DIAL-UP KNOWLEDGE FOR THE DEVELOPING WORLD

by Etienne Barnard

For communities that don't yet have access to the information highway, the telephone can help bridge the digital divide.

Access to information is a key element of individual, group or national success. The truth of this statement has been emphasised by the rapid ascendancy of information technology during the past 20 years. It is now widely understood that the "digital divide" harbours tremendous dangers, by denying many members of society access to modern information sources and thereby excluding them from the potential economic and other benefits that such access presents.

Many non-governmental organisations (NGOs) and other institutions active in the developing world are aware that they can improve citizens' quality of life significantly by providing timely and relevant information to all. For example, material related to health and nutrition, to education and training, and to employment can make a real difference to people's lives. However, the dissemination of such information can be an arduous task – it is no use publishing the latest immunisation advisories on a web site if members of your intended audience do not have access to computers! Conventional, paper-based publication is often too expensive and time-consuming, and less useful for illiterate citizens who have to rely on someone else to interpret the information for them.

The widespread availability of general telephone service makes telephone-based services an attractive option for providing information. A further advantage of telephone-based services is the relatively low levels of infrastructure and user sophistication at which such services can operate. Useful services can be delivered to citizens equipped with nothing but a normal telephone (mobile or fixed-line), and requiring no more than the ability to understand and respond to spoken commands. Such a verbal interface is appropriate from a cultural perspective in many developing countries, where a strong oral tradition exists amongst a pre-literate or semi-literate population. Thus, telephone-based services place a direct focus on the role of human language technologies (HLT) in the developing world, and should assist in drawing resources for the development of such resources.

We therefore envisage services that citizens will access by calling a central toll-free number; they will be given a menu of choices, and be able to obtain or provide information over the telephone. The exact mode of information input raises a number of intriguing research issues – for example, will technologically unsophisticated users fare better using speech recognition or key presses as input to an automated system? Or, how should prompting be structured to assist a user who has no mental model of the functioning of a computer system? It is likely that good answers to such questions will assist us in developing interfaces that can help bridge the digital divide.

The widespread application of telephone-based services will require satisfactory answers to these research issues, as well as low-cost, easy-to-use systems for the development and deployment of these services. Researchers in the CSIR*/University of Pretoria Human Language Technologies research group are studying many of these topics, and collaborating with other members of the CSIR Information Technology Centre to develop the OpenPhone platform – an open, extensible platform aimed at enabling widespread use of telephone-based services in the developing world.

The components of OpenPhone

OpenPhone combines human language technologies, telephony components and workflow-management software in a way that enables organisations such as NGOs to make information available over the telephone network. The tool kit contains components for authoring as well as running "interactive voice response" (IVR) services in several languages. In comparison with commercial IVR tool kits, OpenPhone is distinguished by (a) its open nature; (b) its focus on multilingual services; (c) ease of use (the system must be useable by developers who are not IVR specialists); and (d) the provision of text-to-speech and speech-recognition capabilities as a built-in feature of the tool kit.

A crucial part of the tool kit is training material designed to help general software developers set up such telephony services rapidly and effectively. The training material includes best-practices guidelines and templates for multilingual systems, as well as systems aimed at illiterate users. It is integrated with a workflow-management system to assist developers in creating useable IVR systems for non-conventional audiences.

Since OpenPhone embraces the open-source philosophy, it builds on a number of existing open-source components. These components include the Asterisk platform for basic telephony services and the locally developed CoEfficient platform, supporting workflow functionality.

* Council for Scientific and Industrial Research

Developing local language technologies
Human language technology involves the development of software components that can be used for natural communication between humans and computers. (HLT actually spans a wider arc; it also covers, for example, technology-enhanced human-to-human communication and language-based information access).

In the developing world, the major challenge is to develop algorithms, tools, and systems that are relevant to the cultural background and technological capabilities of highly diverse users – in languages for which there are often not sufficient linguistic resources. For example, a system that can understand speech in a local language such as Setswana would be of great use in OpenPhone, but the development of such a system requires large databases of text and speech, which are not freely available. Components such as pronunciation dictionaries – which provide the spoken form of a word from its written form – are also crucial in HLT.

The CSIR/UP HLT research group is therefore active in the development of a variety of algorithms and tools that support the creation of relevant resources, in addition to the development of technology components. For example, a tool called the DictionaryMaker was developed to allow a speaker fluent in a target language to develop a pronunciation dictionary in an effective way. The tool guides the user through the dictionary creation process, which can be completed in a fraction of the time it usually takes to create a pronunciation dictionary. Along with the dictionary, a related set of grapheme-to-phoneme (G2P) rules is created automatically. This tool forms the basis for the efficient development of speech-processing systems, and uses enhanced machine-learning techniques that were developed by Marelie Davel as part of her PhD studies (see "further reading").

The DictionaryMaker supports technologies such as automatic speech recognition, which gives computer-based systems the ability to understand human speech. Although people have dreamt of computers that understand natural speech for a long time, that goal is still some way off. Currently, the best that we can do is to understand carefully spoken speech on a limited set of topics by any speaker (speaker-independent speech recognition), or speech by one speaker only on a wider set of topics (speaker-dependent speech recognition).

Despite these limitations, speech recognition is already of great practical value. People who are not able to type due to a disability, can for example, use speaker-dependent recognition to operate computers by voice instead of a keyboard. Speaker-independent recognition forms the core of computer systems that interact with people over the telephone, giving them the ability to obtain information or perform transactions through spoken interaction.

The CSIR/UP HLT research group focuses on speaker-independent recognition, principally for telephone-based applications. In collaboration with the University of the North, for example, we have developed an initial version of a speaker-independent recogniser for the Sepedi dialect of Northern Sotho. This system, which forms part of the PhD studies of Jonas Manamela, recognises words in terms of their phonetic decomposition, and was developed using the open-source HTK (Hidden Markov Model) tool kit from Cambridge University.

Prospects
The need for a platform such as OpenPhone, and the human language technologies that support it, has become increasingly clear in the past few years. In addition, the development of tools such as DictionaryMaker, and systems such as the Sepedi recogniser, has shown that these technologies can be developed and supported in the South African environment. Several challenges remain to be addressed.

For one thing, it is still not known how successfully users with limited technological and general literacy can use HLTs to obtain information and perform transactions (although initial experiments in our group have provided encouraging results). Another challenge is one of scale – to serve the population of South Africa - at least the 11 official languages must be supported. And this is still a relatively small task when considering the challenges in countries such as Nigeria or the Congo, where hundreds of languages are commonly used! These challenges will continue to produce interesting technical and intellectual challenges for years to come.

Further reading

More information on the HLT research group and links to the papers listed below are available at www.csir.co.za/hlt

For more information on DictionaryMaker, see:
Davel, M. and Barnard, E., 2004, The Efficient Generation of Pronunciation Dictionaries: Machine Learning Factors during Bootstrapping in Proceedings of the 8th International Conference on Spoken Language Processing, Korea.

Telephone-based systems for developing world applications are discussed in:
Barnard, E., Cloete, L., and Patel, H., 2004, Language and Technology Literacy Barriers to Accessing Government Services, Lecture notes in computer science, Issue 2739, 2003, pp 37-42.

***Professor Barnard**, Department of Electrical, Electronic and Computer Engineering, University of Pretoria.*

ebarnard@up.ac.za