



5th Workshop on Spoken Language Technology for Under-resourced Languages, SLTU 2016,
9-12 May 2016, Yogyakarta, Indonesia

Parallel Speech Collection for Under-resourced Language Studies using the LIG-AIKUMA Mobile Device App

David Blachon^a, Elodie Gauthier^a, Laurent Besacier^{a,*}, Guy-Noël Kouarata^b, Martine
Adda-Decker^b, Annie Rialland^b

^aLaboratoire d'Informatique de Grenoble (LIG) / GETALP group, France

^bLPP, CNRS-Paris 3 / Sorbonne Nouvelle, France

Abstract

This paper reports on our ongoing efforts to collect speech data in under-resourced or endangered languages of Africa. Data collection is carried out using an improved version of the Android application AIKUMA developed by Steven Bird and colleagues¹. Features were added to the app in order to facilitate the collection of parallel speech data in line with the requirements of the French-German ANR/DFG BULB (Breaking the Unwritten Language Barrier) project. The resulting app, called LIG-AIKUMA, runs on various mobile phones and tablets and proposes a range of different speech collection modes (recording, respeaking, translation and elicitation). LIG-AIKUMA's improved features include a smart generation and handling of speaker metadata as well as respeaking and parallel audio data mapping. It was used for field data collections in Congo-Brazzaville resulting in a total of over 80 hours of speech. Design issues of the mobile app as well as the use of LIG-AIKUMA during two recording campaigns, are further described in this paper.

© 2016 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of the Organizing Committee of SLTU 2016

Keywords: Speech collection tool; under-resourced languages; language documentation

1. Introduction

The growing proliferation of smartphones and other interactive voice mobile devices offers new opportunities for field linguists and researchers in language documentation. The ease of collecting large volumes of data lowers the pressure of defining the best sampling selection process, which speakers and what data exactly to collect. We may thus envision endangered language documentation collections growing very large with many speakers and material to study a bunch of linguistic phenomena, from acoustic-phonetics to discourse analysis, including phonology, morphology and lexicon, grammar, prosody and tonal information. Field recordings should also include ethnolinguistic material which is particularly valuable to document traditions and way of living. However, large data collections require well organized repositories to access the content, with efficient file naming and metadata conventions.

* Corresponding author. Tel.: +33-4-76-63-56-95.

E-mail address: laurent.besacier@imag.fr

In the following, we describe our ongoing efforts to achieve an effective mobile device application which easily records primary linguistic data in the speakers' living area. The application automatically generates structured data files with raw speech data as well as field and speaker related metadata. We also describe our experience with collecting speech during two recent fieldtrips.

1.1. Context: the BULB project

The BULB¹ (Breaking the Unwritten Language Barrier) project aims at supporting the documentation of unwritten languages with the help of automatic speech recognition and machine translation technologies. The project relies on a strong German-French cooperation involving both linguists and computer scientists. The aim of the project is to design and adapt NLP methods to speed up linguistic analyses of oral, unwritten languages from the Bantu language family. Bantu includes about 500 languages (figures vary from 440² to more than 660^{3,4} depending on the authors), among those many are exclusively oral and remain unstudied. Towards this aim, we have chosen three languages which are already partly studied and for which a few written resources exist: Basaa (A43), Myene (B11) and Mboshi (C25). In this contribution, we report on our experience in Mboshi data collection using LIG-AIKUMA.

1.2. The respeaking concept

The model of Basic Oral Language Documentation, as adapted for use in remote village locations, which are far from digital archives but close to endangered languages and cultures, was initially proposed by Bird⁵. Speakers of a small Papuan language were trained and observed during a six week period. For this, a technique called respeaking, initially introduced by Woodbury⁶, was used. Respeaking involves listening to an original recording and repeating what was heard carefully and slowly. This results in a secondary recording that is much easier to transcribe later on (transcription by a linguist or by a machine). The reason is that the initial speech may be too fast, the recording level may be too low, and background noise may degrade the content. For instance, in the context of recording traditional narratives, elderly speakers are often required (and they may have a weak voice, few teeth, etc.) compromising the clarity of the recording⁷.

1.3. Paper outline

This paper is organized as follows. Section 2 describes the original application, from which we evolved to LIG-AIKUMA, described in section 3. Then section 4 describes first data collections made in Congo-Brazzaville. Finally section 5 concludes and gives some perspectives.

2. Aikuma - The origins

2.1. The initial Aikuma application and its motivations

The initial smartphone application AIKUMA was developed by Bird et al.¹ for the purpose of language documentation with an aim of long-term interpretability. According to the authors, the application is designed for a *future philologist*¹: it collects enough speech and documentation to allow for a delayed (future) processing by a linguist. Indeed, the authors notice that, in general, language documentation projects lack of resources, especially human resources for the processing of the materials. As a consequence, data may be processed a long time after their collection. Moreover, in the case of an endangered language, there is a risk that the study starts after the language has disappeared. This is why the authors extended the concept of respeaking to produce oral translations of the initial recorded material. Oral translation is performed by listening to a segment of audio in a source language and spontaneously producing a spoken translation in a second language.

AIKUMA is an Android application that supports these requirements: recording of speech, respeaking of audio sources, and oral translation. In the next section, we describe its main features.

¹ <http://www.bulb-project.org>

Download English Version:

<https://daneshyari.com/en/article/485436>

Download Persian Version:

<https://daneshyari.com/article/485436>

[Daneshyari.com](https://daneshyari.com)