

Sustainability of Research and Development: A case of Successful Technology Transfer in Spoken Language Technology

Mucemi Gakuru
University of Nairobi
P.O. Box 30197, 00100
Nairobi, KENYA
mucemi@teknobyte.co.ke
+254-722-306712

Roger C F Tucker
Outside Echo Ltd
Beracah House, Gloucester Rd, Tutshill
Chepstow NP16 7DH, U.K.
roger@outsideecho.com
+44-7963-106173

ABSTRACT

We share our experience on how to establish sustainable research and development based on successful collaborative research between the Local Language Speech Technology Initiative of UK and Kenyan partner - Teknobyte Ltd. Starting in 2003 when we embarked on spoken language technology development, technology and expertise were successfully transferred to the Kenyan partners, culminating in the launch in April 2008 of the National Farmers Information Service (NAFIS), a voice service accessed over the phone which offers a wide range of information in Kiswahili or Kenyan English to supplement the existing agricultural extension services.

General Terms

Management, Human Factors, Languages

Keywords

Spoken Language Technology, IVR

1. INTRODUCTION

Most ICT projects in the developing world have had a difficult history, particularly those with a specific developmental aim [1]. Implementation of these projects usually follows a model where a donor gives funding and expertise to develop and deploy a system. On completion of the system local personnel is trained on how to use it while relying on foreign expertise for support and maintenance. This model often fails to build in certain pertinent aspects that would ensure success in acceptability and therefore proper utilisation. For one, the system may not provide the appropriate solution as the developers do not capture all the important local issues owing to their limited experience in the local environment.

Secondly the project may fail to create local ownership which is key to acceptability that would ensure full

utilisation and hence sustainability. In order to succeed in developing and deploying ICT technologies based on Spoken Language Technology (SLT) a well thought out model for the developing world is even more crucial as the languages are clearly local and hence require a lot of local knowledge. Many challenges would need to be dealt with which are due mostly to the lack of speech data and linguistic resources, the skills shortage in the countries where the languages are spoken, and the general lack of funding for research in the poorer countries of the world. However we still chose Spoken Language Technologies because they offer a great potential in the developing world for making IT more accessible for those without text or computer literacy, enabling widespread access to information on health, markets, agriculture, weather, microfinance etc.

Compared to the use of SLT in the developed world, the potential has yet to be realized in the developing world. There are two key factors that help explain the difficulty – the lack of success of ICT for Development projects generally, and the specific problem of SLT for local languages in the developing world. One way to address these challenges is to nurture local expertise and ownership thereby creating a sustainable base for continued development and resource mobilisation. We therefore set from the beginning to address the key issue of technology transfer and ownership and found out that by allowing the partners in the developing world to take ownership of the technology development process early on, they had the motivation and maturity to see it through to successful deployment.

In this paper we therefore describe how local expertise in Kenya was developed thereby leading to development and deployment of an agricultural information system.

2. INITIAL PARTNERSHIP TO DEVELOP A KISWAHILI TEXT TO SPEECH SYSTEM

Collaborative research was initially set up between the University of Nairobi and the Local Language Speech Technology Initiative [2], an organisation started in 2003 with the aim of creating communities of interest in SLT for Development throughout the developing world. University of Nairobi was approached by LLSTI to undertake the development of a Kiswahili Text to Speech (TTS) system using the open source engine Festival [3]. Sponsorship for the development came from Oneworld, who in turn had received funding from the Vodaphone Foundation to set up a mobile-based social enterprise in Kenya, which subsequently was launched as Mobile4Good.

The initial plan had been for Mobile4Good to incorporate voice into their SMS-based business model, but in the end there was simply not enough financial resource to make this happen. We then searched for sponsorship to re-engineer the TTS to production standards and optimize the execution speed, so it could be made widely available to application designers, and also used in screen-readers for the blind. We were unable to secure any funding and it looked like further development would come to a halt. However the local developers of the Kiswahili system had realised its potential and transferred the development to a Kenyan company Teknobyte.

In addition, it was clear to the local developers that a Kenyan English voice could also have good potential. Most of the English TTS voices available have American or British accents. Even though such voices are fairly intelligible to many Kenyans, a greater majority often face difficulties in picking up entire phrases due to differences in intonation and pronunciation, especially of the large number of borrowed words from local languages. The development of a TTS based on the Kenyan English dialect was therefore deemed necessary in order to produce a system whose output would be more natural to the Kenyan population and hence more easily understood. Teknobyte therefore developed a demonstration voice from an open-source US English TTS by simply rerecording the US English unit-selection database and performing automatic labelling using the Hidden Markov Model Toolkit (HTK). This English voice would be used to test whether there was a good case for further developments.

Until LLSTI started, SLT for Southern languages were most often developed in a Northern lab, with a speaker of the language temporarily joining a team of experts to annotate data and advise on language characteristics. This approach has a number of drawbacks, not least that SLT requires ongoing development and often needs tailoring to a specific application. Unless the team lives and works in the country where the systems will be deployed, it is difficult to do this

effectively. It can be seen that at this early stage in Mid-2005 the creation of local ownership of the project was beginning to bear fruits.

3. THE BANANA PILOT SERVICE

It was now necessary to create awareness of the importance of the SLT systems so as to secure further sponsorship or a business case for further development. We therefore choose to build a pilot application using the Kiswahili TTS that would demonstrate to the development community the potential of a voice-based information system. Kiswahili is the first language of almost 800,000 people and over 100 million rural people in Eastern Africa are second-language users. We identified farming as the sector where it would be easiest to make a direct impact on Kenyan society, since as much as 75% of the population is involved in farming in some way.

We worked in partnership with the National Agriculture and Livestock Extension Programme (NALEP) to provide a phone-based information system for growing bananas. Bananas were chosen because of the demand for knowledge of this particular crop, and the system covers all aspects from deciding whether or not to grow bananas through to harvesting, pest control and market information. In addition bananas are grown in all arable parts of the country, which meant that the system would find use in many parts of the country thereby giving a chance for testing various types of data. We partnered with a small IVR company, SpeechNet, to host the Banana Line.

The content originated in English and was translated into Kiswahili by hand. On phoning the system, the user is given the choice of either language. The English TTS used was a British English female voice (Nina) – the Kenyan English demonstration voice was too buggy to be used at this stage. The pilot [4] included a formal test on a carefully selected group of 10 farmers in the Kirinyaga district of Kenya. 7 out of the 10 chose to listen in English, but then struggled with the British accent.

The farmers were also played some speech from the demonstration Kenyan English TTS system, and found it even clearer than the Kiswahili which confirmed that there was a good case for further development of a Kenyan English Text to Speech System. Furthermore, the pilot also gave a comprehensive picture of what a real service should be and that it should include many different crops, simple update of information among others. After the pilot, in late 2006, Teknobyte took over full leadership of the ongoing developments.

4. DEVELOPMENT OF A KENYAN ENGLISH TTS

Even before the pilot had confirmed the need for a Kenyan English TTS, Teknobyte had set out in Mid-2005 to improve on the demonstration system by creating a lexicon to be used in synthesis using the Festival Engine. With limited resources and little linguistic work done on Kenyan English this would be an onerous task that would take a long time to complete. Firstly it would be to define linguistic features such as a phoneset that would capture all the sounds in the dialect including those of the borrowed words, and then testing the phoneset for effectiveness. Secondly it would be to secure and develop tools that would assist in creating a sizeable lexicon that would yield further improvement of the voices. Finally it would be to generate a comprehensive list of borrowed words and provide for their pronunciation. A lexicon of 60,000 words including 6,000 borrowed words was eventually developed.

The Speech database used to develop the demonstration voice was clearly not suitable for a Kenyan voice and would certainly not be phonetically balanced as it was based on the Oxford Advanced Learners Dictionary. A speech database for the Kenyan English was necessary and this was done by first collecting a large text corpus from various sources such as novels, newspaper articles and written speeches, which were rich in local names. The speech database was developed by selecting 1500-sentence phonetically balanced sentences from the text corpus, in the same way as the Kiswahili TTS system earlier on [5].

With a Kenyan English lexicon and a corresponding speech database, the critical tools were now complete for further development of Kenyan English voices. The first voice to fully utilise these tools was made from a professional Kenyan newscaster, Frank Muiruri and used in the National Farmers' Information Service discussed below.

5. DEVELOPMENT OF A COMPREHENSIVE VOICE SERVICE FOR KENYAN FARMERS

The lessons learnt from the Banana pilot service and the need to harness ICT Technology as a cost effective supplement to enhance the reach of extension services resulted in an idea to develop a comprehensive voice service (automated IVR) for providing agricultural extension information through telephony. The service was anticipated to have a big impact given the importance of agriculture in Kenya's economy, which is mostly done by 4.5 million small-scale farmers.

This automated service named National Farmer's Information Service - NAFIS was built by improving on the banana pilot service and includes more enterprises as well as features. The service is updated through the Web so as to

enable updates by field extension officers, with the information so updated being accessed through the phone, while being available on the website. It uses audio speech generated automatically in both Kenyan English and Kiswahili by the Text-to-Speech Systems developed by Teknobyte.

6. CONCLUSION

The successful development and deployment of Nafis was carried out by Teknobyte based on the experience gained from the Kiswahili TTS and the Agricultural Information pilot, which had been carried out with LLSTI assistance. The five year period taken for development of an information system using SLT for Kenya is long especially considering the development is not yet fully complete. This is because the development has involved imparting skills and ownership in the local developers, a process which is inevitably slow. The faster alternative approach would have been to get a large donor or government department to simply buy lots of expertise from the developed world, put together the system, install it and hand it over to the government of Kenya. However, the experience of many years of ICT for Development suggests that this approach is unlikely to have worked. Instead, we have local expertise and ownership, and every possibility for a successful, locally maintained, appropriate information system for Africa.

Clearly if local expertise and ownership was not courted right from the beginning, the development process would have terminated on completion of the first Kiswahili Text to Speech system, when no more funding could be secured. However, the local developers took up ownership and pushed the project to the current position. Further improvement and maintenance of NAFIS can now be carried out cost-effectively by Teknobyte thereby showing the benefits of technology transfer. We believe that this case clearly demonstrates the approach other research and development projects could adopt so as to ensure maximum utilisation and sustainability of ICT projects.

7. ACKNOWLEDGEMENTS

We would like to thank Ksenia Shalnova for her support during the development of the Kiswahili TTS system, Pernilla Näsforss for the work on the Banana Information line and John Kimani of NALEP for support in NAFIS.

8. REFERENCES

1. R Heeks, "Failure, Success and Improvisation of Information Systems Projects in Developing Countries", 2002. Available from: www.sed.manchester.ac.uk/idpm/research/publications/wp/di
2. R Tucker and K Shalanova, "Supporting the Creation of TTS for Local Language Voice Information Systems", Interspeech 2005, Lisbon, Portugal, pp. 453-456, 2005
3. M Gakuru et al, "Development of a Kiswahili Text To Speech System", Interspeech 2005, Lisbon, Portugal, pp 1481-1484, 2005
4. P Näsfors. "Efficient Voice Information Services for Developing Countries", Masters Thesis, Department of Information Technology, Uppsala University, Sweden, 2007. Also available from www.llsti.org/documents.htm
5. M Gakuru et al, "Design of Speech Data Base for Unit-Selection in Kiswahili TTS", E-Tech2004, Kenya. Also available from www.llsti.org/documents.htm

9. BIOGRAPHIES

Roger Tucker is the CEO of Outside Echo Ltd, a not-for-profit company dedicated to enabling people throughout the world to benefit from ICT through the medium of audio, and a visiting fellow in the Computer Science department at Bristol University, U.K.. He received a B.Sc. degree in mathematics and electronics from Keele University, UK in 1979 and a PhD in the Transmission of Speech over Packet Switched Networks from Aston University, UK in 1984.

Since then he has been a lecturer at Aston University, a consultant with Enigma Ltd, and a technical lead at Hewlett-Packard Labs Bristol until he founded Outside Echo in 2003.

From 2003-2006 he pioneered the Local Language Speech Technology Initiative described in this paper, and has subsequently taken a consultative role in a number of EPSRC projects relating to technology for development. His key research interest is in how spoken language technology can benefit those excluded from mainstream computing.

Mucemi Gakuru is a Senior Lecturer in the Electrical Engineering department of the University of Nairobi and the Managing Director of Teknobyte Ltd, a technology company, he founded in 1994, to develop and promote technology-based solutions in the greater Eastern Africa region. He graduated with a First Class honours B.Sc. degree in Electrical from University of Nairobi, in 1987 and a PhD in the application of Finite-Element method in design of Microwave Devices from Cambridge University, U.K. in 1991.

Since 2003 he has been doing research and development work on Spoken Language Technologies having been introduced and supported initially by LLSTI and later through personal initiative at Teknobyte Ltd. He led the team that developed the first Kiswahili Text to Speech system at the University of Nairobi and later developed a Text to Speech system in English dialect as spoken in Kenya, which he branded Kenyan English. These developments culminated into the development of the Farmers information service, which is the subject of this publication.