

COORDINATEUR: LIG (Grenoble)

PARTENAIRES: DDL, LIA, VOXYGEN

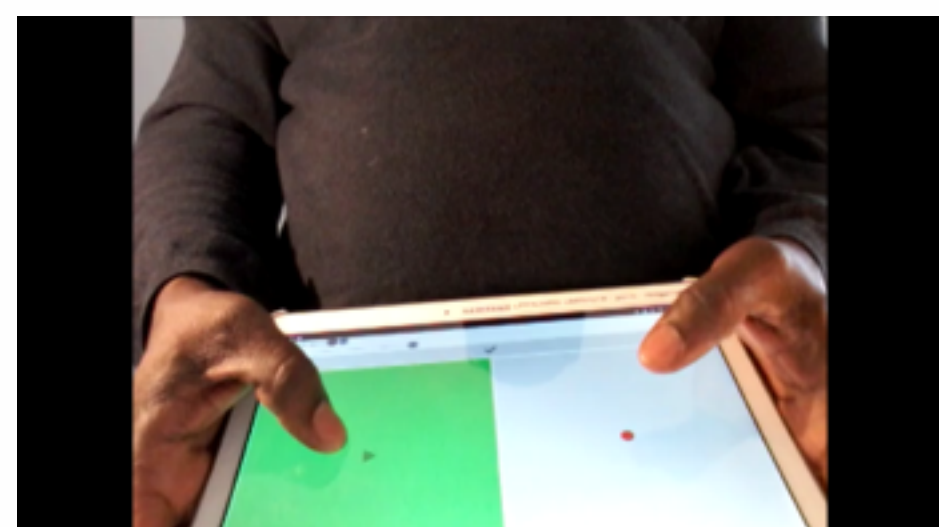
**Abstract:** Today is very favorable to the development of a market for speech in African languages. People's access to ICT is done mainly through mobile and the need for voice services can be found in all sectors, from higher priority (health, food) to more fun (games, social media). For this, overcoming the language barrier is needed and this is what we propose in this 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.

► PROJECT GOALS

- African Languages and Information Technologies
- Address **under-resourced languages** from Africa with immediate or short-term potential
- **Focus on West Africa** (Hausa, Wolof, Fulfulde, Zarma, Bambara and Fongbe)
- **2 East African languages** (Amharic, Swahili)
- **Data collection** methodology
- **ASR** (speech-to-text) and **TTS** (text-to-speech)
- Micro-application **prototypes for mobile** phones

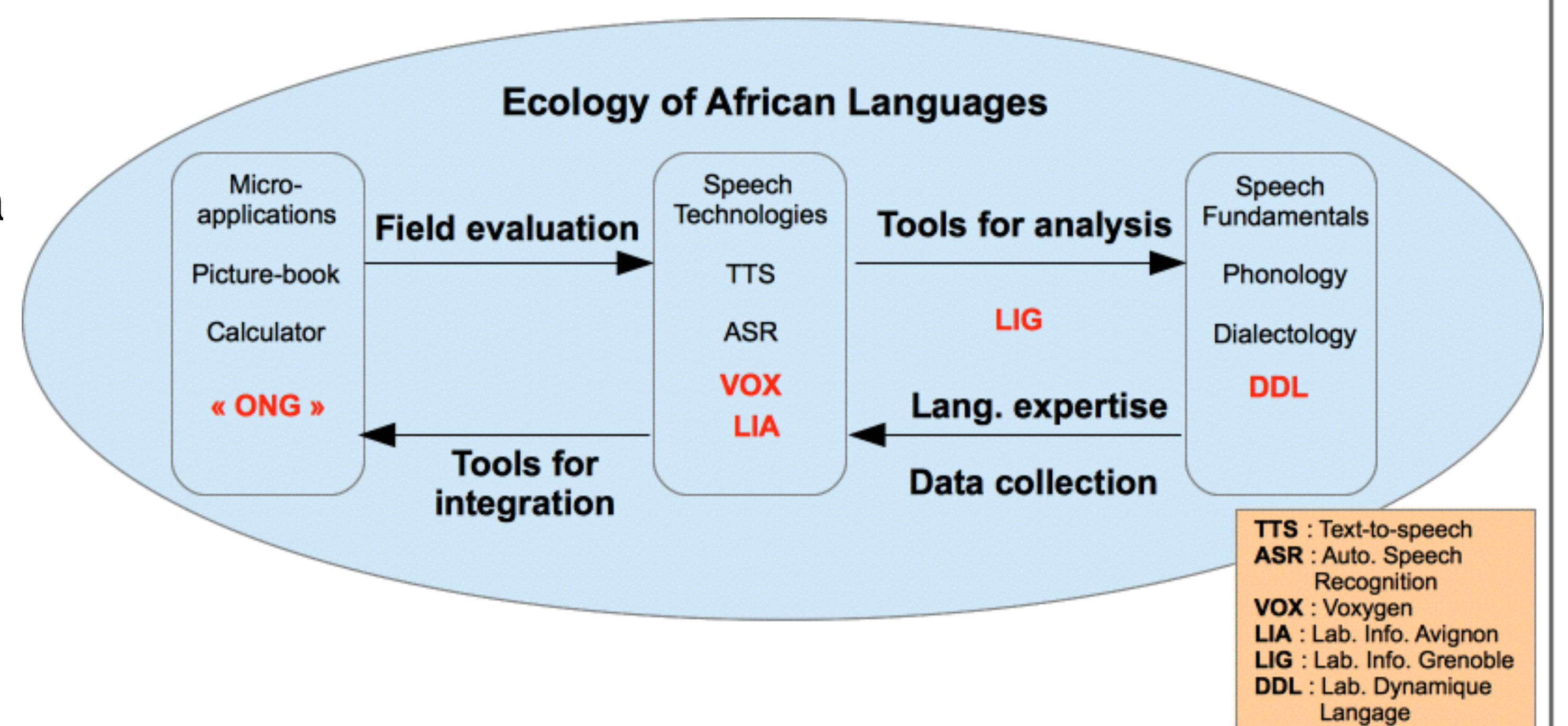
► PROJECT OUTCOMES

- White papers (minimum language description kits for developing voice technologies)  
Hausa, Swahili, Wolof, Fulfulde
- ASR (speech-to-text) data and systems  
Hausa, Swahili, Wolof, Amharic, Fongbe  
[https://github.com/besacier/ALFFA\\_PUBLIC/tree/master/ASR](https://github.com/besacier/ALFFA_PUBLIC/tree/master/ASR)  
(ressources mirrored on OpenSLR repository)
- TTS (text-to-speech) data and systems  
-linguistic modules and lexicons for numerical data restitution (Hausa, Wolof, Zarma, Fulfulde, Pulaar and Serere)  
-TTS voices (Wolof, Hausa)
- Mobile Apps  
-Multilingual Calculator (Voxygen)  
-Data collection app (Lig\_Aikuma)



► PROMOTING R&D ON AFRICAN LANGUAGE TECHNOLOGIES

- **Sénélangues 2015** (spring school - Dakar, Senegal)
- **TALAF 2014** and **2016** (Workshop on African NLP)
- Mailing list on African NLP Research: [talaf@imag.fr](mailto:talaf@imag.fr)
- List of experts compiled (34 names, 20 languages)
- Special session at **Interspeech 2016** (SanFrancisco)



► HIGHLIGHTS

- ALFFA presented at: Francophonie's summit 2014 (**Dakar**), Salon de l'innovation en TALN 2014 (**Marseille**), Africacom 2015 (**CapeTown**), Interspeech 2015 Show&Tell (**Dresden**)
- Voxygen opened an office in **Dakar** (Senegal)
- Voxygen now starts activities in ASR
- ALFFA partners participate to the **Global Learning XPRIZE** (open source software to help children from developing countries to learn basic reading and maths) <http://learning.xprize.org>
- ALFFA/Wolof speech corpus will be used for **ASRU/ ZR-Challenge 2017**

► SAMPLE PUBLICATIONS

- E. Gauthier & al. *Collecting Resources in Sub-Saharan African Languages for Automatic Speech Recognition: a Case Study of Wolof*, **LREC 2016**, Portoroz (Slovenia)
- Elodie Gauthier, & al. *Speed perturbation and vowel duration modeling for ASR in Hausa and Wolof languages*. **Interspeech 2016**, San-Francisco (USA).
- Laurent Besacier & al. *Speech Technologies for African Languages: Example of a Multilingual Calculator for Education*. **Interspeech 2015** Show and Tell session, Dresden (Germany)

