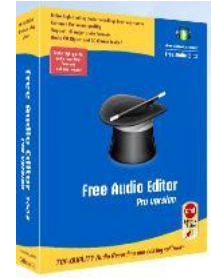Title: **Universal Access using Text-to-Audio & Sound Editing Programs**
Author: Gerry Kennedy © March 2009
Software: Text-to-Audio
Category: Creating voice files using Text-to-Audio Programs

1. Introduction

Creating sound files from recordings and from electronic or scanned text allows users to listen to information on a computer or on a more portable device. These include MP3 players, such as iPods, MP3 players, mobile phones, Portable Digital Assistants (e.g. Palm, HP iPaqs) and even digital cameras. Listening to audio, or listening whilst reading the same text, can dramatically improve access to text-based data. The quality of the audio is dependent on the program that records it, the environment in which it is captured, the synthesised voice use, the format in which it is stored and the player and speakers used in playback.

Students of any age or ability who cannot read a language, or comprehend text are at a distinct disadvantage when studying, researching or following directions and instructions. Traditionally educators and trainers in schools, higher education and work places have used and distributed published materials or provided books and study notes in printed hard copy format. In earlier decades, notes were written on paper, flash cards, blackboards or whiteboards for students to copy or transcribe. Students with poor reading, dyslexia and new arrivals struggling with English (or any other language being used) were disenfranchised and had little or no means to read and decode information.

The creation and mechanics in writing or typing text are complex. Reading and accessing text from paper based sources and electronically from computers, portable readers and now mobile phones can present many different problems and difficulties for some students. If they are able to listen to words, phrases, sentence paragraphs or whole documents, they can function with a degree of independence. Reading and listening to emails, web chat, web sites, blogs, Wikis, Twitter and RSS feeds and other online content is liberating and opens up a world of opportunities for engagement and access to social groups, news, current affairs as well as leisure options.

The ability to read and comprehend is taken for granted yet there are so many students who have difficulty. Translating text to audio is meaningful and practical as the user can then decide when wand where to use the technology to advantage.

Students can also record ideas, thoughts and opinions in voice and have it translated to text (using programs such as *Dragon Naturally Speaking*). Or they can review study notes or literature by using audio files created by themselves, teachers, aides, note takers lecturers, peers or siblings.

Computer applications and utilities provide opportunities to translate text to audio. These programs are either packaged in commercial products, offered as freeware or are utilities that can be accessed in programs such as MS Word (e/g. *Wordtalk V4.2*) or in a browser (e.g. in *Mozilla Firefox* using *TextAloud*).

2. Commonly Used Sound Formats

The most common sound file format in use at the moment is the *MP3* standard. MPEG-1 Audio Layer 3, more commonly referred to as *MP3*, is a digital audio encoding format using a form of lossy data compression. It is a common audio format for consumer audio storage, as well as a de facto standard encoding for the transfer and playback of music on digital audio players. MP3 is an audio-specific format that was designed by the Moving Picture Experts Group. The group was formed by several teams of engineers and was approved as an ISO/IEC standard in 1991.

The use in MP3 of a lossy compression algorithm is designed to greatly reduce the amount of data required to represent the audio recording and still sound like a faithful reproduction of the original uncompressed audio for most listeners, but is not considered high fidelity audio by audiophiles. An MP3 file that is created using the mid-range bit rate setting of 128 kbit/s will

result in a file that is typically about 1/10th the size of the CD file created from the original audio source. An MP3 file can also be constructed at higher or lower bit rates, with higher or lower resulting quality. [Source: http://en.wikipedia.org/wiki/MP3]

A huge range of MP3 capable devices exist with Apple Computer capturing the music market with *iTunes* software and its product line of portable players, the now ubiquitous *iPod MP3* players. Other players exist that offer other functions such as voice or sound recording, radio and video playback. They are usually discreet, small in size and very portable. The first players only offered 128MB of storage. Now they range from 1GB to 80GB of storage. The audio playback quality varies and some have multiple sound format capabilities including:

Windows Media Audio (WMA)	Windows Media Audio (WMA) is an audio data compression technology developed by Microsoft. The name can be used to refer to its audio file format or its audio codecs. It is a proprietary technology that forms part of the Windows Media framework. WMA consists of four distinct codecs.
WAV	WAV (or WAVE), short for Waveform audio format, is a Microsoft and IBM audio file format standard for storing an audio bitstream on PCs. It is an application of the RIFF bitstream format method for storing data in "chunks". It is the main format used on Windows systems for raw and typically uncompressed audio.
Ogg Vorbis	Vorbis is a free and open source; lossy audio codec project headed by the Xiph.Org Foundation and intended to serve as a replacement for MP3. It is most commonly used in conjunction with the Ogg container and is therefore called *Ogg Vorbis*.

Previously, sound files were stored in the WAV format. It is uncompressed format and takes up more storage space than an MP3 file. The *Sound Recorder* application (go to *Start/ All Programs/ Accessories/ Entertainment*) can be used to record sound using an inbuilt or external desktop, headset or handheld microphone. It is a very handy application but has limitations in its ability to record high quality sound or voice production. It is not an editor and has only a few effects (e.g. echo, reverse, or only increase or decrease speed by 100%).

Other more comprehensive audio editing programs exist including:

Audacity	http://audacity.sourceforge.net/ - is free, open source software for recording & editing sounds
WavePad Sound Editor	www.nch.com.au/wavepad/ - is a full featured professional audio editor
Free Audio Editor 2009	www.free-audio-editor.com/ - record audio from a microphone or any other input device
Audio Blast	www.moor-software.com/0.php?lang=English&page=blast.php - another free program
Soliton II	*http://biphome.spray.se/baxtrom/soliton.htm*
Reaper	www.cockos.com/reaper/ - a reasonably priced shareware program
Cool Edit 2000	*www.mp3-converter.com/cool_edit_2000.htm* -
Acid Pro V7	www.sonycreativesoftware.com/acidpro - a professional audio editor
Sony Sound Forge Audio	www.sonycreativesoftware.com/audiostudio - another powerful audio editor

3. **Recording Voice or Sound Effects**

You can use *Sound Recorder* or these other programs to record voice, music or sound effects whilst at the computer. The USB microphones or analogue models that have a USB converter provide the best quality. Some have noise cancelling or peripheral noise reducing functions. Microphones that are unidirectional are usually better than omni-directional. Price usually dictates performance and thus higher sound reproduction.

Digital Voice Recorders can be used to record voice files away from the computer. These portable devices vary on price and performance as well as features. The memory storage is an important consideration for the recording of lectures or

performances. Memory storage ranges from 256MG to 4GB and above. Some will allow for different sound format recording and transfer to a PC. These models usually have a USB facility whilst others have Bluetooth connectivity. Models from companies such as *Sony*, *Olympus*, *Panasonic* and *Sanyo* vary in price.

Most devices have microphone input and headphone jacks. Other features include storage 'folders', slow or fast playback speeds, counter(s), erase and hold functions. The *Hold* function stops inadvertent turning on or off. Audio files can also be uploaded to these devices and played back. Some allow for MP3 format whilst others use WAV or WMA. As with all technologies, some research is required before purchase. The user needs to determine and assess specific needs and identify the relevant functions that are critical or necessary before committing to a particular model.

4. Commercial Programs that Provide Text-to-Audio Translation

These programs will convert any text that can be selected or highlighted on screen to one or more audio or sound formats. Keeping in mind that some formats have inbuilt compression, there may be a loss f 'data' or audio quality. Choosing different synthesised voices may result in better outcomes. By selecting a male or female voice, increasing or decreasing sound pitch and speed may also provide more desirable results.

It depends on the user, the type of data being converted, the intended purpose and audience and ultimately what the final product will be used for in a setting or situation. Private listening is different from broadcasting. Many schools and tertiary institutions are streaming content or creating Podcasts. Higher quality audio would be desirable in these instances, as it will be recorded for a wider audience and more discerning listeners.

Some applications that can be used to convert text to audio include:

TextAloud   www.nextup.com

TextAloud uses voice synthesis to convert text into spoken audio. Students and other users can listen to text on their PCs or create MP3 or WMA files for use on portable devices such as iPods, Pocket PCs, and CD players.

- ❖ Users can directly open Word, PDF, and HTML files.
- ❖ It has automatic iTunes/iPod syncing.
- ❖ It has some advanced pronunciation tools
- ❖ Toolbar plug-ins are included for *Internet Explorer, Mozilla Firefox*, and *MS Outlook*
- ❖ Optional premium voices for a wide variety of accents and languages can be downloaded
- ❖ TextAloud is for Windows and a MAC OS version is available

Power Text to Speech Reader   *www.1speechsoft.com/?gclid=CMbbtvGQiJkCFSMSagodViRHmg*

Power Text to Speech Reader is a popular Text-to-Speech tool. It lets the user listen to documents, e-mails or web pages instead of reading on screen. It uses voice synthesis to create spoken audio from text with natural voices and converts text to MP3 format.

Natural Reader   www.naturalreaders.com/products.htm

The Free version no longer converts text to WAV files but the Personal and Professional versions do, at a cost. They also have 2 or more AT&T voices packaged with the commercial versions. The powerful audio output function allows users to convert a large text file containing up to 4 million letters to an audio file. Users can create single large-size audio files (up to 3GB audio files) and can convert a text file into a single audio file containing up to 3GB of information. NaturalReader allows students to convert text into MP3, WAV and OGG Vorbis audio files.

Verbose Text to Speech   www.nch.com.au/verbose/index.html?gclid=CO3lnqWViJkCFSMSagodViRHmg

Converts text to voice and can save as MP3. Verbose is a text to speech program that will read aloud any text or save it as MP3 or WAV formats. After users have installed this text reading software they can assign a system-wide hot key. Then whenever they want Verbose to read the text on the screen they just push that key and the software will read it aloud.

5.   Freeware Text-Audio Utilities

WordTalk V4.2   www.wordtalk.org.uk/Home/

This is a very powerful program for students with reading and writing difficulties, as having text reinforced by hearing it read aloud can be very useful. Specialised programs have existed to do this for a long time, and in many cases are extremely helpful and highly appropriate and should be seriously considered, perhaps in consultation with professional advice where necessary.

WordTalk is a free text-to-speech plug in developed for use with all versions of Microsoft Word (from Word 97 upwards). It will speak the text of the document and will highlight it as it goes. It contains a talking dictionary to help decide which word spelling is most appropriate.

Siting neatly in the MS Word toolbar it is highly configurable, allowing users to:
   * Adjust the highlight colours
   * Change the voice and the speed of the speech
   * Convert text to speech and save as a .wav or .mp3 file so that it can be played back on an iPod or mp3 player.

Spokentext.Net   *http://spokentext.net/*

SpokenText.net is a free text to speech converter. It allows users to record (English, French, Spanish or German) text in PDF, Word, plain text, PowerPoint files, RSS feeds, emails and web pages and converts them to speech automatically. Students can easily create audio recordings (in English, French, Spanish and German) of any text content. Students and educators can use this web site or the Firefox Extension to record books, articles, web pages, papers, class notes or any other text content that users need to access in audio format.
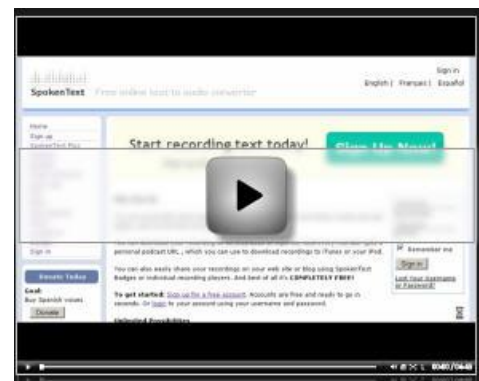
AT&T Labs   www.research.att.com/~ttsweb/tts/demo.php

This web site allows a user to type in up to 300 characters. It will then voice this text in a number of different AT&T voices (including Crystal, Rich, Mike, Claire and many others). The user can then download the WAV formatted file. It is particularly useful in creating sound grabs, narrated or voiced text to insert into MS PowerPoint or other multi media programs.

ISpeech   www.ispeech.org/convert.text.php

Convert any text to speech quickly and easily online at this web site. Simply enter text into the field and click "listen" or can cut and paste text into the box or drag and drop text for conversion. iSpeech will quickly convert text into natural sounding spoken audio. After users convert text to speech, they can listen to it with iSpeech's flash widget, embed it on websites, share it with friends, or download it and put it on users' iPods. The text to speech conversion can be used as a tool to catch up on useful information, get ahead at work or sit back and relax while users listen to their favourite author. iSpeech text to speech conversion is fast and high quality and always available on the Web.

Dspeech   http://dimio.altervista.org/eng/

Allows students to save the output as a .WAV, .MP3 or OGG file and quickly select different voices, even combine them, or juxtapose them in order to create dialogues between different voices. DSpeech integrates a vocal recognition system that, through a simple script language, allows users to create interactive dialogues with the user. Users can also configure the voices in an independent way. Cleverly integrating apposite TAGs, it allows users to dynamically change the features of the voices during the playback (speed, volume and frequency), to insert pauses, emphasize specific words, or even to spell them out. It also has a portable App version that will run from a USB memory thumb drive. This is available from RSC Scotland: www.rsc-ne-scotland.ac.uk/accessapps/.

6.   Assistive Software to Access Electronic Text or Text from Scanned Documents

Numerous other programs and web services exist. It is a matter of locating a text-to-speech program, utility or web service that converts text to an audio format of your choosing. The text, once converted from a third party source or from personal writing in MS Word or word processor, email or web site to a suitable sound file, can then be played back:

- ❖ On a computer using speakers, amplification or sound system for private or public performance
- ❖ On a computer for a student to read along with the text on screen
- ❖ At any time to read a paper-based version for editing, fluency, practise or study
- ❖ For memorising content
- ❖ For rehearsing purposes (e.g. in practising the delivery of a speech)
- ❖ Using headphones (for privacy or reflection or for use in public spaces)
- ❖ To read a section, part or whole book
- ❖ To listen to a web site offline
- ❖ To listen to an email offline
- ❖ To proof read a student's own writing
- ❖ To attend to instructions, directions or
- ❖ To follow a list of ingredients or method in a recipe
- ❖ To listen to a 'talking book' for leisure or fun
- ❖ To listen for meaning, clarity or to assess and appraise

7.   Features in Most Text-to-Audio Programs

Most programs have a number of voices, both male and female, from which to choose. Users can dictate the speed of delivery and often change pitch and/or tone. Some programs offer 'word pause' and even sentence or paragraph pause that will add extra pauses to text, where necessary.

Commercial literacy tools such as *textHELP Read & Write V9 (www.texthelp.com), ClaroRead 2008 for Windows* (www.clarosoftware.com) and *Wynn* ™ V5.1 (www.freedomscientific.com/LSG/products/wynn_new.asp) have text-to-audio facilities within these programs. Extensive features provide for high quality audio files using a range of human-sounding voices.

8.   In Conclusion

Text-to-Audio technology has matured over the past few years. It is being used in the business, telemarketing, telecommunications and telephony markets and should be an important aspect in education for all students. The portability and ease of access is available to students of all abilities. Programs use either SAPI 4 or SAPI 5 voice technologies to speak or voice the text in a document on the computer and then also convert it to an audio file.

Audio formats have improved and synthesised voice technologies are now more advanced with human sounding voices available such as Australian voices, *Karen* and *Lee*. Portable players can be used to listen to MP3, WMA or WAV files in any location, at any time for any purpose. It is liberating for students to access information in a format that is socially acceptable. Headphones provide for private listening or in busy, noisy environments. The text can be read on paper or in a book as well.

Students often fatigue due to poor reading skills, acquired brain injury or vision impairment. The ability to listen has been a huge benefit to students and people who are blind. All literature becomes universally accessible. Students of all abilities, though, can benefit from this technology. It can be performed quickly online or more formally using free or commercial applications.

Listen to your students, as they voice their opinion, in how they wish to access and listen, to your content!

Resources

A useful guide to audio editors with links to some more free programs as well as fact sheets and guides can be located at http://www.thefreecountry.com/utilities/audioeditors.shtml. For information on MP3 conversion programs with links to Commercial, Shareware and Freeware applications link to: http://www.mp3-converter.com/mp3_converter_freeware.htm. Some excellent presentations (as well as many other papers and resources) can be downloaded from the Call Scotland website at http://www.callscotland.org.uk/Resources/Presentations/.

A very handy web link is http://www.freedomscientific.com/LSG/resources/industry_links.asp#elec. Here users can locate a number of other links to web based text, books and free libraries of eBooks.

Some eBook sites such as the Gutenberg Project (www.gutenberg.org) offer books that are out of copyright, as MP3 files. This saves time for parents, teachers and aides in having to convert the thousands of free books to audio files as they have already been professionally converted for students.

NCH Software has developed a number of professional sound recorder programs for Windows, Mac and Pocket PC. Each sound recording program is specifically designed for particular recording tasks including general audio recording, voice recording, music recording and more. Link to: www.nch.com.au/software/soundrec.html?gclid=CJHqpbqfiJkCFRwDagodhFMnmQ for more information and a list of programs that all work with audio files.

Author: Gerry Kennedy ICT Consultancy © March 2009