AfricaSign - A Crowd-sourcing Platform for Lexical Documentation of African Sign Languages

Abdelhadi Soudi [1], Kristof Van Laerhoven [2], Elmostafa Bou-Souf [1]

[1] Sign Language Lab and Center for Languages and Communication,
ENSMR-Rabat, Morocco
{asoudi, bousoufelmostafaensmr}@gmail.com

[2] Ubiquitous Computing, Electrical Engineering and Informatics
University of Siegen, Germany
kvl@eti.uni-siegen.de

Abstract

Research in sign languages, in general, is still a relatively new topic of study when compared to research into spoken languages. Sign language research in Africa is still at its infancy, so standardization of communication tools has rarely been attempted. Most of African sign languages are endangered and severely under-studied (Nyst 2010). In an attempt to (lexically) document as many endangered sign languages in Africa as possible, we have developed a low-barrier, online crowd-sourcing platform (AfricaSign) that enables the African deaf communities to document their sign languages. Like many under-researched sign languages, the relationships between African sign languages and other sign languages are not well understood. American Sign Language and European sign languages, including French Sign Language, certainly have some influence on African sign languages, but are unlikely to be related genetically. Our lexical documentation efforts will likely lend insight to the genetic relationships between Africa sign languages and other sign languages. It is also likely that this research will uncover typological features exhibited by African sign languages, thus furthering our understanding of the typology of sign languages and language generally.

1 Introduction

Sign language is a visual/gestural language that is distinct from spoken language and has no text representation. SL is not gestural spoken language. Sign language does not represent in any direct way the form of the spoken language either by visually representing sounds or syntactic sequences of words of the spoken language. One sign may mean an entire spoken language phrase and vice versa. Sign language can only be described, animated or videotaped.

Building on early works by Stokoe (1960, 2005), Losson and Vannobel (1998) and Lebourque and Gibet (1999), several computer applications have been developed to describe sign languages. By way of example, two major EU projects ViSiCAST and eSIGN built a technology for signing avatars. These projects used an XML description language (SiGML) which is based on HamNoSys notation and Lebourque and Vannobel's (1999) gesture specification language (GessyCA). A system was developed to convert HamNoSys code of the given word to its SiGML form, to enable the animation of the avatar.

In this work, we present AfricaSign, a novel and easy-to-use crowd-sourcing web-based platform for Early Grade and STEM Vocabulary in African Sign Languages.

2 System Description

Documentation is categorized by country and region/city to accommodate regional variation for the same language. The design of the crowd-sourcing platform takes into account the quality assurance of the documented signs, and the platform enables users to login and describe themselves demographically by country and region and deafness affiliation (e.g., deaf themselves, have deaf parents (CODA)). As a starting point, the database of words for which we are seeking signs in African Sign languages is based on an enhanced version of MacArthur-

Bates Communicative Development Inventory Words (Fenson et al., 2007). Users can choose the language of the platform, including the vocabulary in the database (English, French, Arabic, or German).
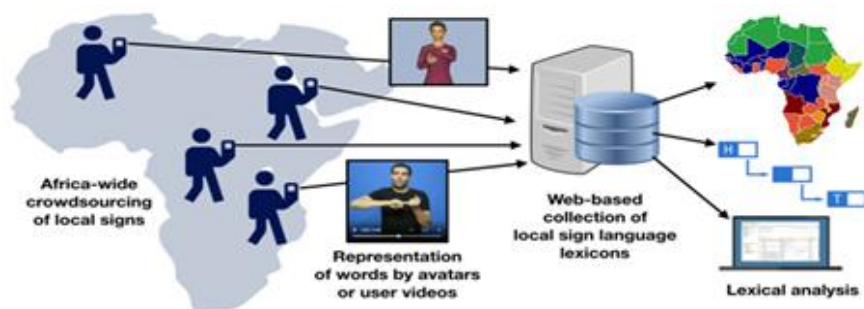

Figure 1: Overview of the AfricaSign platform

Users can add their signs using one of the following three modes:

(1) Uploading a video sign if a user already has it

(2) Videotaping a sign using Laptop/Phone cam: After their consent, the users' phone/laptop cam will be automatically activated and they will be asked to provide a sign for a particular word.

(3) Describing the sign using an avatar: this mode operates by having users identify:

(i) Manual markers/ the four cheremes for each hand (Handshape, Orientation, Location and Movement) for the sign for which they want to find the spoken/written language equivalent.

Users can select these cheremes by using drop-down pictorial menus. The description of signs using the four cheremes is based on work done by the Institute for Disabilities Research and Training (IDRT) for American Sign Language as well as work on Moroccan Sign Language within the framework of two recent projects (Soudi and Vinopol, 2016c). The novelty of the current work is that we are using the chereme-based approach to animate an avatar.

(ii) non-manual markers (head movement and facial expressions (eyebrows, eye aperture and mouthing). Users can select the non-manual marker type by using drop-down pictorial menus.
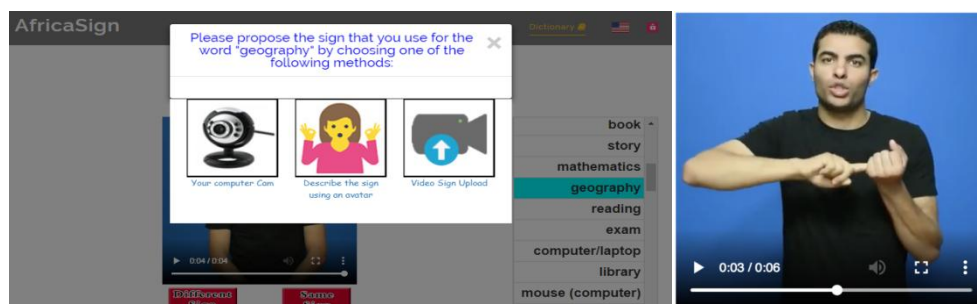

Figure 2: The web-based tool allows new signs to be captured via multiple options (see the pop-up window requesting the user to add the sign for the word "geography"). One example is to use the local camera (embedded in phone, tablet, or PC) to record the sign.

2.1 Application's Research Implications

The data collected through this platform will help answer fundamental technological and linguistic research questions about African sign languages and sign language typology more generally. Core research questions that are addressed are twofold:

(i) Research questions in linguistics:

● How can research on African sign languages further best practices for creating large scale digital resources of minority languages?

● What is the nature of variation in signing across the African continent?

● Are there typologically unusual lexical features of African sign languages?

● What are the theoretical implications of such features?

- To what extent can language (lexical) documentation and conservation address negative attitudes about Deafness in Africa?

(ii) Technology related questions:
- Sign language representation and animation: How can a virtual signing avatar be designed to render it most suitable for displaying sign language, especially for a heterogeneous set of languages and cultures? In this context, we will address the 2 challenging tasks: (i) how to improve the quality of the signing avatar and (ii) in the absence of a standardized evaluation protocol, how to best evaluate a signing avatar.
- Research framework for crowd-sourcing in Africa: How crowd-sourcing can be effectively used as a strategy to collect sign language data in a robust and reliable way. This also involves addressing interface design through usability testing and specific user studies.

2.2   Initial Results

The Moroccan Sign Language (MSL) data collected so far shows a high degree of regional variation. Deaf signers from five regions were given 1500 words signed in MSL by Rabat-Salé region Deaf signers. The former were asked to go through these signs and provide signs that are different from the ones used in their region. Over 500 sign variants have so far been documented. Several typological features have already emerged. For example, some unusual inter-regional and intra-regional typological features have been observed in the numbering system. A large degree of variation is exhibited in the order and sequence of the Tens and Units when signing some numbers. A preliminary investigation reveals that language contact (regional spoken dialects but also contact with foreign sign languages) could be behind this variation, but this observation should be scrutinized by further rigorous linguistic analysis.

References

Battison, Robbin. 1978. *Lexical Borrowing in American Sign Language*.

American Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. 2007. *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual (2nd ed.)*. Baltimore, MD: Brookes.

Hanke, T. and Schmaling, C. 2001. *A hamnosys-based phonetic transcription system as a basis for sign language generation*. In Gesture Workshop 2001, London.

Huenerfauth, M., Zhao, L., Gu, E., & Allbeck, J. 2007. *Evaluating American Sign Language generation through the participation of native ASL signers*. In Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS) (pp. 211–218). Tempe, AZ.

Kipp, M., Heloir, A., & Nguyen, Q. 2011. *Sign language avatars: Animation and comprehensibility*. In Proceedings of the 11th International Conference on Intelligent Virtual Agents (pp. 113–126).

Losson, O. and J.-M. Vannobel. 1998. *Sign language formal description and synthesis*. In Proc. 2nd Euro. Conf. Disability, Virtual Reality and Assoc. Tech., pages 129–137, 1998.

Lebourque, T. and Gibet, S. 1999. *High level specification and control of communication gestures: the GESSYCA system*. In Computer Animation, 1999. Proceedings, page 24ˆaA¸S35, 1999.

Nyst V.A.S. 2010, *Sign languages in West Africa* In: Brentari D. (Ed.) Sign Languages - A Cambridge language survey.Cambridge: Cambridge University Press.

Soudi, A., and Vinopol, C. 2016. *A Digital Moroccan Sign Language STEM Thesaurus*. In Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining. LREC'16. Portoroz (Slovenia), May 23-28, 2016.

Stokoe, W. C., 1960, *Sign language structure: An outline of the visual communication system of the American deaf*, in Studies in Linguistics, Occasional Papers 8, University of Buffalo Press [reedition 2005 in JOURNAL OF DEAF STUDIES AND DEAF EDUCATION, n°10, pp. 3-37] pp. 1-15.

Stokoe, W. C., Casterline, D., & Croneberg, C. 1965. *A dictionary of American Sign Language on linguistic principles*. Silver Spring, MD: Linstok Press.

Smith, R., & Nolan, B. 2015. *Emotional facial expressions in synthesised sign language avatars: A manual evaluation*. Universal Access in the Information Society, 1–10. (Retrieved from http://dx.doi.org/10.1007/s10209-015-0410-7 (last accessed November 20, 2015)).

Sutton, V. 1984. *Sign writing updates*. Boston, MA: The Center for Sutton Movement Writing.