

Describing the Ngbugu language (Banda, CAR) at a distance : face-to-face and remote cooperation with speakers

Due to the country's political instability, Central African Republics' (CAR) languages have been little studied in the last thirty years. As a result, important parts of these languages remain underdescribed to this day. Of those, the Ngbugu language (Banda; CAR), could be considered a serious example: while most of its only available data are lists of isolated words (Mbomate and Olson 2003, 200 words, Boyeldieu 1997, 1520 words), no global grammatical description has been proposed yet. I am currently engaged in a PhD dissertation whose aim is precisely to work on a description of this language. In this talk, I will present the method I use to describe some aspect of the Ngbugu language without going in CAR and the impact this method has on the resulting analysis. I will also discuss some ethical issues relative to the type of contact I have engaged with speakers.

The description of the Ngbugu language is possible thanks to two elements : Native speakers living in France and those who use social media. I met Ngbugu speakers in face-to-face encounters (diaspora organization, central African restaurant) and through social media (Facebook). I currently collaborate with three consultants - Ngbugu speakers who live in France – and I receive data from Ngbugu speakers living in CAR through online services such as WhatsApp or Messenger.

Speakers living in France help me transcribe and translate texts as well as performing elicitations. They also have a central role in data collection: they mobilize their personal network for audio recordings, which are recorded directly by them or sent via social media apps. The collection and analysis of data in Ngbugu therefore relies on the interest of speakers living in France in having a scientific description of their language done. The recent creation by a Ngbugu speaker of a facebook page called "bichi ngbougou – je suis ngbougou – I am ngbougou", on which discussions about the Ngbugu language are frequent, and on which people living in France or in CAR are present, shows the interest of some speakers for the transmission of their language to future generations.

This cross collaboration between speakers from the diaspora and remote speakers facilitates the description of the language, in a context where it is complicated to physically go to the areas where it is spoken. It does, however, have repercussions on the type of data collected that should be interrogated. The fact that these speakers are responsible for the data collection has an impact on what exactly is collected: they are the ones who decide on the type of text collected and on who is actually recorded. Their choices are decisive in the corpus constitution on which the language description is based. Considering the distance with some recorded speakers, (geographical and in the interpersonal network), data collection is not planned but opportunistic. One of the most important consequences is that the data collected may come from speakers of different Ngbugu dialects. Distant data collection doesn't permit direct observation of the context of Ngbugu speaking. As a result, some parts of the language are not studyable (dialectology, spatial deixis, fauna and flora lexicon, ...). At the technical level, audio files emanating from social media apps do not meet the quality standards normally expected. This situation also leads to ethical issues regarding the diffusion of the data, as the speakers' consent in this regard is not always clearly expressed.

References (selection) :

- Mbomate, Jacques Vermond, Kenneth S. Olson. 2007. « Ngbugu digital wordlist: presentation form ». *Linguistic Discovery* 5(1):40-47. doi: [10.1349/PS1.1537-0852.A.317](https://doi.org/10.1349/PS1.1537-0852.A.317).
Boyeldieu, Pascal. 1997. « Lexique ngbugu ». <http://www.reflex.cnrs.fr>