Developing a Somali Dictionary Application

Nick Thieberger and Nadia Faragaab

New technologies offer access to unprecedented amounts of information and, while the equitable cost of access has been a major problem for distribution of such information, which is now changing. Mobile devices are becoming cheaper so more people from a wider range of backgrounds and speaking a wider range of languages are using the Internet. Support for the many less commonly spoken languages of the world has become a focus in the academic discipline of linguistics. This includes developing a presence for these languages on the web and in mobile devices. This brief report discusses one such example: the Somali-English Dictionary application (app), released in June 2014 by a Melbourne, Australia, team headed by the Somali artist Nadia Faragaab.

Somali Language and Technology

In Somalia, the number of Internet users doubled between 2000 and 2011, but that still only represents two percent of the population. There were some 600,000 mobile users at the end of 2007,[1] and over a million in 2011, with projected estimates of over six million in 2015.[2] However, the speed of the mobile phone Internet service does not exceed 50 kilobytes per second at best.[3]

Although access remains slow and expensive in Somalia, there is a large population of some 1.5 million Somalis living elsewhere in the world in whose everyday lives we can expect the Internet and mobile telephones to be a normal part. Thus, a dictionary can be used to assist

in communication where Somali is not the dominant language. It can also support the use of Somali by younger Somalis in the diaspora for whom the lure of other languages can sometimes overwhelm everyday use of their ancestral language.

Recognizing that dominant languages in the world are putting pressure on the languages of smaller population groups and that Somali in particular has little in the way of a web presence or technological support, we wanted to redress this situation through the publication of a free Somali dictionary. We assessed what was currently available and decided there was a need for a detailed and authoritative dictionary that included recordings of spoken words as examples. With the help of various small grants, we built this free dictionary for use on phones and other portable devices.

Our Procedure

As part of her 2010 research into pictorial representations of Somalis and Somalia for her 2011 Kronologies exhibition,[4] Nadia Faragaab used search engines like Google (Image) to see what resulted when the words "Somalia" or "Somali" were entered into the search bar. She also searched to ascertain what sort of apps were available with a connection to the Somali language. At the time, the apps that were available were mainly games that had names and key words such as "shoot the pirate," as well as a very good translation of the Bible, which has since disappeared. The Bible was probably the most functional and well-produced app in the Somali language. The Bible app had Somali written text as well as audio that allowed the user to listen to the reading of the text. Faragaab sought and found app developers and provided a small sample of Somali and English word entries with the intention of adding to the data in later versions. The developers produced a working app, but it become clear very early on that the app needed more work and, more importantly, required a team that included linguists and developers with prior experience working on such projects.

Faragaab got in touch with Nick Thieberger, a linguist in the School of Languages and Linguistics at the University of Melbourne, who has an interest in documentation of languages and the tools for doing this type of work. Together, they explored existing Somali (print) dictionaries and approached a publisher for permission to use the text. The dictionary selected was Zorc & Osman (1993) and Dunwoody Press not

only gave permission, but after the project was explained, it also very kindly provided a digital version of the text in Microsoft Word format.

We knew of an existing dictionary app that provided a framework in which we could create a Somali version. We asked the developers of that app for details on cost and data format requirements to apply for funding. In December 2013, after several unsuccessful funding attempts, we received a University of Melbourne Vice Chancellor's Engagement Award, which covered the cost of building the content for the app. Nick Thieberger has a background in lexicography (for example, he created the South Efate online dictionary[5]), and converted the MS Word document to the structured format required for importing it into the dictionary app. In this way, we are able to show some 26,000 words and their definitions. Each part of the dictionary (headword, definition, part of speech, and so on) had to be explicitly marked and then entered into a spreadsheet for use in the app.[6]

The Ma! Framework[7] and the Somali App

Initially developed for Iwaidja, an Australian Aboriginal language, the app framework called *Ma!* has caught the imagination of a number of different language groups, and projects are underway with the Aboriginal languages Yorta Yorta and Pitjantjatjara/Yanykunytjara, and with the African languages Bena Bena and Mokpe. A generic version of the phrase book function of *Ma!* is being developed for some twenty different indigenous Australian languages. In part, this is because it is available for both iOS and Android, and a major factor is that most dictionary apps created in this framework are provided for free.

What does the Somali version of the framework provide? The functions the user sees are, as would be expected, a list of words in the language, with definitions in English. Words can be searched for and the result will usually have an example sentence showing how to use the word. Selected Somali words have audio files to allow users to hear them. Users can register themselves if they want to be able to upload content, and there is a web-based dashboard for Faragaab to administer changes offered by users. The app includes a page outlining brief information about the Somali language and details about abbreviations used in the dictionary.

The Future of the App

Having developed Version 1 of the app, we are aware of the possibilities offered by the medium and want to keep working on additional versions. Exposing the dictionary has also shown us where the text and the structure of the framework can be improved; for example, allowing cross-references in the dictionary to be clicked on to navigate the text. We want to include common idioms, greetings, and other useful expressions. For any of this work to continue we will need further funding, which is requested at the Burji website.

Notes

1. mbendi.com/indy/cotl/af/so/p0005.htm, accessed online on 28 May 2014.

2. shuraako.org/sites/shuraako.org/files/hogan_lovells_-_somalias_success_industry.pdf.

3. hiiraan.com/news4/2012/Oct/26384/internet_accessibility_growing_steadily_in_somalia.aspx, accessed online on 28 May 2014.

4. See the video of this exhibition at vimeo.com/33063726.

5. paradisec.org.au/SELexicon/index-english/main.htm.

6. See also the blog entry paradisec.org.au/blog/2014/05/somali-phone-dictionary/.

7. themaproject.org/?p=28.

Bibliography

Zorc, R. D. P., and M. M. Osman. *Somali-English Dictionary with English Index* (3rd ed.). Kensington, MD: Dunwoody Press, 1993.