Issue 40
December 2007

ICT Update

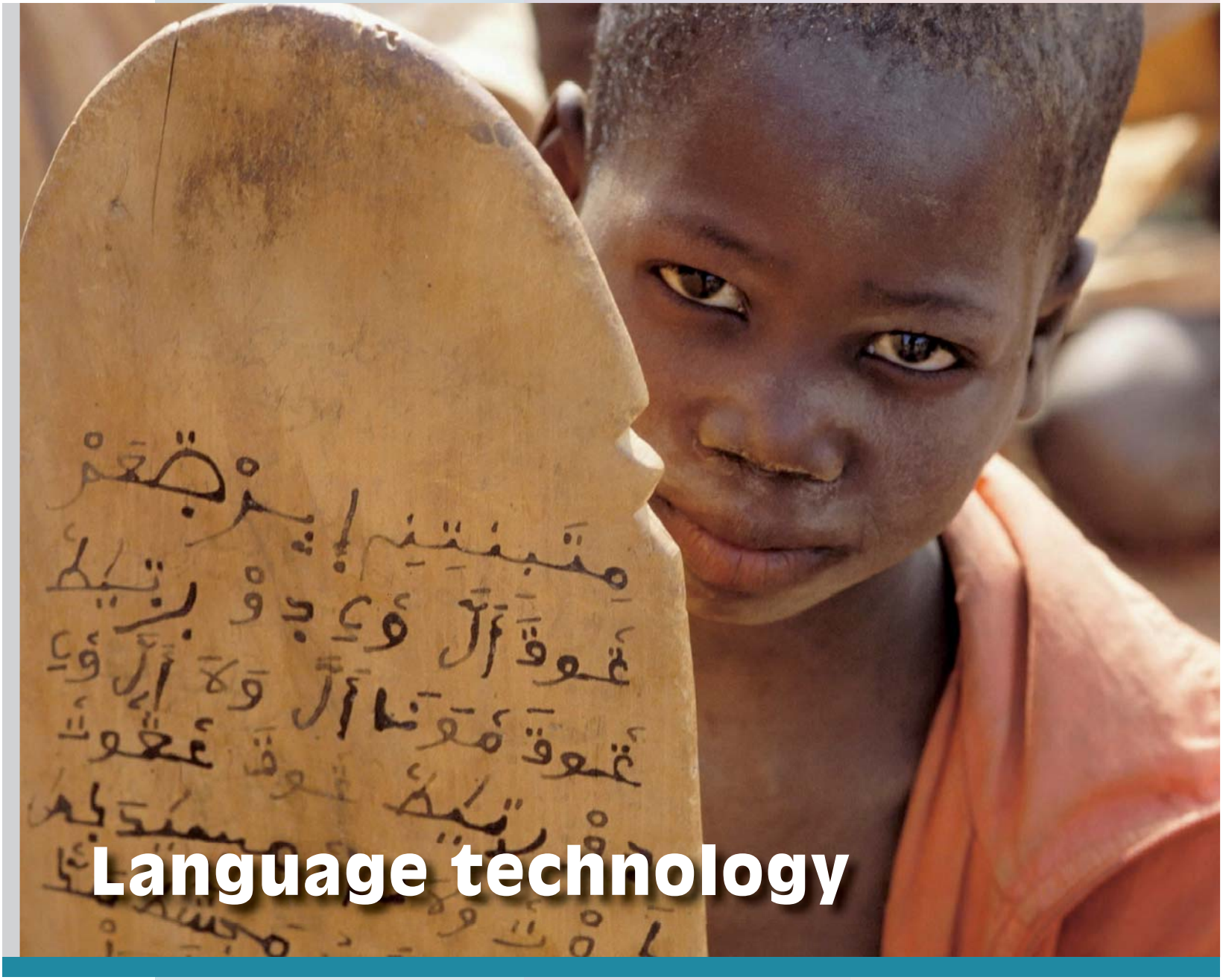
a current awareness bulletin for ACP agriculture

CTA
http://ictupdate.cta.int

A system that turns text to speech provides information to **Kenyan** farmers

Traditional **Nigerian** music inspires speech recognition technology

A **pan-African** project makes technology more accessible to local communities



Language technology

Contents

ICT Update

ICT Update issue 40, December 2007. *ICT Update* is a bimonthly printed bulletin with an accompanying web magazine (http://ictupdate.cta.int) and email newsletter. Each issue of *ICT Update* focuses on a specific theme relevant to ICTs for agricultural and rural development in African, Caribbean and Pacific (ACP) countries, and includes feature articles and annotated links to related web resources and projects. The next issue will be available in February.

Publisher: CTA Technical Centre for Agricultural and Rural Cooperation (ACP-EU). CTA is an institution of the ACP Group of States and the EU, in the framework of the Cotonou Agreement and is financed by the EU. Postbus 380, 6700 AJ Wageningen, the Netherlands. (www.cta.int)

Production and content management: Contactivity bv, Stationsweg 28, 2312 AV Leiden, the Netherlands. (www.contactivity.com)

Coordinating editor: Rutger Engelhard / Editor: Jim Dempsey / Copyediting: Valerie Jones / Magazine design: Frissewind / Layout: Robert-jan Cornet / Translation: Patrice Deladrier / Cover Photo: Jorgen Schytte/ Still Pictures / Editorial advisory committee: Peter Ballantyne, Oumy Ndiaye, Dorothy Okello, Kevin Painting

http://ictupdate.cta.int

Editorial

Coming to an understanding

For centuries, humans have been using technology to communicate. From using drums to send messages to the next village, telegraph to reach neighbouring countries and radio to exchange words across oceans. Today, the internet, email and text messages serve the same purpose. They enable people to communicate over large distances, but only if they understand each other's language. And only if they understand the technology. However, numerous efforts are now under way to make technology available in more languages, giving more people the opportunity to use it and adapt it to their own local needs.

These initiatives include giving access to the many people who cannot read or write. More than 40% of the populations of ACP countries are illiterate, and the rates are even higher in rural areas where there is also a problem of access to technology. Here, the difficulty is delivering relevant information – which may already exist in printed publications or on the web – to people in remote areas and in a form that most can understand.

A system that automatically reads out prepared text, on demand, could solve this. This type of 'text-to-speech' (TTS) technology has already been developed for many languages, particularly for the major European and Asian languages. But a few organizations are now investing time and money to produce TTS for other languages. Our feature article describes how such a system was developed for Kiswahili to give advice to farmers in Kenya on all aspects of growing bananas. The Local Language Speech Technology Initiative expects that the work will be replicated in other regions and that, eventually, the information could be delivered directly, and cost effectively, to farmers via mobile phone.

Traditional solutions

But real communication isn't only one-way. In Nigeria, the African Language Technology Initiative (Alt-i) has developed a system that can recognize the Yoruba language which, when spoken, can be automatically transcribed to text. Among many other uses, such an automatic speech recognition system is particularly valuable when documenting traditional knowledge from village elders. In fact, the development of the Alt-i speech recognizer was inspired by the sounds of the traditional 'talking drums' of the local Yoruba culture.

This is an excellent example of using traditional knowledge to develop new technology which then, in turn, helps to preserve that knowledge. This approach is also a feature of the work of the Hul'qumi'num Treaty Group (HTG) in Canada. Fewer than 100 people are still fluent in the Hul'q'umi'num' language, but the group is now using video to record those last few speakers and preserve the language for future generations. HTG hope that their methods will be reproduced and adapted elsewhere to help save some of the other 3000 or so languages that could disappear in the next 100 years.

But while some languages are dying out, there are growing attempts to increase the use of other minor languages in technology. The Pan-African Localisation project is looking at ways in which websites, computer interfaces and software can be adapted or produced for African languages. At the moment, most computer operating systems and program menus are displayed in only English or French, which can be disconcerting to many people. Presenting technology in more familiar languages would promote involvement and increase the number of people using computers and the internet in the future.

Science fiction writers predicted long ago that we would have computers that could talk and understand our every word. While that hasn't quite happened yet, it is certainly encouraging to know that work is being done to make sure that more people are able to use new technology. Even those who haven't learned to read and write can get involved. And, most important, they can do so in the languages they best understand. ■

Adama Samassekou (asamassekou@acalan.org) is executive secretary of the African Academy of Languages (www.acalan.org) and chairperson of the Maaya network (http://maayajo.org)

Linguistic diversity

Linguistic diversity in human societies is the equivalent of biodiversity in the natural world. It is highly prized because it is the key to the survival of many of the world's cultures, especially in the new digital era which is currently growing so rapidly. At the moment, the language most frequently used on the web is English (45% of all pages). French accounts for less than 5%, and Spanish only 4.6% (figures from http://funredes.org). It would seem though, that cultural diversity is on the increase because in 1998 English accounted for 75% of all web content. However, if we disregard the 10 or 12 languages most commonly used on the internet (which make up 90% of the total), the lack of diversity is a worrying phenomenon, given that more than 6000 languages are spoken worldwide.

The representation of African languages, such as Kiswahili, Hausa, Yoruba, Mandingue and Peul, in computer technology is so marginal that they do not feature in the survey findings. And while universal access to the internet was a goal of the World Summit on the Information Society (WSIS) in 2005, this should not be taken as merely the ability to connect to the web. Equally important, is the ability of citizens to access content in their own language and also to produce some of that content themselves.

There is a real risk that the development of cyberspace may hasten the disappearance of the least spoken languages. These languages lack standardized terminology that can be applied to information technology, at least as far as the written forms of the languages are concerned. The sheer difficulty of introducing vocabulary from these languages into computer technology is often considered impractical. This is the case with the African languages.

VANESSA VICK/THE NEW YORK TIMES/ HH

If national and international decision makers were to devise strategies to safeguard these languages, and to promote them among the public through widespread literacy training, then technology could offer a number of opportunities for putting them into cyberspace. The development of Unicode Standard, an encoding system recognized by all computer operating systems, programs and in all languages, is a promising means of computerizing languages and integrating them into office and management software. Although the makers of open source software were first on the scene to produce products for the various linguistic communities, it is a welcome fact that companies such as Microsoft are now also beginning to take an interest in this field. Some of these new ventures are, of course, commercial in nature, whereas the guiding motive for the governments should be the national interest, as well as the needs of the general public. There is, moreover, scope for creating international domain names, allowing the use of alphabets other than English to be recognized on the internet. That would be genuine progress. Fortunately, at the Internet Governance Forum (IGF), a by-product of the WSIS, the issue of international domain names is often a key subject for discussion.

Digital technology can also be put to use to develop mutual understanding between languages. The development of automatic translation and voice recognition tools is especially noteworthy, and was discussed at the meeting of the Dynamic Coalition for Linguistic Diversity during the IGF conference in Rio de Janeiro, Brazil, in November 2007. The participants said it was regrettable that public authorities and software developers were so inactive in promoting machine translation, even though this is the key to ensuring universal internet access. Of course, developing language technology and creating text corpuses and dictionaries are costly affairs, but governments should be playing a greater role in this area.

We may hope that 2008, proclaimed the 'International Year of Languages' by the General Assembly of the United Nations, will prove to be a turning point for the development of the marginalized languages in every country of the world. ■

Feature

Accessing information by voice

Getting up to date agricultural information to rural farmers has always been a challange, especially in areas with low literacy rates. A system to read out text, available via mobile phone, could help reach many more communities.

A Kenyan farmer growing bananas for the first time is excited when the bunches of fruit appear, ready to be harvested. When he discussed expanding into this new crop, the local agricultural extension officer gave him a telephone number to call. The phone line would provide all the information needed to guide him through each step of the process. As this first batch ripens, the farmer needs advice on how best to harvest and look after the bananas once cut. He decides to call the number. A mistake now could cost him dearly.

The farmer calls using his mobile phone, and a voice offers him the choice of listening in Kiswahili or English. Although neither is his first language, he is more confident in Kiswahili, so he chooses that. After a few minutes the farmer has the information he needs, and he even recognizes the voice on the line as Ken Walibora, one of Kenya's best known TV anchormen.

He does not know that Ken Walibora has never spoken a single word over the Banana Information Line; the information was generated automatically from text thanks to a 'text-to-speech' (TTS) system developed by a team led by Dr Mucemi Gakuru at the University of Nairobi. The advantage of this service is that the information can be kept up to date simply by editing web pages. All Ken Walibora had to do was to spend about 45 minutes reading aloud some carefully selected sentences. This is the power of speech technology.

Local input

The Banana Information Line is a project of the Local Language Speech Technology Initiative (LLSTI), produced in partnership with National Agriculture and Livestock Extension Programme (NALEP) of the Kenyan Ministry of Agriculture. It ran as a pilot for several months in 2006, investigating the use and possible functionality of a voice information service for Kenyan agriculture.

LLSTI first looked for a partner in East Africa to help develop a Kiswahili text-to-speech system, and contacted Dr Gakuru in 2004. Interestingly, until then, the team at the University of Nairobi knew little about speech technology. In the following months, LLSTI supplied the tools, training and expertise to enable Dr Gakuru to develop the text-to-speech system used in the Banana Information Line.

This project is typical of the work done by LLSTI. The organization began in 2003 as a global initiative led by Outside Echo, a UK not-for-profit organization which facilitates audio access to information, and works together with partners from India, South Africa, Kenya and Nigeria. LLSTI provides the support needed for a team with no prior knowledge of speech technology, to produce usable, natural-sounding voices with a TTS system. The main requirements are a linguist, a software engineer, and a motivated team leader who preferably is an engineer as well. Each team is normally based in the area where the language is spoken, and part of a university or research institute. This is in contrast to the commercial development of TTS, where a speaker of the language temporarily joins a team of experts in a company's lab in Europe or the USA.

Why is it important to encourage local involvement in the development of speech and language technology? One reason is motivation – people feel strongly about their own language. Then there is maintenance – language technology needs ongoing development to overcome problems that may arise during its use. Relevance is also very important – academics need to be encouraged to focus on problems that are most relevant to their own communities.

This last point is an important one. Academics are assessed on the basis of their publications in international journals and conference presentations, which all too often focus on small, barely significant, contributions to mainstream research. LLSTI therefore encourages a rigorous approach to local language technology development so that academics can publish their work internationally. So far, the partner organizations have been very successful, and their work in the local languages has generated several quality publications.

LLSTI also prefers to make the software code of a new language system, or at least the initial prototype, available for anyone to use, including for commercial purposes, through a Berkeley Software Development (BSD) open source licence. This makes it easier for the technology to be used, and also for new researchers to pick it up and enhance it.

The Meraka Institute in South Africa, one of the early partners of LLSTI, has gone on to develop a number of new local language TTS and related applications. In Botswana, for example, the Institute has set up an English/Setswana Aids Caregivers' Helpline, which will be piloted in early 2008. Indeed, the Meraka Institute itself has become a centre of expertise in speech technology, and now offers training and support for researchers from as far away as Nigeria. The Meraka team has also extended the original project to encompass automatic speech recognition (ASR), with the long-term goal of perfecting automated translation. The system they are currently working on, called 'Lwazi', is an ambitious phone-based, speech-driven information system commissioned by the South African Department of Arts and Culture. With Lwazi, citizens will be able to access government information and services in any of South Africa's 11 official languages, using either landline or mobile phones.

Dr Roger Tucker (roger@outsideecho.com) is director of the Local Language Speech Technology Initiative (www.llsti.org)

Difficult languages

All of this may sound easy, but a good TTS system requires some extremely clever software. It requires very specific language knowledge, a lot of hand-annotated text and audio data, and skilled engineering judgements that are specific to the particular language. The ultimate aim is to develop a system with a voice that sounds just like a human reader, already achieved for many European languages. However, a TTS system is good enough to be used as long as any unnaturalness does not significantly affect intelligibility. It turns out that this is easier for some languages than for others.

For instance, if the script does not include vowels, as in Arabic, how can the system know how to pronounce the word? Or if the language has free (unpredictable) stresses, as in English, how can the system know which syllable of a word should be the strongest? And if a comprehensive pronunciation dictionary is part of the solution to these problems, how can that be built and used if the morphology of the language is complex and allows a large number of variations on any given word stem, as in Russian?

To understand these problems from the very start, the LLSTI project began by conducting a survey of 105 languages to identify all the script and language features in each case that can create complications for a TTS system. The team catalogued these features in a TTS-related multilingual database that would enable them to predict the issues that specific languages would raise. The results are summarized in the TTS development complexity scores (see table).

Difficulty of developing TTS systems in various languages (0=easy, 10=difficult)

Language	Basic TTS	Good TTS
Pashto	9	9.5
Arabic (classical)	7	8.5
Russian	6	9
Tibetan	6	7.5
isiZulu	6	8
Ibibio	5	7
Thai	5	8
English	4	6
Hindi	2	4
Welsh	1	4
Kiswahili	0	4
Tamil	0	2.5

The ideal language, from a TTS point of view, has a complexity score of 0. This is a language where the text to speech process can be defined in a straightforward set of rules that any linguist can write down from their existing knowledge of the language. In practice, such rules can never completely define the process, but there are some languages where they can produce a basic system – one in which phrasing, loan words, abbreviations and other such details may not be perfectly rendered, but the meaning is still quite intelligible.

In general terms, developing a TTS system involves the following steps:

- Defining the language characteristics: phone-set (i.e. the sounds used in the language), letter-to-sound rules, the rules of syllabification (the separation of words into syllables), etc.
- Selecting a set of phonetically balanced sentences, from a large database of phonetically transcribed texts that cover all the different phone combinations of the language in as few sentences as possible. This is an automatic process, but some compromise is always needed between the number of sentences and the coverage of rare phone combinations.
- Selecting a speaker. The choice of voice is the single most important decision in the development of a TTS system. This may appear to be counter-intuitive – after all, doesn't almost everyone speak clearly

Related resources

Try the Text To Speech (TTS) system used for the Banana Information Line (in Kiswahili)
➜ www.llsti.org/demos-interactive.htm.

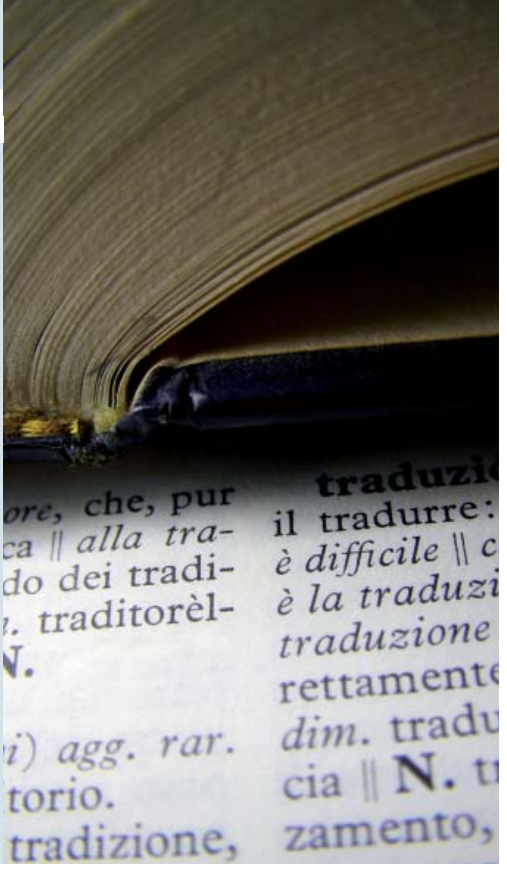
The TTS-related language database from LLSTI
➜ www.llsti.org/languages-database.htm

The Meraka Institute's Lwazi language project
➜ www.meraka.org.za/lwazi/

Examples of European TTS system
➜ www.nuance.com/realspeak

The Dictionary Maker
➜ http://dictionarymaker.sourceforge.net

Learning the Morphology of Complex Synthetic Languages
➜ www.cs.bris.ac.uk/Research/MachineLearning/morph



ANDREA MATONE / ALAMY

enough to be understood? But TTS requires a clear, precise rendition of every word in the database. The system constructs its output by combining small segments of words in the database, so each one has to be exactly right. Besides, the speech needs to be as intelligible as possible to start with, so that any deterioration in quality that may occur in the process of joining up these segments has the least possible impact.

- Recording the phonetically balanced sentences. For Kiswahili, there were about 400 sentences, which took around 45 minutes to record; for most other languages more spoken sentences are needed.
- Making phonetic annotations of the recordings by hand. Although this can be done automatically, any resulting errors can create problems for the TTS output.
- Compiling all the data into a TTS system using Festival, an open source software package developed at the University of Edinburgh.
- Testing the system.

If it is not possible to define the rules in the first step, there are a number of data-driven techniques that can be used instead. A data-driven technique takes a large database of annotated textual data, usually from public sources (like the internet) if they exist in that language, and attempts to derive 'rules' automatically. Annotating data by hand is laborious and time consuming, and over the years researchers have tried to minimize the proportion of data requiring accurate annotation, the ultimate aim being acceptable performance with no hand annotation at all. In South Africa, for instance, the Meraka Institute has developed a pronunciation dictionary builder that employs an iterative technique to build the entire dictionary with a minimum of effort.

For the majority of languages, even a basic TTS system requires morphological analysis (MA) of each word to derive its part of speech – a process that is language-specific and usually quite complicated, and is the major show-stopper for many languages. Consequently, LLSTI is currently involved in a major research project at the University of Bristol, UK, to develop a machine-learning MA system that can be applied to languages with very little known linguistic data.

Words talk

Back in Kenya, the Banana Information Line was formally evaluated with the help of a carefully selected group of 10 farmers in Kirinyaga district. The evaluation revealed some interesting problems. For example, seven out of the 10 participants chose to listen to the information in English, but then struggled with the British accent. Those who chose Kiswahili loved the voice, but then struggled with the formal Kiswahili grammar used in the translation. All of them said that they liked the voice system and that they preferred it to written material, but it was clear that the accent and translation issues would need to be fixed before it could be put to wider use. Dr Gakuru played some samples from a Kenyan English TTS under development, and the farmers found this even clearer than the original Kiswahili version. That TTS is now almost completed, and will be used in future.

During the pilot project and the evaluation, the LLSTI team consulted NALEP to build up a picture of what a phone-based agricultural information service should offer (see box). Such a service could allow farmers to get the specific information they need, whenever and wherever they need it, and in a language they can understand.

With the introduction of mobile data services across Africa, a future version of the information line becomes possible where the TTS system runs on the phone itself. Only the text data would be transferred, which the phone would then convert to speech. Not all farmers will have a mobile phone that can support this, but for those who do, it could provide a very attractive option, for a number of reasons. First, the cost of using the information line is much lower – mobile calls are still quite expensive in Kenya. Second, users can access the information they need using visual (pictorial/icon) menus, with a text search function for those who are happy to try it. Third, pictures and key numbers/words can be displayed along with the voice, making it easier for users to absorb the information and remember it afterwards.

All of this can be built with today's technology. The part that is still missing is a TTS system in the local language that people are comfortable using. Some of these languages present difficulties, and there is a lot of interesting and challenging work to do, but once these problems are resolved, the systems will be there to be used. The South African government is already putting resources into this; surely it is time for others to follow. ■

Reviving a language

With half of the world's languages likely to die out this century, a Canadian initiative shows how multimedia technology can regenerate interest.

In Canada, as in other former colonial nations, many indigenous languages are close to being lost forever. One of these is Hul'q'umi'num', a language spoken by a number of communities on southern Vancouver Island, including, among others, the Cowichan tribes, Chemainus First Nation, Penelakut Tribe, Lyackson First Nations, Halalt First Nation, and Lake Cowichan First Nation. As there are now fewer than 100 fluent Hul'q'umi'num' speakers remaining, the majority of whom are elderly, the survival of the language now lies at a critical juncture.

The Hul'qumi'num Treaty Group (HTG), an organization that acts of behalf these tribes and First Nations, and researchers from the Universities of Victoria and of British Columbia, have collaborated in a significant project to revitalize the language. The work of the project is guided by an advisory board of elders and together, they are encouraging the reincorporation of the Hul'q'umi'num' language into the everyday lives of the community members.

Digital multimedia technology, employing audio, visual and text-based resources, is visually appealing. This is particularly important for engaging elders and youth in both the creation and evaluation of language learning materials. Importantly, digital media such as digital video disks (DVDs) are becoming easier to produce and distribute, and so have enormous potential to contribute to the increased use of the Hul'q'umi'num' language in schools, language classes and communities.

Involvement

One effective technique employed in this project is participatory video (PV) which provides a training component aimed at building the skills of community members in film-making and producing DVDs. The elders' advisory board considered it important that, as well as language, the subject matter of the films and DVDs should contribute to the revitalization and strengthening of traditional cultural practices. The DVDs therefore, focus on several themes, including the documentation of traditional forms of public ceremonial speech, and skills such as cedar-bark weaving.

At the launch of the project in May 2004, the idea was to provide community members with highly interactive language-related information and other resources that they could use in their own homes. Most internet-based services involve both streaming and downloadable multimedia content, so users require expensive high-speed broadband internet connections. But most Hul'q'umi'num' community members live on reserve lands where access to broadband services is limited. At the time, it was estimated that only 10–15% of community members had a high-speed internet connection at home, whereas 90% had access to a DVD player. Under these circumstances, the decision to opt for DVD rather than web-based technology was an easy one to take.

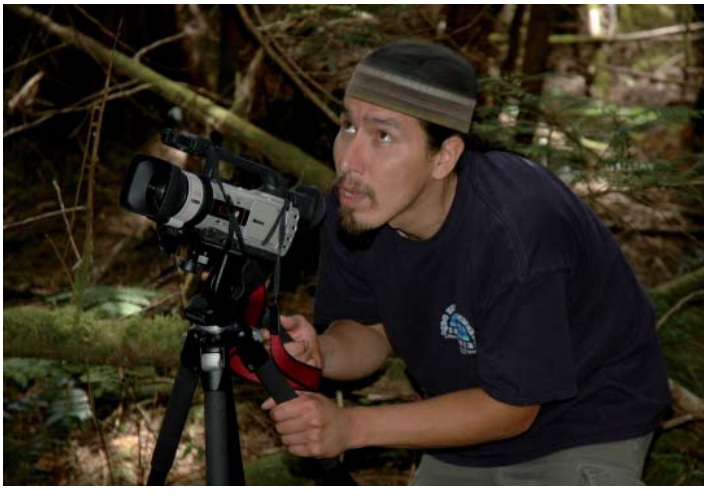
The DVDs were designed to give a level of user interactivity that is more commonly found with web-based technologies. The content includes language drills, word exercises and other materials, as determined by each user. The interactive DVD menu gives the user the ability to be selective about the type of information they access and receive. This element of choice also increases the value of the DVDs as people are unlikely to access all the content in one sitting but will go back to the disc and reuse it many times. The DVDs are, therefore, flexible and robust language-training tools.

New direction

Recently, it has been proposed to take the video footage from earlier film sessions and use it on the internet. This would help to make the language accessible to many more people. In particular, it would encourage the inputs of Hul'q'umi'num' youth who are increasingly using online social-computing technologies for entertainment and communication. The growth in the number of internet users has been helped, in part, by the rising number of people with a broadband connection in their home, as well as the availability of publicly accessible computer services.

In particular, the project is developing a blog based mash-up that combines text, photos, video and audio materials housed on Picasa Web Albums and YouTube. The blog can be easily updated to provide fresh material on a regular basis, encouraging users to return to the site to access new language material. Users can also comment on material which, in turn, could help to contribute to the development of an online Hul'q'umi'num' language learning community.

As technology becomes easier to access, hardware cheaper to buy and software easier to use, there has been a proliferation in the number and variety of special interest groups who are using digital media to document cultural information and communicate their views. And digital multimedia lends itself particularly well to language learning. Given the rapid rate at which many minor languages are being lost, it offers an unprecedented opportunity to help revitalize interest and, in the case of the Hul'q'umi'num', perhaps to help bring the language back from the brink of extinction. ■



Jon Corbett (jon.corbett@ubc.ca) is assistant professor at the University of British Columbia, Canada (http://web.ubc.ca/okanagan/ccgs/welcome.html). **Tim Kulchyski** (hemutth@gmail.com) is Hul'q'umi'num co-coordinator of Cowichan Tribes, British Columbia, Canada (www.cowichantribes.com). **Tom Hukari** (hukari@uvic.ca) is professor emeritus at the University of Victoria (http://web.uvic.ca/ling)

Hitting the right tone

The traditional Yoruba 'talking drums' of Nigeria provided inspiration for a unique approach to speech recognition technology. The technique could be applied to other African languages and used in literacy programmes.

People tend to perform more efficiently and effectively when they work and operate in their own language system. This is why the nations that taught science and technology in their mother tongues had significantly more success in the industrial age. Those nations that continue to live with the social and cultural consequences of colonization, and teach science and technology in foreign languages, are still struggling to compete.

Dr Túndé Adégbólá (tadegbola@alt-i.org) is executive director of the African Languages Technology Initiative (www.alt-i.org) and associate lecturer at the Africa Regional Centre for Information Science, University of Ibadan, Nigeria (www.arcisng.

If language had such a great impact in the industrial age, it will have even more important consequences in the information age. This is the basis of the work of the African Languages Technology Initiative (Alt-i); developing ICT resources in African languages to help take African cultures into the knowledge era.

But that will not be an easy task. Fewer than 10, of the 2000 or so African languages, are used by more than 10 million speakers, and many of the others are spoken by less than a few tens of thousands of people. But the apparent enormity of this problem should not overwhelm us, as even a journey of a thousand miles starts with just one step.

Original source

Alt-i started developing aspects of human language technology using Yoruba as a pilot language. Yoruba was chosen because it has a wide speaker base, numbering more than 25 million. Despite its many dialects it has a standard form, from which a standard orthography has evolved over the past 100 years into a stable writing system. Yoruba is used as a language of instruction in higher education and doctoral theses have been written in it. The academic literature on Yoruba studies is substantial, and there is significant literature covering various scientific aspects of Yoruba that are relevant to human language technology. Given our limited

BRIAN SEED / ALAMY

resources, Alt-i chose to start with a language with such a large amount of available resources.

Our organization had already developed an efficient and ergonomic computer keyboard layout for Yoruba. We also adapted an HIV/Aids information kiosk, developed by the Africa Information Society Initiative (AISI) of United Nations Economic Commission for Africa (UNECA), for Yoruba. The kiosk provides HIV/Aids training for people who can neither read nor write, and delivers all the information using speech in Yoruba. It is used by the staff of local NGOs working in the health care sector.

But Alt-i's major achievement has been the development of a unique and more efficient system of speech recognition for African tone languages, in which the pitch of the speaker's voice contributes to the meaning of a word. The task of identifying these tones, in addition to those of identifying consonants and vowels, is normally viewed as an extra workload for speech recognition technology.

However, our research has shown that it is possible to use tonality to increase the speed and accuracy of speech recognition systems. Such a system can be used to convert speech to text and can be incorporated into laptops and desktop computers. More practically, it can also be applied to low-cost information devices such personal digital assistants (PDAs) or converted game consoles. These devices can be particularly useful in collecting and documenting valuable local knowledge from community elders who have not learned to read and write. Coupled with other speech technologies, such as speech synthesis, this knowledge can then be shared and refined by others.

Bang on

We first got the hint that tonality could be an important feature in speech technologies, particularly in speech recognition, after observing Yoruba 'talking drummers'. Based on an age-old traditional practice, these drummers are able to use the drum for functional communication, simply by varying the pitch of drumbeat. This showed that there must be sufficient information in Yoruba tones if listeners can still understand the meaning of a drum statement that has no consonants or vowels. To research this further we worked with members of the Ayan clan who are the repositories of talking drum knowledge. Studying under these traditional drummers was both an enlightening and a humbling experience.

The knowledge obtained from what was, essentially, an apprenticeship, was later analyzed using modern scientific theories from fields as diverse as ethnomusicology, psychology, linguistics, computer science and information coding theory. Based on this knowledge, we were then able to build a Yoruba speech recognizer that is more efficient than any of those derived from the existing literature on human language technology.

Because research efforts in the field of speech recognition began with English, a non-tone language, most attention is given to the recognition of consonants and vowels. At Alt-i, however, we found that if tones are accurately recognized they can suggest the vowels that bear them and the consonants that they accompany. Interestingly, tones are a lot easier to recognize than vowels and consonants.

Such features are common among many African tone languages. It is likely that tools and techniques can be developed and shared for other languages, including Asian tone languages. We are aware of the pressing needs of many other African languages and we are approaching the development of this product in such a way that will make it relatively easy to adapt to other African tone languages in the near future.

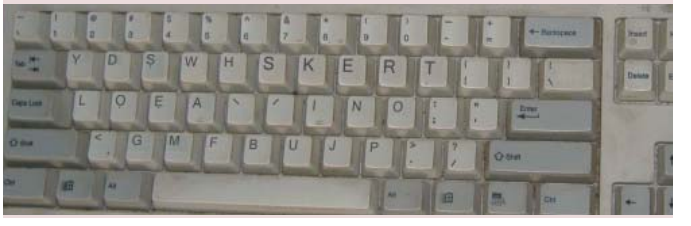
Words aloud

This work on speech recognition systems was conducted as part of a bigger Alt-i project, called Redefining Literacy, funded by the Open Society Initiative for West Africa (OSIWA), a part of the George Soros Open Society Institute. The project is working to develop speech recognition technology and speech synthesis to enable all Africans, whether literate or not, to interact with written literature. With this approach, Alt-i hopes to change the definition of literacy away from the 'ability to read and write', to the 'capacity to interact with literature'. Redefining literacy in this way will lead to the inclusion of people who would otherwise have been classified as illiterate. Their status is therefore transformed from that of illiterate to e-literate.

On a personal note, as an ordinary African village boy whose family has moved from a totally oral culture into the world of technology in merely two generations, I understand the problems of adapting. And, being an integral part of the world in the computer age I understand its possibilities. Modern ICTs have the power to keep Africa relevant in the knowledge era. It will be very unfortunate if we do not take full advantage of these opportunities. ■

The Yoruba keyboard

One of Alt-i's early achievements was the development of an efficient and ergonomic computer keyboard layout for the Yoruba language. To achieve this, a large Yoruba corpus was analyzed to derive the frequency of occurrence of each Yoruba character. Based on this, the most frequently used characters are placed in the most convenient typing positions. The keyboard was introduced to local typesetters and it has been a huge success. Before this, a typical page of Yoruba took three to five times longer to type than a page of English. This has now been greatly reduced. Ergonomic computer keyboards such as these can help reduce the problems of repetitive strain injury (RSI) associated with heavy computer use, and Alt-i is encouraging similar approaches for all languages for which keyboards do not yet exist.

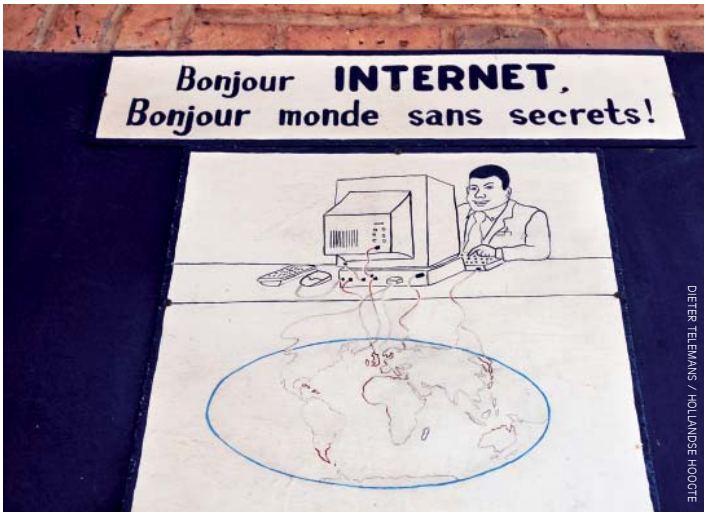


Related links

International Institute for Communication and Development (IICD)
➜ www.iicd.org
Commission économique pour l'Afrique des Nations unies (CEA/UNECA)
➜ www.uneca.org
Africa Information Society Initiative (AISI)
➜ www.uneca.org/aisi
Open Society Initiative for West Africa (OSIWA)
➜ www.osiwa.org
George Soros' Open Society Institute
➜ www.soros.org

Localizing languages

A partnership to provide technology in African languages is helping to make content and software more accessible to people across the continent.



DIETER TELEMANS / HOLLANDSE HOOGTE

The African continent is linguistically complex. Most countries have numerous indigenous languages and lingua francas. Some estimates suggest there are more than 2000 African languages, according to ethnologue.com (depending on how one distinguishes between closely related and inter-intelligible tongues). The various European languages inherited from the colonial era and retained for official use, add another layer to this complexity.

In such multilingual settings, the language(s) used is always a matter of choice, and that choice may have consequences. For example, deciding to work in one language rather than another affects who can effectively participate within a community, or can make a difference to how indigenous knowledge is used. Also, much traditional development work relies on multilingual extension officers or locals for translation, but for projects using ICTs, reliance on intermediaries is not always practical or desirable.

Fortunately, translation can be done before users access ICTs. Computers can, in principle, operate in any and all human languages. At the most basic level, this means ensuring that computer systems can handle special characters or non-Latin alphabets. Adding African language interfaces and content is not only possible but practical, as well as desirable, as people tend to access technology and information more readily in languages they know best. These languages, however, are often not the ones that dominate in the field of ICT.

Don Osborn (dzo@bisharat.net) is the founder and director of Bisharat (www.bisharat.net)

Pan-African support

The Bisharat initiative was launched in Mali in 2000 to help focus attention on the use of African languages in ICT projects. At that time, it was clear that African language content was barely visible on the emerging agenda for what is now called ICTs for development. Also, in the case of languages like Bambara, whose alphabets include modified Latin characters, there were challenges related to the use of locally developed 'special fonts'. These realizations in turn pointed to two needs – to emphasize the importance of African language content and computer interfaces, and to explore ways to overcome the technical issues connected with alphabets like that of Bambara.

The Bisharat website and discussion forums were the most extensive efforts, at the time, to explore the issues related to the use of African languages in ICTs. This work eventually led to discussions in 2004 between Canada's International Development Research Centre (IDRC) and Bisharat, and later also involving the NGO, Kabissa. The talks resulted in a new Pan-African Localisation (PAL) project to enhance the localization of technology in Africa, with particular focus on development and education.

By the time the PAL project began in April 2005, web-based content in African languages and some Africa-based projects to localize software were starting to appear. A PAL workshop in Pretoria, South Africa, in November 2007, brought together various people working to adapt technology to local languages. Their presentations described efforts to localize open source software in languages such as Wolof and Kinyarwanda, to increase web content in local languages (including blogs and a dictionary in Swahili), as well as efforts to develop terminology in languages such as Lingala, and research on speech recognition for Yoruba and Somali.

Add to these the efforts of Microsoft to localize much of its Windows and Office software into some African languages, and the commitment of the One Laptop per Child project to accommodate languages of the countries in which it will work, and an emerging trend is clear. A new IDRC-funded PAL project, led by www.translate.org.za, will begin in 2008 to focus on developing some key elements for localization in Africa, such as local data files, keyboard layouts and terminology. It will also look at how governmental policies, with regard to language and ICTs, can affect localization.

Full circle

The payoff for all of these efforts will be in their use on the ground. New and existing projects designed to use ICTs for development and education will need to incorporate these localization schemes. As the availability of software in more languages increases through local efforts and projects like PAL, there will be less technical justification for overlooking African languages in ICT projects in Africa. Fonts, keyboards, and, for some African languages, entire software office applications are already available as free or low-cost value-added elements to enable computer systems in Africa to fully accommodate the languages of their intended users.

Part of the challenge now is the marketing of African language products, and part is overcoming an apparent mindset that adding a new African language capacity to computers somehow detracts from the existing one, usually English or French. Ultimately, the technical issues for including African languages in ICTs may be less daunting than changing perceptions regarding the potential of using these languages on computers and the internet, and the lack of information about what is already being done. Multilingual computing is a reality – how Africans can exploit it optimally and appropriately is the question that now needs to be addressed. ■

Related links

PanAfrican Localisation
➜ www.panafril10n.org
Translate.co.za
➜ www.translate.co.za
KiLinux
➜ www.kilinux.udsm.ac.tz
Alf@net (Senegalese initiative for localization)
➜ www.alfanet.anafa.org

Technology that talks all languages

A number of language-related applications are free and easily available on the internet. The most widely used of these are translation tools.

Translation

There is a motto among professional translators that comes from a saying in Latin by St Jerome: *non verbum e verbo sed sensum exprimere de sensu*, which means they do not translate each and every word, but an overall sense of the original text. Online translation tools, however, do not work in this way. They tend to translate each word individually. Take the following quote in the original French, and compare the English translations:

'Tout ce qui n'est point prose est vers ; et tout ce qui n'est point vers est prose'
Molière, French dramatist and actor (1622–1673)

All that is not prose is towards; and all that is not towards is prose.
AltaVista – Babel Fish (http://babelfish.altavista.com)

Anything that is not prose is to, and everything that is not verse is prose.
Google Translate (www.google.com/translate_t)

All what not at all is prose is towards; and all which not at all is towards is prose.
Free translation.com – claims to have a basic understanding of the text

While all of the online translations are similar, none of them make any sense. The correct translation should be; All that is not prose is verse; all that is not verse is prose.

These tools are more useful when translating single words, but problems can arise here too, especially with words that are spelled the same but have different meanings – homographs.

Here is an example of a French homograph with its English translations using the online sources:

Café: coffee. Only Google Translate gives the alternative meaning; café, a place to drink coffee. This might seem trivial but consider the following sentence to see how confusion could arise:

French: Un homme entre dans un café.

Google Translate and AltaVista – Babel Fish translate this correctly as: 'A man enters a café'. However, both PROMT online (http://translation2.paralink.com) and Free translation.com both translate it as: 'A man enters a coffee'.

While this shows that translation software is unrelibale, it also indicates the difficulty of developing language technology.

Literacy

Individuals who unable to read or write can use various software packages, such as screen readers and speech recognition systems, to access content on the internet, compose emails and even create text documents.

Screen readers
Screen readers use a voice synthesizer to read out the text from a computer screen. Although originally developed for the blind and visually impaired, they can also be used by individuals who cannot read or write. There are many proprietary screen readers, but these tend to be very expensive. A few can be downloaded for free. Additional requirements include a sound card and speakers and/or headphones.

NonVisual Desktop Access (NVDA)
NVDA is an open source program for Windows XP and Vista operating systems. It reads text aloud from Mozilla Firefox and Internet Explorer web browsers, from documents in Word and Excel, and emails in Outlook Express.
www.nvda-project.org

Thunder ScreenReader
Thunder also runs on Windows XP, 2000 and Vista. Supported programs include Word, Excel and Internet Explorer. It is free to download for individual use; organizations and companies have to negotiate a price.
www.screenreader.net

Simply Web 2000
Simply Web works only with Internet Explorer (version 4.01 or later) and allows web navigation using a speech synthesizer.
www.econointl.com/sw

Note: A free speech synthesizer is available on Apple Macintosh computers that, when combined with an extra software plug-in for the Netscape browser, will read web pages out loud.

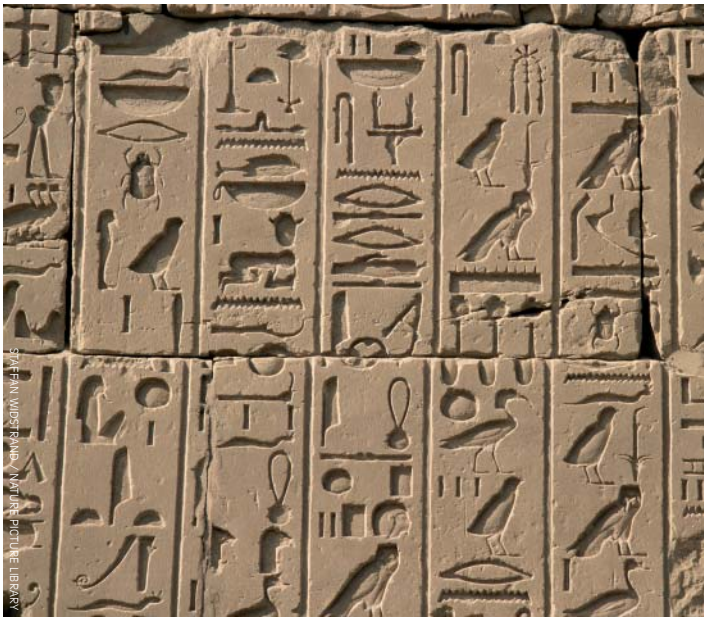
Automatic speech recognition (ASR)
With speech recognition software and a microphone, users can give vocal commands to open programs, create emails, dictate text documents and browse the internet. Most ASR programs are available in several (European) languages, but some voices and accents may present recognition problems.

E-speaking runs on Windows 2000 and XP, and works with Microsoft Office programs. After a 30-day free trial period, users can buy an upgrade licence for US$14, which gives the option to add or delete user commands.
www.e-speaking.com

SpeechVibe costs US$15 after a 30-day free trial period and operates on Windows 2000, XP and Vista.
http://speechvibe.com

SpeechTools is available for Windows 2000 and XP, SpeechTools costs US$9.95 and works with Microsoft Office programs.
www.speechtoolscenter.com ■

Note: Windows Vista includes a voice recognition capability, as does Apple Macintosh OS X.



Steven Bird (sb@csse.unimelb.edu.au) is associate professor in the Department of Computer Science and Software Engineering at the University of Melbourne, Australia (www.csse.unimelb.edu.au), and senior research associate in the Linguistic Data Consortium at the University of Pennsylvania, USA (www.ldc.upenn.edu).

Using the right words

What are the difficulties in producing technology that understands human language?

➜ Simple tasks that are second nature to us – and mastered in the early years of life – are surprisingly difficult to automate. The first step in the process is to work out what words are being spoken; to 'recognize speech'. Try saying 'wreck a nice beach' quickly to see that the computer's task isn't trivial. Add in dialect variation and some background noise and the computer's task becomes immense.
The second step is to work out how the words relate to the world around us. Suppose we read a news story containing the sentence: 'The thieves stole the jewels'. Imagine that the next sentence is any of the following, and work out what **them** refers to:

*a. Some of **them** were subsequently sold.*
*b. Some of **them** were subsequently caught.*
*c. Some of **them** were subsequently found.*

For a computer to understand the text, it needs to establish who did what to whom. Was it the thieves or the jewels that were sold, caught or found? To answer such questions we need to draw on our knowledge of the world and our ability to make inferences. It is extremely difficult to transfer our knowledge and reasoning abilities to a computer.
The third step is to respond appropriately. This may range from the trivial – such as playing a favourite piece of music or answering a simple factual question – to the impossibly difficult, such as engaging in extended discussion and critical analysis.

DIETER TELEMANS / HOLLANDSE HOOGTE

Is the problem partly human? Do linguists and technology producers need to cooperate more?

➜ There has already been cooperation in developing systems to process human language, including question answering systems and spoken dialogue systems. In this field of 'natural language processing', computer scientists and linguists are investigating human language and developing new techniques for processing language automatically. The field needs more people with computing skills who are also interested in language, and my forthcoming book *Introduction to Natural Language Processing* (downloadable free from www.nltk.org), is written with exactly this audience in mind.

What kinds of language technology are there?

➜ There are basically two kinds. One aims to bridge the gap between humans and computers by providing more naturalistic interfaces, including predictive text, speech recognition, and detecting emotions. The other is intended to bridge the gap between the vast amount of information available on the web and our personal information needs, including providing document summaries and automatic translation. These technologies exist for English and a handful of major languages where there has been substantial research and development investment.

Do these types of technology exist for most world languages?

➜ Today, according to ethnologue.com, about 7000 distinct languages are spoken worldwide. Only about half of these have an agreed writing system, and far fewer have any literary tradition. Some research institutions and development organizations are actively involved in 'language development', starting with the design of writing systems followed by mother-tongue literacy education. For example, in the 1990s I worked in western Cameroon to analyze several unwritten tone languages, and published a dictionary for one of them. If a local language has a well-established writing system, a literary tradition and government support, it is ripe for language technology.
The initial tools are as simple as keyboards and spell-checkers. More than this is rare: developing a speech recognizer, a search engine, or a translation system for the language requires considerable investment. Useful technologies exist to help expatriates learn a local language, for example, and Transcriber (trans.sourceforge.net) is useful for recording, transcribing and studying dialogues between native speakers.

Can you imagine a low-budget approach to providing technology for these other languages?

➜ The biggest challenge is to obtain large amounts of text and transcribed audio in each language, harnessing the free labour of many speakers for each one. I imagine a kind of YouTube, augmented with time-aligned transcriptions and translations, being widely used by speakers of local languages to document their linguistic heritage.
Also, a type of Wikipedia for each language could provide a useful body of text covering a wide range of topics. These materials would add up to a rich source of data to be used in developing technologies for each language.

Open Language Archives Community (OLAC)

OLAC is an international partnership of institutions and individuals who are creating a worldwide virtual library of language resources, including dictionaries, annotated texts and grammars. To date, OLAC has over 30 archives and about 30,000 catalogued items, for over 3000 languages from all over the world. Some items are digital and available online, while others are physical artefacts that can only be accessed by visiting the archive in person.
www.language-archives.org