear idea of the complete forms of words, and understand later how principles and specific rules apply to those full forms to derive their surface forms. But a more forceful argument in favor of the decision to mark all the tones is the fact that the presence of every tonal constituent will be relevant in any attempt to develop for example a Machine Translation tool for each specific dialect. The tones are very important in the process of discriminating a Mandenkan variety from another, helping both researchers and learners understand the prosodic specificity of each language.

With more than 10 million speakers, the Bamanankan lexicon could be the one to boost the interest for the Manding languages.