



Making DTV for All with TTS

Technology could help the blind get vital information

In my column, "Speech in a Digital World" (September 2008), I addressed the transition to digital TV, culminating on February 17, 2009, as it relates to the power that speech recognition offers for making on-screen menus accessible for people with vision disabilities.

Another concern is access to emergency information. The Federal Communications Commission (FCC) requires broadcasters to provide emergency information that is intended to further the protection of life, health, safety, or property, in a form accessible to individuals with disabilities. For deaf users, the familiar text scrolls, or crawls, across the bottom of the TV screen to describe pertinent emergency information accomplish this directive. But for someone who is blind, the visual crawl is of no value; for emergency information to be accessible to him, it needs to be aural.

Under FCC regulations, if the programmer interrupts programming to provide emergency information that appears via a crawl, it must be accompanied by an aural tone to alert people with vision disabilities that emergency information is being provided and they should tune in elsewhere to get it.

In Los Angeles, for example, officials at several local news affiliates state that "audio is always heard" when the Sheriff's Department issues an emergency statement via a news crawl. However, viewers in all cities may not share this experience.

While the FCC has clearly recognized the importance of providing emergency information to people with disabilities, "its rules regarding blind people and television...[don't] give any meaningful information about what's going on, and, from what I've learned from the community, are rarely, if ever, used," says Larry Goldberg, director of media access at WGBH in Boston.

Prototype in Place

Because the crawl is really just text, it is possible to translate the on-screen information from text-to-speech, facilitating compliance and eliminating the need for a human resource to record the message. In a federally funded project, Geoff Freed, project director for the Carl & Ruth Shapiro Family National Center for Accessible Media at the WGBH Educational Foundation, is doing just that.

Freed's group created a prototype that turns on-screen information, such as tornado warnings, school closings, and winning lotto numbers, into speech generated by a

text-to-speech engine. The prototype uses off-the-shelf software, Text Allowed, but networks could also purchase voices from other companies. The software takes the text source of the on-screen information, turns it into speech, and reinserts it back into the broadcast stream. The volume of the newscaster or actor's voice is lowered during transmission of the emergency information and automatically restored upon conclusion of the message.

Taking the concept further, the group is attempting to incorporate TTS into all audio for all segments of a show. On an all-news network such as CNN, for example, the newscaster's voice is always accompanied by a crawl. Typically,

Because the crawl is really just text, it is possible to translate the on-screen information from text-to-speech.

there are also one or two other windows on the screen displaying additional text. Prioritizing information for TTS processing gets tricky. Does the prototype default to the newscaster's voice, or does the crawl take priority? How does the system know whether the crawl contains emergency information or stock quotes? If the program defaults to one source of information,

how can a viewer using TTS switch from listening to the newscaster to listening to the information in a text box? And if the user can switch to TTS, how does she choose which source to activate for it?

Freed's team has no plans to sell the prototype. Instead, they will use the technology to show broadcasters it is possible to integrate text-to-speech with on-screen text. They will also write a white paper, including guidelines, so broadcasters can integrate TTS solutions into their own networks.

"The benefit to the networks," Freed says, "is that...this technology can help networks comply [with existing regulations]."

According to Freed, all broadcast stations already have the hardware necessary to create the graphics, and the prototype software works with existing equipment. The network needs only to write the software programming so it will work with its own internal systems.

As with other accessibility obstacles, finding solutions will take time and consideration, not just money. Speech recognition will likely contribute to the solution. Whether broadcasters will embrace this new method of making information useable remains to be seen, but Freed's reminder is still appropriate: "It's a public service." ☒

Robin Springer is president of Computer Talk, a consulting firm specializing in the design and implementation of speech recognition and other hands-free technology services. She can be reached at (888) 999-9161 or contactus@comptalk.com.