Speech Technologies for African Languages: Example of a Multilingual Calculator for Education

To cite this version:
Laurent Besacier, Elodie Gauthier, Mathieu Mangeot, Philippe Bretier, Paul Bagshaw, et al.. Speech Technologies for African Languages: Example of a Multilingual Calculator for Education. Interspeech 2015 (short demo paper), Sep 2015, Dresden, Germany. hal-01170505

HAL Id: hal-01170505
https://hal.archives-ouvertes.fr/hal-01170505
Submitted on 15 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Speech Technologies for African Languages: Example of a Multilingual Calculator for Education

Laurent Besacier¹, Elodie Gauthier¹, Mathieu Mangeot¹, Philippe Bretier², Paul Bagshaw², Olivier Rosec², Thierry Moudenc², François Pellegrino³, Sylvie Voisin³, Egidio Marsico³, Pascal Nocera⁴

¹Laboratoire d’Informatique de Grenoble (LIG), Univ. Grenoble Alpes, Grenoble, France
²Voxxygen S.A. - Pleumeur Bodou, France
³Laboratoire Dynamique Du Langage (DDL), CNRS - Université de Lyon, France
⁴Laboratoire d’Informatique d’Avignon (LIA), Univ. Avignon, France

laurent.besacier@imag.fr

Abstract

This paper presents our achievements after 18 months of the ALFFA project dealing with African languages technologies. We focus on a multilingual calculator (Android app) that will be demonstrated during the Show and Tell session.

1. Introduction

1.1. Context: the ALFFA project

Today is very favorable to the development of a market for speech technologies in African languages. People’s access to ICT is done mainly through mobile (and keyboard) and the need for voice services can be found in all sectors, from higher priority (health, food, education) to more fun (games, social media).

For this, overcoming the language barrier is needed and this is what we propose in the ALFFA project where two main aspects are involved: fundamentals of speech analysis (language phonetic and linguistic description, dialectology) and speech technologies (ASR and TTS) for African languages. In the project, developed ASR and TTS technologies will be used to build micro speech services for mobile phones in Africa. For this, speech fundamental knowledge for targeted languages has to be upgraded while African language technologies are still at their very beginning.

For these reasons, the ALFFA project is really interdisciplinary since it does not only gather technology experts (LIA, LIG, Voxxygen) but also includes fieldwork linguists/phoneticians (DDL).

1.2. Demo Content

This paper describes our achievements after 18 months of project. We notably present the last version of a multilingual calculator prototype in several African languages (including Wolof, Hausa and accented French). This calculator - developed by the industrial partner of the ALFFA project (Voxxygen) - was already presented at the Francophonie’s summit 2014 in Dakar (Senegal). Voxxygen is currently looking for partners, especially in the world of education, who could deploy this educational tool at a large scale.

1. Introduction

1.1. Context: the ALFFA project

Today is very favorable to the development of a market for speech technologies in African languages. People’s access to ICT is done mainly through mobile (and keyboard) and the need for voice services can be found in all sectors, from higher priority (health, food, education) to more fun (games, social media).

For this, overcoming the language barrier is needed and this is what we propose in the ALFFA project where two main aspects are involved: fundamentals of speech analysis (language phonetic and linguistic description, dialectology) and speech technologies (ASR and TTS) for African languages. In the project, developed ASR and TTS technologies will be used to build micro speech services for mobile phones in Africa. For this, speech fundamental knowledge for targeted languages has to be upgraded while African language technologies are still at their very beginning.

For these reasons, the ALFFA project is really interdisciplinary since it does not only gather technology experts (LIA, LIG, Voxxygen) but also includes fieldwork linguists/phoneticians (DDL).

1.2. Demo Content

This paper describes our achievements after 18 months of project. We notably present the last version of a multilingual calculator prototype in several African languages (including Wolof, Hausa and accented French). This calculator - developed by the industrial partner of the ALFFA project (Voxxygen) - was already presented at the Francophonie’s summit 2014 in Dakar (Senegal). Voxxygen is currently looking for partners, especially in the world of education, who could deploy this educational tool at a large scale.

Abstract

This paper presents our achievements after 18 months of the ALFFA project dealing with African languages technologies. We focus on a multilingual calculator (Android app) that will be demonstrated during the Show and Tell session.

Index Terms: ASR, TTS, Android app, African languages

2. ASR and TTS systems developed

2.1. Target languages

Language choice for the project is mainly governed by population coverage and industrial perspectives. We focus on Hausa spoken by around 60 million people, as first or second language. Enlarging the sub-region coverage, we will also consider Bambara, Wolof, and Fulfulde languages, to cover major West Africa languages. Bambara is largely spoken in West Africa by around 40 million people. Wolof is mainly spoken by 10 million people. Fulfulde is a set of dialects spoken in all West Africa countries by 70 million people. Those languages cover more than half of the 300 million people of West Africa. They include both tonal and non-tonal languages.

As far as East Africa is concerned, we designed ASR system for Swahili which is the most widespread language in the East of the continent: more than 60 million people. We also work on Amharic, mostly spoken in Ethiopia (20 millions).

A linguistic description of the targeted languages (white paper) is planned for each target language. So far, white papers were written for Wolof and Swahili.

2.2. ASR

ASR systems for Swahili, Hausa and Amharic have been built so far. All the data and scripts to build a complete ASR system for Swahili and Hausa are already available to the public on a github repository. We used Kaldi speech recognition toolkit for building our ASR systems. For the Swahili and Amharic ASR systems, the transcribed speech corpora, pronunciation lexicons and LMs are also made available while for Hausa ASR, users need to buy the corpus and the lexicon at ELDA first.

More details on the Swahili corpus and how it was collected can be found in. For Hausa, the GlobalPhone Speech Corpus was used. The Amharic system was retrained from the corpus described in. A summary of the ASR performance obtained for the three languages is given in but more experimental details can be found in the README files of the github repository. Current data collection and ASR developments include Wolof language.

2.3. TTS

A viable manner of addressing language development for TTS stems from the incremental approach followed by Voxxygen for several years now. The key point of this methodology is that instead of developing the entire (under-documented and under- resourced) language at once, only part of the language is ad-
onsets while it is going from 90 to 240 milliseconds for open
in closed syllabic context is mostly between 20 and 80 millisec-
lighted with automatic tools. We see on the figure that duration
and /o/, the vowel length contrast (short / long) can be high-
labic context. Figure 1 shows that for two hausa vowels /e/
and contexts, to provide fluid speech with adequate prosody. In
the ALFFA project, first step TTS prototypes have been de
devolopd for Hausa and Wolof and some fillers have been recorded
for the multilingual speaking calculator needs. New develop-
ments concern Fulfulde, Zarma en Serere.

3. Multilingual calculator for education
Education is one of the straightforward domain for which
speech enabled micro-applications are needed. In the ALFFA
project, we propose a multilingual speaking calculator to
demonstrate the potential of speech technologies for African
languages. The prototype, which is an android app, offers dif-
ferent features: a simple voice-enabled calculator where the
user enter the operation from the numeric keyboard (or vo-
cal); then the application reads the operation and the result
; a multiplication table mode where the user chooses a number
then the calculator recites the corresponding table ; and a quiz
mode where the user is invited to submit a complete equation
and the calculator confirms if the answer is correct or not. A
button allows to switch the language at any time or to make
the calculator repeat the last utterance in the new selected language.

The ALFFA project will progressively extend the speaking calculator
with new African languages: Fulfulde, Zarma and Serere are coming shortly. Bambara and Swahili will rapidly follow. Field testing of the multilingual speaking calculator pro-
totype is envisaged in the project.

4. Machine-assisted language analysis
Automatic speech technology offers great opportunities for in-
vestigating a wide range of issues in laboratory phonology. How-
ever, its availability hardly extends beyond a small num-
ber of languages. In the ALFFA project, we develop ASR for
African languages, so linguists can, in the meantime, benefit
from automatic tools (automatic annotation tools, forced align-
ment and segmentation, etc.). For instance, we did a prelimi-
ary study on vowel duration in Hausa depending on their syl-
labic context. This figure confirms, at a large scale (5k utterances),
linguists’ description of vowel contrast phenomenon in Hausa.

<table>
<thead>
<tr>
<th>Task</th>
<th>WER (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swahili broadcast news</td>
<td>20.7</td>
</tr>
<tr>
<td>Hausa read speech</td>
<td>10.0</td>
</tr>
<tr>
<td>Amharic read speech</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Table 1: ASR performance for Swahili, Hausa and Amharic
- HMM/SGMM acoustic modeling - all scripts available on

http://kaldi.sourceforge.net/

http://papillon.imag.fr/

http://www.youtube.com/watch?v=vpIxY2N8rt4