Localising for Yoruba: Experience, Challenges and Future Direction

Tunde Adegbola
African Languages Tech.
Init. (Alt-i), Ibadan, Nigeria

taintransit@hotmail.com

Kola Owolabi
University of Ibadan
Ibadan, Nigeria

kolawoleowolabi@yahoo.co.uk

Tunji Odejobi
Obafemi Awolowo
Univesity, Ile-Ife, Nigeria

oodejobi@yahoo.com

Abstract

This paper discusses the localisation of the MS Vista operating system for the Standard Yoruba language. A list of English lexical items used in MS Vista was provided and the task is to generate equivalent Yoruba terminologies. The Yoruba terminologies are required to convey the semantic connotation of the concepts presented in the MS Vista user interface. A number of techniques based on the linguistic structure of the Yoruba language were applied in generating translations and deriving equivalent lexical item words. The use of the glossary to develop the Yoruba version of MS Vista indicates that they capture the essence of the task required in the user interface. Although some user expressed reservation on non-technical aspect of the work, on the whole majority of the users expressed satisfaction on the outcome. We conclude from our experience that the task of localising for a language is a multi-disciplinary endeavour which will demand the expertise of Linguists and computer scientists.

1. Introduction

To improve the quality of human-computer interaction for Africans in the MS Vista environment, Microsoft Corporation extended its Local Language Program (LLP) to cover the localisation of Microsoft Windows operating system in various African languages. Yoruba is one of these languages. The African Languages Technology Initiative was invited into a partnership to moderate the development of Yoruba terminologies from a glossary of computer terminologies in English. A number of challenges were encountered in the process of developing the appropriate Yoruba glossary from the English glossary. This paper presents the efforts made in addressing these challenges and thereby suggests strategies that may be used to facilitate similar future work in Yoruba and other African languages. In section 2 we describe the standard Yoruba Language briefly, with focus on its linguistic properties that are of interest in the morphology of lexical and related terms in the language. In Section 3, we discuss the methodology that underlies the technique that guided the generation of the lexical terms. Section 4 documents our experience as well as suggestions on extending the work to other African languages. Section 5 concludes this paper.

2. The standard Yoruba language

The Yoruba language comprises several dialects but this work is based on the Standard Yoruba (SY) language which is the language of trade, education, mass communication and general everyday interaction between Yoruba people, whatever their dialects of the language might be. SY is spoken by over 40 million people, mainly in West Africa. In Nigerian, it is spoken in Lagos, Osun, Ogun, Ondo, Oyo, Ekiti and Kwara, as well as some part of Kogi state. Yoruba is also one of the major languages spoken in Benin Republic and it is also spoken in some parts the Republic of Togo. In addition, the Yoruba language survived the transatlantic slave trade and is therefore used largely as a language of religion in a number of countries of the Americas such as Cuba and Brazil as well as some Caribbean countries.

Prior to the localisation of Microsoft Vista, Yoruba had never been widely used in the domain of modern technology in general, and computer technology in particular. However, there had been concern among

members of Egbe-Onimo-Eded-Yoruba (Yoruba Studies Association of Nigeria) about the need to widen the domains of use of the Yoruba language. This led to the development of *Ede-Iperi* Yoruba (Yoruba Metalanguage) Volume I in 1984 and Volume II in 1990. Between the dates of the publication of these two volumes of *Ede-Iperi* Yoruba, The National Language Centre of the Ministry of Education produced "A Vocabulary of Primary Science and Mathematics" in 9 Nigerian languages including Yoruba. *Ede-Iperi* Yoruba was addressed at the teaching of Yoruba through the medium of Yoruba in tertiary institutions while A Vocabulary of Primary Science and Mathematics was addressed at the teaching of mathematics and Science at the primary school level in Nigerian languages. Hence, even though these earlier efforts provided useful background materials, the localisation of MS Vista (or any other computer software for that matter) for Yoruba demanded more than mere translation of computer terms into Yoruba. The project necessarily demanded the creation of Yoruba equivalent terms, which involves application of scientific strategies and principles for technical-term creation.

3. Methodology

3.1 Strategies for Term compilation and terminography
Since the project was seen as one of term compilation and terminography rather than mere translation, basic strategies based on the following concepts used in term compilation and terminography were employed.

a) Base forms
b) Derivations
c) Obtainable collocations

These strategies were adopted in the building of equivalent Yoruba terminologies from the glossary of terms supplied by Microsoft. Examples of their deployment are as follows:

Modulation as presented in the glossary is a concept that suggests changing of the behaviour something. Hence, to derive the Yoruba terminology for modulation, the base form; **modulate** was first derived as *yí (ìṣesí) padà*. From this base form, other English derivations from modulate were appropriately translated into Yoruba as follows:

Modulate	yí (ìṣesí) padà
Modulator	ayíṣesí-padà
Modulation	ìyíṣesí-padà
Demodulate	dá (ìṣesí) padà sípò
Demodulator	adaṣesí-padà
Demodulation	ìdáṣesí-padà
modulator/ demodulator	ayíṣesí-padà/ adaṣesí-padà

Other similar examples are:

Active	Aṣiṣé
Directory	àkójọpò fáìlì
Domain	àgbègbè ìkápá
Services	àpèsè
Active document	àkọsílẹ̀ aṣiṣé-lé-lórí
Active object	ohun àṣiṣé-lé-lórí
Active field	Ìdá rẹ́kọ̀ọ̀dù aṣiṣé
Active window	wínńdò aṣiṣé

Identify	ṣèdámọ̀;(ṣè ìdámọ̀)
Identified	àṣèdámọ̀; àdámọ̀
Identification	ìṣèdámọ̀; ìdámọ̀
Identifier	aṣèdámọ̀
unique identifier	aṣèdámọ̀ àso
globally-unique identifier	aṣèdámọ̀-àso káríayé

3.2 Formulation devices employed:
a) Semantic extension
This involves extending the meanings of indigenous Yoruba words or items, e.g.

Buffer	Àká (original Yoruba meaning: barn, store)
Chart	àtẹ (original Yoruba meaning: tray for displaying things)

8

b) Description
This involves describing English terms based on their salient attributes, such as functions/purpose, manner of application/ production, appearance, behaviour and other noticeable peculiarities.

Function/purpose, e.g.
Certification   Ìjérìísí (lit. to bear witness to or provide evidence for something)

Manner of application or production, e.g.
Attachment   Àsomọ́ (lit. that which is attached to something else)

Appearance, e.g.
angle brackets   àkámọ́ onígun (lit. Brackets with angles)

Behaviour, e.g
Brackets   Àkámọ́ (lit. That which encloses things)

c) Coinage
This involves inventing a word or phrase, e.g.
Alphabet   Ábídí

Loan translation or calque, e.g.
d) Conditional   Kání (lit. *if we say ….*)
This is a calque from the protasis of a Yoruba conditional sentence as opposed to the apodosis.

e) Borrowing
This involves adoption or borrowing of a word or linguistic expression. The borrowed words are integrated into the phonological structure of Yoruba, e.g.
Megabyte   Mẹ́gábáìtì

f) Composition
This involves combining two or more Yoruba words or morphemes, e.g.

Active   Aṣiṣẹ́ (a-ṣe-iṣẹ́ lit. Something that is working)

3.3 Derivational tactics
The most prominent tactic used is nominalisation (i.e. noun formation) via the highly productive morphological processes of, e.g.

a) pre-fixation
identifier   Aṣèdámọ̀ (from ṣèdámò prefixed with a)

b) compounding
Index   ètò ìtọ̀ kasí (from ètò and ìtòkasí implying a collection of indicators)

All these procedures were followed in order to ensure harmony and consistency in the formulation of Yoruba terms, thus enhancing prospects of transparency and acceptability of the created terms as well as convenience in pedagogy.

4. Some of the problems identified and our efforts at addressing them
The development of the Yoruba language in the direction of technological development has suffered over the years due to the use of English as the language of technology in Nigeria. Hence, many of the English terms used in MS Vista do not have corresponding Yoruba terms. It was necessary therefore to develop terms that can convey the meanings of the original English terms to the Yoruba user of MS Vista.

There were a number of challenges encountered in the project. One of the main challenges was that due to the use of English as the main language of education in Nigeria, most Yoruba computer scientists do not have sufficient competency in the Yoruba language to facilitate enough knowledge to produce appropriate terminologies. On the other hand, most Yoruba linguists do not have sufficient knowledge of computing to understand the

9

proper contexts of use of many words in the domain of computing. Hence, a linguist translated a data field in the sense of a playing field for data and a computer scientist translated stand-by mode as *moodu sitandibai*.

Another major challenge was the influence of Yoruba news casters who in course of their daily work have to translate English terms encountered in news items to Yoruba on the fly. These newscasters sometimes use inappropriate terms which usually stick. Hence, there was the challenge of changing such prevailing terms without causing confusion in the minds of computer users.

One of such is the use of the term *Èrọ ayára bí àṣá* for the computer in Yoruba news on radio and television. *Èrọ ayára bí àṣá* describes a machine that is as fast as the hawk. The order of speeds in computing and the flight of the hawk are different and so this term may not be as appropriate as *kọ̀mpútà* which is derived by borrowing and phonologising the original; computer.

5. Experience and suggestions on extending our work to other African languages

There were many other challenges faced during the project. Some of these stem from the low level of infrastructure development in Nigeria. Such facilities as constant electric power supply and Internet connectivity that are usually taken for granted in other parts of the world were not always available. Efforts to make up for these deficiencies resulted in increased cost of the project.

After the launching of the Yoruba version of Windows Vista, user response did not reflect the anticipated level of enthusiasm from Yoruba computer users. First, users might not have been sufficiently aware that the Yoruba component is available due to low publicity of the project. Second, most users who are aware of the product seem to prefer to continue working in the English Language environment in which they are already familiar.

User reviews were mixed. They vary from excitement through ambivalence to outright condemnation of the project.

Also of importance is the need to approach such localisation project from an interdisciplinary point of view. The project required expertise from three key areas: (i) linguistics, (ii) language technology, (iii) computer science.

5. Conclusion

In conclusion, the localisation project is a worthwhile project. The project is both viable and important. The curiosity aroused by the project has attracted some desired attention to the Yoruba language in and around West Africa. The documentation that accompanies the work will also serve as a valuable archival and research tool, if not now, in the future. Based on the experience in this project, we consider expanding the work to other software products such as word processors and spread sheets. The idea that anybody who wants to use the computer should be able to speak English is not tenable. There is no reason whatsoever why any Yoruba that need to get money out of an ATM machine must learn to speak English.

References

Oladele Awobuluyi. 2008. *Eko Iseda-Oro Yoruba*. Montem Paperbacks, Akure.

Kola Owolabi. 1989. *Ijinle Itupale Ede Yoruba: Fonoloji ati fonetiki*. IUP, Ibadan.

Friedel Wolff. 2011. *Effective Change Through Localisation*. Translate.org.za.