

# **Strategic Importance of Speech Technology for NGNs**

## **White Paper**

**Stephen Rondel, Chairmand and CEO, Conversay**  
**Sr. Pat T. Pattabhiraman, Senior Computational Linguist**  
**Sr. Aravind Ganapathiraju, VP Core Technologies,**  
**Dr. Peyvand Khademi, COO Conversay**  
**Jennie Rondel, Managing Director CDN**

*Conversay*  
*15375 NE 90<sup>th</sup> Street*  
*Redmond, WA, 98052, USA*

## **Abstract**

We advocate the adoption of speech technology as a *strategic* component of Next Generation Networks (NGNs), and urge development and implementation of regulations and public policy for world-wide deployment of the technology embedded **IN** the devices accessing the NGNs. We highlight several key benefits brought by embedded speech technology to all nations of the world and describe how they fulfill important policy objectives cherished by the ITU and its member nations.

## **Strategic Importance of Speech Technology for NGNs**

### **1. Introduction**

This white paper highlights the strategic importance of speech technology for Next Generation Networks (NGNs), and advocates urgent development of regulations and public policy for world wide deployment of speech technology in the devices accessing the NGNs.

Next Generation Networks (NGNs) represent convergence between traditional public switched telephone networks and modern digital data networks. Network operators find NGNs attractive owing to substantial cost savings (brought about by a single converged network) and ease of sustainable growth. The attraction of NGNs for consumers arises from innovative new services, greater control and personalization, ease of migration between services, and so on.

NGNs are currently in embryonic stage. In the recent years, they have been increasingly occupying the thoughts of network operators, researchers, developers, industrial concerns, public policy makers, regulators and consumers alike. Indeed, on the matter of NGNs, the International Telecommunication Union (ITU) has initiated a study group [8], held several workshops in the last three years (for example, [14]), launched a global standards initiative [9], and set up many resources for strategists and policy makers.

Speech technology, encompassing Automated Speech Recognition and Text-to-Speech, enables humans to interact with electronic devices through human language. It is the most human and benign of all technologies. Formulation of regulations and policies aimed at including speech technology as an essential ingredient of NGNs will ensure that all nations of the world, developing or developed, reap the handsome benefits brought by ICTs, the internet, and the increasing computer power of cellular phones. Given the current embryonic stage of NGNs, it is critical that speech technology is included in the regulatory thought and action.

The overall objective of this paper is to enable regulators and policy makers to recognize the importance of speech technology for NGNs. Ensuring inclusion of speech technology in NGNs will help meet many public policy objectives recognized by the ITU and indeed, the United Nations at large, and held dear by the nations of the world. In section 2, we describe speech technology and its current state. In section 3, we discuss the impact of speech technology on

various issues pertaining to NGNs, and highlight the key benefits brought about by speech technology. We conclude the paper with a discussion in section 4.

## **2. Speech Technology**

Speech technology is a collective term referring to the intricate and sophisticated technologies of Automated Speech Recognition (ASR) and Text-to-Speech (TTS). ASR, also termed Speech Recognition, is a technology that enables computers, mobile phones and other software-enabled electronic devices to recognize and act upon spoken input in human (natural) language. TTS technology gives these devices the capacity to output their responses to human users in natural language.

An electronic application involving both ASR and TTS brings about complete human-device interaction in natural language. By enabling communication with electronic devices in spoken natural language, speech technology brings about a distinctly human character to the device, and has been praised for its ease and universality of use that cuts across national, cultural, social and economic divisions. It is spoken about as the great equalizer among modern technologies.

Speech technology is robust. Nevertheless, it continues to undergo improvement and evolution in the hands of researchers and developers of the technology focusing on such issues as improving accuracy of recognition, increasing speed and electronic memory usage, improving robustness of performance in noisy environments, reduction in power consumption, improving naturalness and intelligibility of synthetic speech, expanding the technology to work with more natural languages, and adapting the technology for use in the developing world.

Speech technology is applied and widely used on a daily basis by consumers, businesses and governments, and involves spoken interaction in dozens of languages around the world. Users of Information and Communications Technologies (ICTs) in the industrialized, developed world are familiar with its role as a beneficial component. However, a technology gap still exists as a manifestation of the well-recognized digital divide. Speech technology is yet to make its positive impact in most of the developing world, and in many rural and less-developed parts of the industrialized, advanced nations as well. The digital divide is likely to widen more adversely if future developments involving NGNs do not make the presence of speech technology a universal norm.

## **3. The Benefits of Speech Technology**

This section is devoted to the beneficial impact of speech technology on some key inter-related issues of information access and global development.

### **3.1. Speech Technology and the Digital Divide**

The digital divide is commonly used to refer to the perceived gap between those who have access to the latest information technologies and those who do not [3]. It is recognized that the digital divide exists within a developed society (such as the United States), as well as between the developed and developing or under-developed parts of the world. According to Compaine [3], the notion has substantial economic and political implications. Reduction of costs, increase of availability and natural acculturation of infrastructure as well as end products is acknowledged to help move in the direction of widespread adoption, thereby narrowing the digital divide.

New technologies have a pattern of being used first in the urban centers of the developed, industrialized world and spreading subsequently to the rural areas of the developed world and to the developing and least-developed nations. We may separate the factors that entrench the digital divide into (a) availability of technology and (b) usability of technology. The former factor has to do with whether there is robust infrastructure and whether devices (such as computers) are widely available to everyone. The latter has to do with whether the user is literate and possesses adequate technical skill to use the available technology.

The less-developed countries of the world have benefited by the advent of every aspect of modern ICTs. However, the benefits to their societies as a whole have been limited by the shortcomings in the infrastructure supporting the ICTs, as well as by the high rate of illiteracy and personal economic conditions of the people. The introduction of speech technology widens the availability, and as well, increases the effective usability of ICTs in those countries. With NGNs forming the next wave of change in the world of ICTs, a systematic inclusion of speech technology as a key ingredient of NGNs, especially as embedded in the devices, will ensure maximal access, and therefore, propagation of the benefits, to every nation in the world.

### **3.2. Speech Technology and Illiteracy**

Innovatively-designed user interfaces that include speech technology (ASR and/or TTS), often complemented by simple pictorial representations, have proven effective in making computerized information systems accessible to illiterate milieu. Successful efforts around the world have demonstrated the great potential of speech technology as the enabler of computer use and information access.

Plauché et al [10] describe a system with a speech-based user interface used in rural Tamil Nadu (India) giving agriculturally-pertinent information such as crop market prices to illiterate farmers. Timely access to crop price information boosts prices at which farmers are able to sell their crops. The authors describe how data collection and user interface design techniques have to be adapted to developing regions of the world, and how simple designs can yield flexible interfaces that enable tailoring the information system for use in different villages with different dialects of Tamil. The small vocabulary recognizer is inexpensive, scaleable, rapidly deployable, and well suited to limited resource environments.

[1] describes a web access system for illiterates in South Africa. South Africa straddles the highly-developed and underdeveloped worlds, and has been a well-suited locale for projects on illiterate-empowerment via technology. The project is concerned with the presentation of text-based internet information in audio format. Central to the information system is a *literacy server* which is interposed between the web client (browser software) and the web server (the internet site that the user wishes to access), and which processes every web page requested by the user. The literacy server employs TTS technology for producing audible output.

The SAMBAD project in rural Nepal [11] has demonstrated how non-literate can benefit from knowledge on the internet and in fact, contribute to that knowledge. Their technology focuses on speech and images, and allows creation of web pages containing spoken language and images. A TTS system for the Nepali language is being developed with a view to enabling rural non-literate folks to access written information that is valuable to them. The project has researched interaction methods that are suitable for people unfamiliar with writing, computers or advanced technology. The ultimate aim of this project is to enable non-literate people to participate fully in the information society and thus, reduce the digital divide.

### **3.3. Speech Technology on Cell Phones**

Cellular (mobile) phones have caught on like wild fire around the world. Most of the developing and under-developed countries with deficient conventional phone line infrastructure have gladly embraced cellular phones as a robust and affordable means of communication. In fact, countries of the developing world such as India and Brazil represent some of the fastest growing markets for mobile phones.

In the more educated and economically enriched populations of the developed world, cellular phones have expanded greatly in their function from being mere person-to-person talking devices into those that access the internet. The convergence of computing with the cell phone and the internet is spurring tremendous socio-economic growth through information sharing in those societies. But the digital divide grows if the benefits are denied to those who need it most.

Speech technology embedded in cellular phones enables easier (and potentially hands-free) dialing as well as access to various functions and software applications on the cell phone through spoken natural language input. As well, it enables access to information on the internet through spoken natural language commands to the cell phone. Regulations that would assure principled inclusion of speech technology (coupled with intuitive user interfaces) in cellular phones used in all countries will greatly help towards making the growth benefits of the technology potentially available to every cell phone user in the world, and meet the global development objectives of international policy makers.

### **3.4. Speech Technology and Economic Development**

Speech technology contributes valuably to global economic development by giving better access to education, health and e-government, and in general, by giving the advantage of information access and communication to everyone.

Speech technology is a vital educational aid. The direct use of speech in computer-aided instruction has great potential to reduce illiteracy simply by combining a simple workbook with a speech driven tutorial. [5] and [6] highlight the use of speech technology as a learning tool, in such educational tasks such as first- and second language learning. Parents and early childhood specialists use it to improve pronunciation, diction, accent and spelling in young children.

Cellular phones are contributing greatly to economic growth in the developing world [4] and [7]. Their contribution will be accelerated and further enhanced if they include speech technology as discussed in section 3.3.

### **3.5. Speech Technology and Linguistic Diversity**

Speech technology contributes to preserving linguistic and cultural diversity. Linguistic diversity is essential to preservation of cultural diversity. Speech technology has been implemented and used in many languages of the world. As it happens, the areas of the world with low levels of literacy, education and material standard tend to be the ones with high levels of linguistic diversity [2]. Currently, there is particular attention being paid to development of speech technology in “local languages” of the developing world [13], [11], [10]. Assuredly, it is a momentous task to develop speech technology in all languages of the world at once. However, manageable scenarios exist in which speech technology is introduced into the developing world in the dominant language or *lingua franca* of the region, and gradually extended to other languages or dialects based on demand. Countries and regions in the developing world enjoying

the benefits of speech technology can assume regional leadership in extending and transferring the technology to neighboring countries or regions. This process of technology is itself full of local industrial opportunities for the developing nations, and holds potential for economic development. In sum, speech technology contributes to the preservation of cultural and linguistic diversity, consistent with the principles adopted in the World Summit on the Information Society [15].

#### **4. Concluding Remarks**

Speech technology has several other valuable uses in addition to the ones discussed in section 3. It is critically helpful in emergency and disaster relief operations. It recently demonstrated usefulness in relief operations in the United States in the wake of hurricane Katrina [12]. It can transcend language barriers with such attractive applications as speech-to-speech translation. It has universal appeal, applicability and accessibility. With demonstrated success across diverse language groups, cultures and individual physical differences, speech embraces all nations of the world in a global information society. When made integral to mobile phones, the technology brings concomitant benefits to all nations in which mobile phones are used.

Technology, in general, has to be combined with appropriate skills and sound business models in order to sustain development [16]. Speech technology, as discussed in earlier sections, is not at all demanding of high literacy levels or technical skills, and is ideally suited for use in the developing world for sustainable development.

The exact terms of the regulations that would help accelerate world wide introduction of speech technology in NGNs is not fully clear at this point, and is certainly a matter for future discussion and deliberation, presumably at working sessions and workshops under the auspices of the ITU. What is clear is the strategic value of the inclusion of speech in mobile devices to overcome barriers to ICT access, preserve and transcend language diversity, and provide fresh new opportunities for education and information. This paper has provided a first step in underscoring the strategic importance of speech technology and the reasoning behind why it is important that regulators and policy makers consider speech technology as an essential component of NGNs.

#### **Acknowledgment**

We gratefully acknowledge Scott Randal's help in research and discussions.

#### **Bibliography**

1. ACACIA, International Development Research Centre, Exploring a Basic Illiterate Web Access System: Discussion and Demonstration of Technical Concepts, and Pointers to Future Research, 1997, [http://www.idrc.ca/en/ev-9515-201-1-DO\\_TOPIC.html](http://www.idrc.ca/en/ev-9515-201-1-DO_TOPIC.html).
2. Allwood, J., Linguistic Diversity and the Digital Divide, Crossing the Digital Divide: Shaping technologies to meet human needs, SCALLA Conference, Nepal, 2004.

3. Compaine, B., Re-Examining the Digital Divide, Internet and Telecom Convergence Consortium, MIT, 2000.
4. Fraser, M., Cell Phones Vital in Developing World, Associated Press news article, January 27, 2007. <http://www.forbes.com/feeds/ap/2007/01/27/ap3368377.html>
5. Gipe, T.S., Speak to me! Speech Technologies are Making their Mark, TechKnowLogia, Knowledge Enterprise, Inc., Nov-Dec 2001, pp. 46-48.
6. Gipe, T.S., Speaking in the Future Tense: A Look at the Speech Technologies of Tomorrow, TechKnowLogia, Knowledge Enterprise, Inc., Nov-Dec 2001, pp. 55-57.
7. GrameenPhone, 2006, <http://www.grameenphone.com/>.
8. Next Generation Networks: Lead Study Group on NGN and Satellite Matters, ITU-T Study Group 13 (Study Period 2005-2008), <http://www.itu.int/ITU-T/studygroups/com13/index.asp>.
9. Next Generation Network Global Standards Initiative (NGN-GSI), ITU, <http://www.itu.int/ITU-T/ngn/index.phtml>.
10. Plauché, M., Nallasamy, U., Pal, J., Wooters, C. and Ramachandran, D., Speech Recognition for Illiterate Access to Information and Technology, Proceedings of the 1<sup>st</sup> International Conference on ICT and Development, UC Berkeley, 2006.
11. Sambad: Access to Computers for Non-Literate People, <http://www.sambad.org/>.
12. Telecom Tech Aids Hurricane Efforts, Sep 30, 2005, <http://www.tmcnet.com/submit/2005/Sep/1188555.htm>.
13. Tucker, R. and Shalanova, K., The Local Language Speech Technology Initiative, Localisation of TTS for Voice Access to Information, Crossing the Digital Divide: Shaping technologies to meet human needs, SCALLA Conference, Nepal, 2004.
14. What Rules for IP-enabled NGNs? ITU Workshop, March 23-24, 2006, Geneva. <http://www.itu.int/osg/spu/ngn/event-march-2006.phtml>.
15. World Summit on the Information Society, Declaration of Principles, Geneva 2003 & Tunis 2005, Document WSIS-03/Geneva/Doc/4-E, 2003.
16. Yunus, M., Remarks at the Opening Ceremony of Telecom World 06, Hong Kong, 3 Dec 2006, <http://www.itu.int/partners/telecom2006/yunus.html>.