

Creating MP3s

High Tech Center Training Unit

of the California Community Colleges at the
Foothill-De Anza Community College District

21050 McClellan Road
Cupertino, CA 95014
(408) 996-4636

www.htctu.net

Table of Contents

Training Objectives.....	4
What is an MP3?.....	4
How does MP3 technology work?.....	4
Why is the MP3 format appropriate for student use?	5
How do I prepare a file so I can convert it to the MP3 format?	6
After I have prepared a file, how do I create an MP3?	8
How do you create multiple MP3 files in <i>TextAloud</i> ?	10
How can you use the File Splitter Utility in TextAloud MP3?	12
What other software packages can create MP3?	13
How do you play MP3 files?	14
How is the material stored?	17
Burning (Writing) MP3 files to CD-Rom.....	18
Advanced creation of MP3 files:	18
Other Compressed File Formats	20



Training Objectives

This is a one-day training designed to explain the use and creation of MP3 audio files as an alternate media format. At the conclusion of this training, participants will be able to do the following:

1. Understand the general technology behind MP3 sound files,
2. Create MP3 sound files from e-text using TextAloud and Kurzweil 1000, and
3. Understand how to play back MP3 files on computers or portable devices.

This training is designed for alternate media specialists but open to anyone who is interested in learning about the use of MP3 technology.

What is an MP3?

MP3 (MPEG-1 Audio Layer-3) is a standard technology and format for compression of a sound sequence into a very small file. The resulting file is commonly referred to as “an MP3.” Files in the MP3 format have an extension of **.mp3**. MP3 technology is one example of the broader technology known as digital audio.–

The goal of the MP3 format is to compress a sound file by a factor of 10 to 14 without noticeably affecting the sound quality, allowing for the storage of hundreds of audio files on a computer’s hard disk. The ability to compress the file size while preserving the original level of sound quality of the audio file is what has made MP3 music compression the most popular format for transferring, storing, and listening to music on the Internet.

How does MP3 technology work?

There are two techniques that are used together to help compress the size of the sound file into an MP3 without losing noticeable sound quality:

1. **Perceptual noise shaping:** The MP3 format uses characteristics of the human ear to design a compression algorithm. For example, there are certain sounds that the human ear cannot hear and certain sounds that the human ear hears much better than others. Using facts like these, certain parts of a sound sequence can be eliminated without significantly hurting the quality of the sound for the listener.
2. **File compression.** Compressing the rest of the sound sequence further shrinks the file size considerably.

Why is the MP3 format appropriate for student use?

The MP3 format has characteristics that are ideal for students with reading disabilities or others who have trouble with the printed word. Recording spoken text onto MP3 files allows for:

- **Better audio quality** – Digital audio produces recordings with significantly less noise than analog recordings (e.g., it minimizes "tape hiss"). Also, unlike analog recordings, the copies sound exactly the same as the original recording.
- **Better storage** – The MP3 file format allows for the encoding of very large amounts of spoken text into a very small sound file. For example, on a CD you can store 74 minutes of audio. However, if you compress the same data into the MP3 format, you can store over 40 hours of recorded material.
- **Convenience and portability** – The MP3 file format enables students to play spoken text without a computer. There are many devices that allow students to store, transfer, and play their MP3 files portably. [Some of these devices will be described later in this manual.] As a result, students have the flexibility to use spoken text with as much freedom as a person carrying around a textbook.
- **Instant access and improved navigation** – Unlike an audiotope, the MP3 file format allows a student to instantly jump from one part of a sound file to another. The audio does not have to be navigated from beginning to end, as is the case for tapes. Digital audio enables you to start the book at any given point. For example, you will be able to go directly to a certain page number or move from chapter to chapter at the press of a button - there is no need to fast-forward or rewind through beep tones as you do with audiotapes. A user can also "bookmark" certain sections of a sound file so they can later reference important sections of a book, article, or manuscript.

How do I prepare a file so I can convert it to the MP3 format?

Step 1: Create a profile for each student

With text-to-speech technology, there are many variants such as the playback rate and the synthetic voice of the text that is read back. It is important to find the best fit for the student using the audio file. Even with the same student, he or she may prefer that certain subject matter be read at a different rate or in a different voice than other subject matter. Before creating MP3 files for your student, set some time so that you and the student can listen to the different synthetic voices and different reading speeds. It would also be a good idea to keep track of different students' preferences in a document so that you do not need to consult the student every time you want to create an MP3 file for him or her. Here is an example table:

Name	Subject matter/ Textbook	Voice	Reading Rate (Words Per Minute)	Encoding Rate	Notes
John Doe	<i>The American Experience</i>	“Jennifer”	150	32 kbps (mono)	Uses a screen reader
John Doe	<i>The Art and Science of C</i>	“Mary”	100	48 kbps (mono)	Uses a screen reader
Amanda Hugandkiss	<i>Alice in Wonderland</i>	“Mike “	200	32 kbps (mono)	

Step 2: Prepare the document and save it as a .txt file

Make sure that the text you want to encode is in plain-text format (i.e., no formatting) and that any ASCII art has been removed from the file. If your file is not already in the .txt format, you will need to re-save it in the .txt format.

Step 3: Set up the pronunciation editor

The Pronunciation Editor allows the user to specify how *TextAloud* will pronounce words. In general, the synthetic voice reads most words correctly, but it may mispronounce certain words, such as acronyms, names, or other unique words. To change how *TextAloud* pronounces a word, you can follow these steps:

- Go to **Options** and select **Pronunciation Editor**.
- A new window will appear with a text box labeled **Word**. In the Word text box, add the word that you want *TextAloud* to correctly pronounce.
- In the **Pronunciation** text box, type in the word in such a way so that when you click the **Test** button it sounds correct (Hint: Try to spell the word phonetically.) This will change the way the word is read in all applications that use text-to-speech.

Step 4: Set up predefined pauses.

When some documents are read, they need pauses in them. *TextAloud* allows you to add predefined pauses to a document. To set up the predefined pauses, go to the **Options** menu and select **Pre-Defined Pauses**. A window will appear with two input boxes that allow you to edit two pause intervals, a Short Pause interval and a Long Pause interval. The number values are in seconds, so if you were to put 2.5 in the Short Pause interval and 4 in the Long Pause interval, *TextAloud* would pause for 2.5 seconds at a given location at which you inserted a Short Pause interval and for 4 seconds anywhere you inserted a Long Pause interval

To insert the pause intervals, right-click with your mouse where you want a pause inserted. Then, select **Insert Short Pause** or **Insert Long Pause**. After inserting the pause, *TextAloud* will insert a string of text into the document. The text inserted will look like this:

" {{Pause=x}} "

with a blank space between the string. The value of x will be a number that tells *TextAloud* to pause for x seconds.

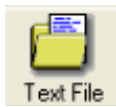
After I have prepared a file, how do I create an MP3?

Once your electronic text document has been prepared, you are ready to create an MP3 file.

TextAloud MP3 is one software program that uses “Text-to-Speech” technology to record text information to an audio format. You can create ordinary .wav files, or save the output in compressed MP3 format. You can create MP3 files from your word documents, email, news articles, or any text you want, and download the content to your computer or portable MP3 player. *TextAloud* is developed by NextUpTech. The software can be purchased for \$24.95 (per user). More information about the software can be found at <http://www.nextup.com>

Step 1: Open the plain-text file in TextAloud.

You can do so by going to the **File** menu and selecting **Open Text File**, or by pressing **Ctrl + o**.



Step 2: Select a file title.

After opening the text file, you will need to select a name for the MP3 file you are about to create. (You can change the name later if you want to.)

Step 3: Pick the synthetic voice that you would like to read the text file.

To select a synthetic voice, go to the **Engine/Voices** tab of *TextAloud* and use the **Test** button to get a sample of what each voice sounds like.



Step 4: Set the rate of the synthetic voice.

The rate refers to how many words per minute (WPM) will be read. Again, you should consult the student for whom the MP3 is being created because students will differ in what rate they can easily understand the text and depends on what subject matter is being read. To configure the rate of the synthetic voice, go to the **Engine / Voices** tab of *TextAloud* and click on the **Edit** button. A box should appear with a slider that lets you adjust the rate of the speech. To hear the current rate setting of the synthetic voice, press the **Test** button. When done setting the rate, click the **Ok** button.

Step 5: Set the encoding rate.

The encoding rate can be changed in the **File Options** tab of *TextAloud*. The file output can be saved in .wav or .mp3 file formats. If you pick the output to be .mp3, then you need to decide the encoding rate you want for the MP3. The encoding rate is anywhere from 16 kbps to 128 kbps. The frequency ranges from 11.02 kHz to 48 kHz. You can also select mono or stereo sound. The lower the encoding rate, the lower the sound quality and the smaller the file will be. The same is true about recording in mono sound, which will result in a file about half the size of a stereo encoded file with the same bit rate.

Step 6: Set the Write-to-File speed.

This option is in the **File Options** tab of *TextAloud*. The larger the number is, the faster the MP3 will be created. However, the greater speed will also place a greater demand on your computer's processor. If you have an older computer or have other applications running while you are creating the MP3 file, it is recommended that you keep the write-to-file speed around 10x.

Step 7: Select a file output directory.

The File Output Directory is where the finished .mp3 or .wav file will be stored on your computer. It is located in the **File Options** tab of *TextAloud*.

Step 8: Create the MP3 file.

The MP3 file can be created by pressing **Ctrl + F7** or by pressing the button with a play button and disk image on it.



There is a status bar in the lower right-hand corner of the *TextAloud* window that will show you when the file is done being created.

How do you create multiple MP3 files in *TextAloud*?

With *TextAloud* you can create many MP3 files simultaneously instead of having to wait for each file to complete. This section discusses the steps to do this.

Step 1: Change the settings from Single Article Options to Multi-Article Options.

To do this, go to **Options** on the menu bar and select **Multi-Article Options**. You can also do this by clicking on the **Multi** button.

Step 2: Set voice and file output options.

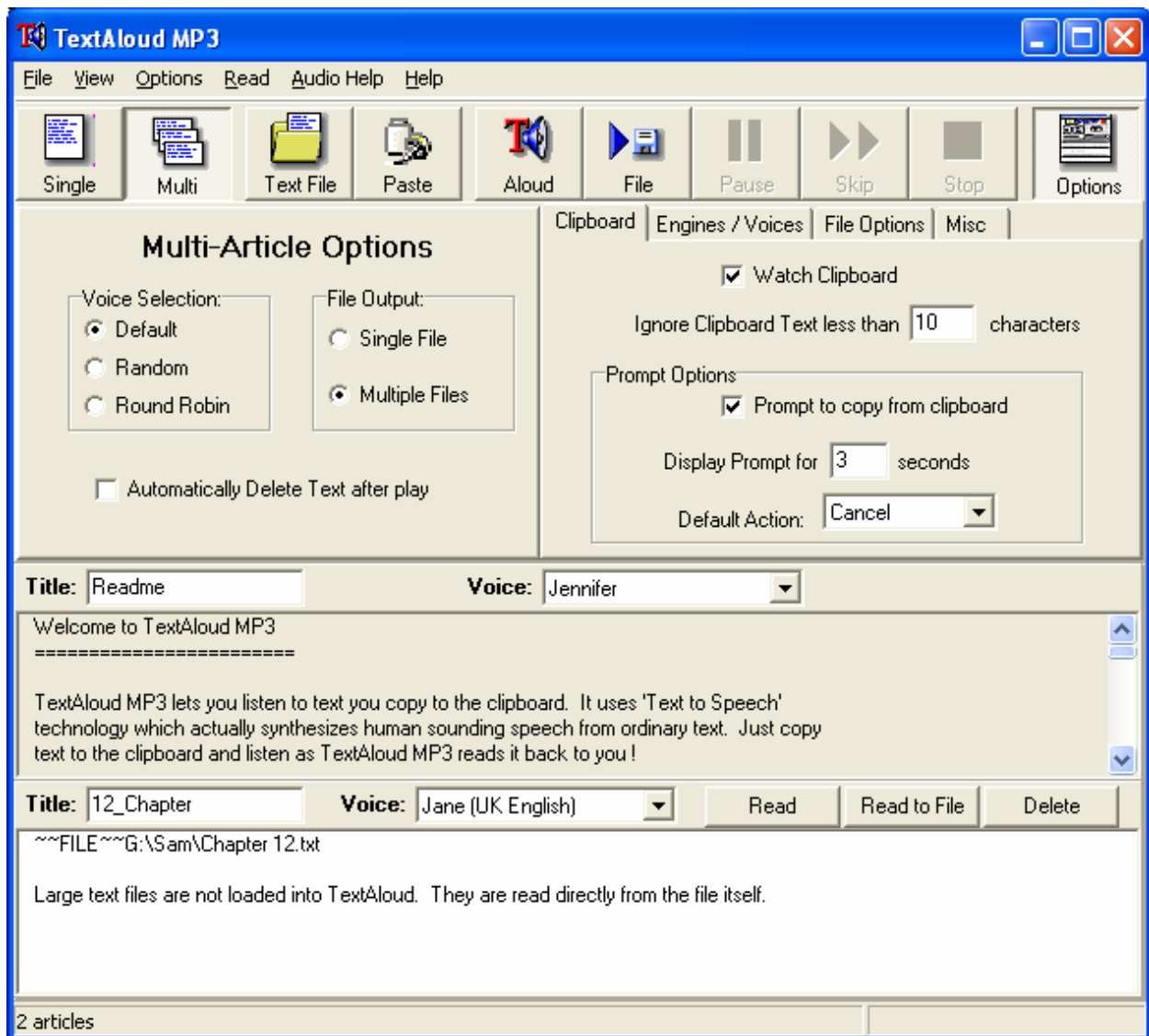
The Multi-Article Options pane to the upper left of *TextAloud* has two radio button menus:

- a) The **Voice Selection Menu** has three options **Default**, **Random** and **Round Robin**.
 - When the Voice Selection Menu is set to Default, all the documents will be read with the default synthetic voice.
 - If Random is selected, *TextAloud* will pick a random synthetic voice to read the documents.
 - Round Robin will cycle through all the voices that are available to *TextAloud* and will use a different voice for each document that you have open.

- b) The **File Output Menu** has two options: **Single File** and **Multiple Files**.
 - When **Single File** is selected, all the documents that are open in *TextAloud* will be read and combined into one single file.
 - When **Multiple Files** is selected, an individual MP3 will be created for each of the documents that are open in *TextAloud*. An advantage of the Multiple Files option is that you can, for example, create individual MP3 files that correspond to chapters or sections of a book. Having an MP3 file for each chapter will allow the student to more quickly navigate their reading material than it would be if all of the information was contained in one large MP3 file.

Step 3: Open the text files to be created into MP3s

Go to **File** on the menu bar and then select **Open Text File** (or press **Ctrl + o**) select the files that you want and repeat this step until all the text documents you want are open.



Step 4: Repeat steps 3-8 for creating a single MP3 file in *TextAloud*.

How can you use the File Splitter Utility in TextAloud MP3?

As of version 1.446 of *TextAloud MP3*, *Nextup.com* added a utility to *TextAloud MP3* that can be used to split files into smaller files. One limitation of earlier versions of *TextAloud MP3* is that if you had one large e-text file, converting it to MP3 format resulted in a file that was large and hard for the user to navigate. A good example is to think of trying to create an MP3 file for a 200-page chapter of a textbook. The MP3 file that would be created in TextAloud would be over an hour and hard for the student to navigate within the chapter. A way to alleviate this problem is to split the e-text of the chapters of the book into smaller files so you can create many small MP3 files. With small MP3 files, the user can easily navigate within the chapters of the textbook.

Users have two options for determining how the file will be split into multiple smaller ones. The first option is to specify a target size for each of the small files in kilobytes. The second option is to set-up a list of defined characters (delimiters) that will split the file each time that they occur in the document. The following steps will walk you through the instructions for either set-up

Step 1: Open the File Splitter Utility in TextAloud MP3 and select the input file name.

In TextAloud MP3, go to the menu and select **File, File Splitter Utility**. In the Input File Name, type the name of the file or select the Browse button to search your computer for the text file that you want to split into smaller files.

Step 2: Select the directory where you want the split files to be stored.

In the Output File Directory, type the path of the directory or select the Browse button to specify where the split files should be stored on your computer.

Step 3: Set the starting output file sequence number.

The Starting Output File Sequence Number is the starting number that the split files will be named. For example, if you have a large text file called Alice.txt and you set the Starting Output File Sequence Number to 4, the split files will be named Alice_004.txt, Alice_005.txt and so on.

Step 4: Select the split method.

As described earlier, there are two ways to split a text file. To split them by output target file size, set the size (in kilobytes) that the File Splitter Utility will split the input file into. To split the larger file by specified strings in the input file, specify a list of delimiters which can be set in the comma delimited list of strings text box.

Step 5: Preview and split the input file.

Select the Preview button to display how many files the input file will be split into. Select the Split button to complete the process.

What other software packages can create MP3?

TextAloud MP3 is not the only software program that will create MP3 files from text documents.

***Kurzweil 1000 ver.8* from Kurzweil Educational Systems.**

- This program is a scan and read application for users with visual impairments. It can be purchased for \$955. More information about the program can be found at http://www.kurzweiledu.com/products_k1000.asp
- Users can create MP3s from files that they open in Kurzweil by selecting File -> Utilities -> Create MP3 files. A window will open with options.
- The Range option can be set to Pages or All. If you choose All, then all the pages in the document will be made into MP3 files. If you choose Pages, then only the page numbers that you specify will be made into MP3 files.
- The Pages per MP3 file is the number of pages that you want read per MP3 file created. For example, if you have a 10 page document and you select 2 pages read per MP3 file, there will be 5 MP3 files created.
- The Audio quality option allows the user to pick between 3 encoding rates of the MP3 file.
- The Destination is where the created MP3 files will be saved after they are created.
- Windows Media Encoder or the LAME MP3 encoder must be installed to create MP3 files with Kurzweil.
 - The Windows Media Encoder can be downloaded at: <http://www.microsoft.com/windows/windowsmedia/9series/encoder/default.aspx>
 - The LAME encoder can be downloaded at: <http://mitiok.cjb.net/>

***Text-to-Audio* from Premier Programming**

Text-to-Audio is another application that can create MP3 files from electronic text. It can be purchased for \$50. More information about the software program can be found at: <http://www.premier-programming.com/content/productdetails/texttoaudio.html>

How do you play MP3 files?

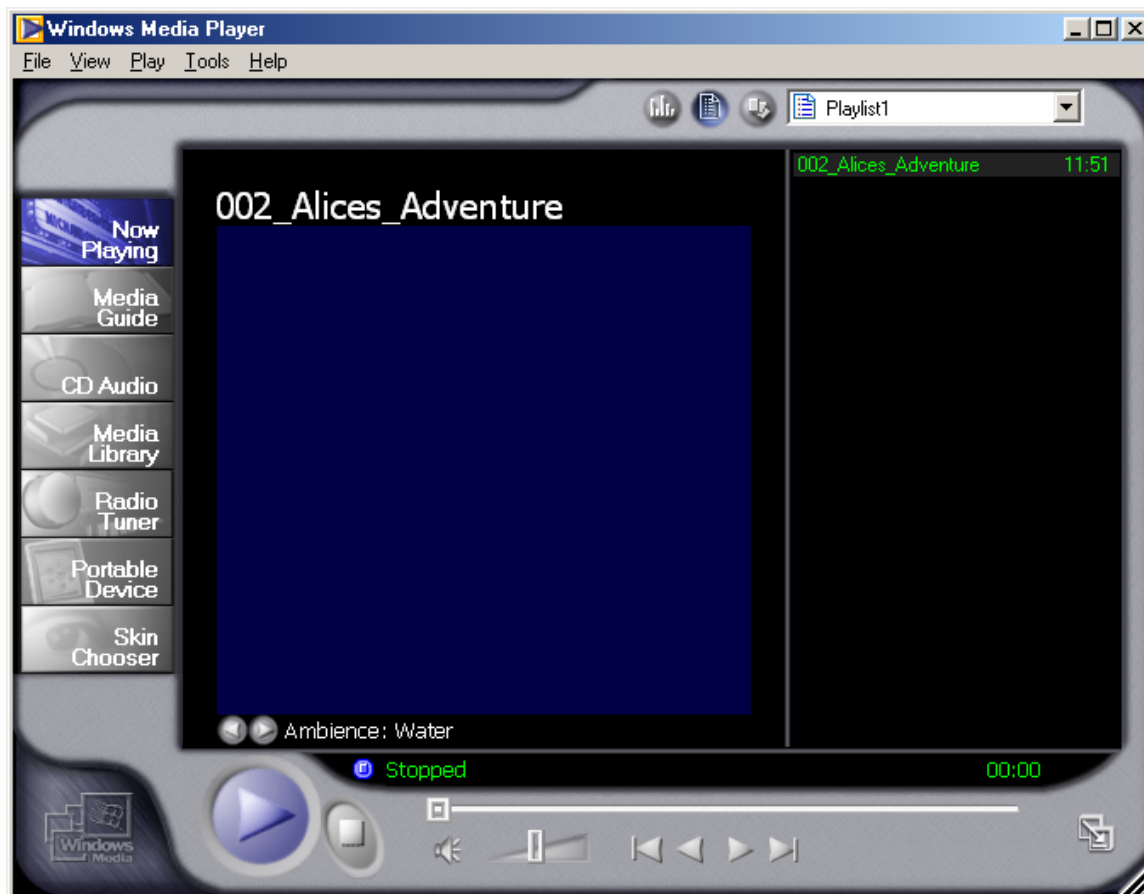
SOFTWARE PLAYERS

There are many different programs that will play MP3 files, including Windows Media Player and Winamp.

Windows Media Player:

Windows Media Player is a free application that is already installed in Windows 98, Me, 2000, and XP. To open Windows Media Player, click on the **Start** button -> **Programs** -> **Accessories** -> **Entertainment** -> **Windows Media Player**, or locate it manually on your harddrive.

Windows Media Player includes function commands that are similar to the buttons found on VCRs and CD players. To access the commands, click the **Play** menu, and then click the command. You can also access the commands by using the buttons along the bottom edge of Windows Media Player. Drag the slider, which is just above the commands, to specify where you want the file to begin playing.



If your computer does have Windows Media Player already installed on it, you can download the current version of Windows Media Player at <http://www.microsoft.com/windows/windowsmedia/players.aspx>

Winamp:

Winamp is an audio player. It plays music, or any other recorded sound, on your computer. Winamp also has a number of popular plug-ins for visualization (cool effects that synchronize with the music), input, output, and other purposes. It also allows the look of its interface to be customized through the use of skins. Though Winamp is most popular as a player for MP3s, it also supports a number of other formats.



You can download Winamp for free at <http://www.winamp.com>

You can download Winamp plugins at <http://classic.winamp.com/plugins/>

Other MP3 software players:

Other MP3 players can be found at:

<http://www.mpeg.org/MPEG/MPEG-audio-player.html>

HARDWARE PLAYERS

The following hardware players allow you to play MP3 files without the use of a computer and also allow you to use them portably, much like a personal walkman.

When picking a portable MP3 player, keep the user in mind. For instance, if a person has a disability that makes it difficult to see a visual display, certain hardware players may be better suited than others. [For example, the MuVo might be better than an iPod for a blind or visually-impaired user since it does not rely on a visual display.]

Apple Computer iPod



Apple Computer has a portable MP3 player called the iPod that will work with Mac or Windows computers depending on the module that you buy. The iPod comes in three storage capacities: 5 GB, 10 GB, or 20 GB. The iPod has a physical hard drive with moving parts inside of it. This means that the iPod is more fragile than an MP3 player that uses flash memory. The iPod connects to the computer with a Firewire cable. Firewire is faster than USB so you can transfer files faster. Unfortunately, most PC computers do not come with Firewire ports so you will need to install a Firewire PC card to connect your iPod to your computer. The iPod has an easy visual navigation for picking the files that you want to play.

The starting model of the iPod (with a storage capacity of 5 GB) costs \$299.

More information is available at <http://www.apple.com>

Creative Nomad MuVo

The MuVo has two models: a 64 MB and a 128 MB version. The MuVo uses flash memory, which means that it does not have moving parts inside. The flash memory makes the player more durable. The MuVo connects to the USB port on a computer and the operating system treats it as a removable hard-drive. The MuVo uses touch button controls and does not rely on visual display controls, so it is a more appropriate device for people with visual disabilities. In addition to playing MP3s, the MuVo can also play .WMA files.

The MuVo 64 costs \$89.

The MuVo 128 costs \$147.

More information can be found at <http://www.nomadworld.com/products/muvo/>

Other MP3 Hardware players:

A list of portable MP3 players can be found at <http://www.mpeg.org/MPEG/mp3.html#portable>

How is the material stored?

One of the many reasons that MP3 files are ideal for alternative media is the size of the files and how they can be stored on many different media.

For example, if you encoded Chapter 2 of *20,000 Leagues Under the Sea* by Jules Verne (1583 words long) and you have the synthetic voice read it at a rate of 175 words per minute, the file is about 8.53 minutes long. The size of an MP3 encoded at 32kbps mono is 2.4 MB.

The following table shows some of the different storage options for storing MP3 files and the time (in minutes) of spoken word that can be stored.

Media	Storage Space (in Megabytes)	Time in minutes of spoken word in MP3 format with 175 WPM reading rate encoded at 32kbps mono
3.5 Floppy Disk	1.44	4
64 Megabyte MuVo MP3 Player	64	224
100 Megabyte Zip disk	100	350
128 Megabyte MuVo MP3 player	128	448
250 Megabyte Zipdisk	250	875
650 Megabyte CD	650	2,275
700 Megabyte CD	700	2,450
DVD-R	4,700	16,450
5 GB iPod	5,000	17,500
10 GB iPod	10,000	35,000
20 GB iPod	20,000	70,000

Burning (Writing) MP3 files to CD-Rom

There are two types of CDs that you can create: an audio CD or a data CD. Both types have advantages and disadvantages. An audio CD can be played on almost any CD player. The disadvantage of burning an audio CD is that it can only hold 74 minutes of audio sound. An advantage of having a data CD is that you can store over 74 minutes of spoken text and also store other files on the CD. For example, you could potentially store several MP3 files plus the documents or text files that corresponds to the audio file on the same CD. That way, the student can have both the audio as well as the text document on a single CD. The disadvantage of data CDs is that you have to play the audio files on a computer or a special CD player that is made to play MP3 files.

Advanced creation of MP3 files:

Encoding from .wav to .mp3

An MP3 encoder is a program to generate .mp3 files from .wav files. There are many different MP3 encoders. All encoders will output a playable MP3 file, but the size, sound quality, and speed at which the MP3 file is encoded differs greatly depending on the encoder you use.

- The **LAME encoder** is a free MP3 encoder that is used in many applications. *TextAloud* uses the LAME encoder to make its MP3 files. The LAME encoder has many options that are not available in *TextAloud*. Some of the options include encoding rates, constant or variable bit rate encoding, setting high and low pass filters. The LAME encoder is free and can be downloaded at <http://lame.sourceforge.net/links.html>
- The **Fraunhofer encoder** is one of the pioneers of MP3 technology working alongside other entities like Thomson Multimedia to develop the format. There are many different versions of the Fraunhofer encoder. Some versions of the encoder will not let you encode files at a high bit rate. More information can be found at <http://www.iis.fraunhofer.de/amm/download/index.html#2>
- The **Xing encoder** is a commercial encoder from RealAudio. The Xing encoder is fast and can encode some files faster than LAME and Fraunhofer encoders. Xing has many of the same options as the LAME encoder. There is also a command line encoder that can be used to encode MP3 files. The Xing MP3 encoder is \$19.95 and can be bought at <http://www.xingtech.com/>

Editing

There are programs that will let you split MP3 file into smaller pieces and also let you join them. This can be helpful if a student's MP3 player cannot hold large files. If you want to join two related document together you can use a joiner. *MP3 Workshop* is one a piece of software that

will allow you to split and join MP3 files. It cost about \$20. More information can be found at <http://www.audio2x.com/ezmp3/index.htm>

Other Compressed File Formats

WMA

Windows Media Audio file format is a digital audio format created by Microsoft. All Windows computers that have Windows Media Player 6 and up can play WMA files. Reviews of WMA rate them comparable to MP3 files for size and quality. Some of the portable MP3 players can play WMA files.

OGG

Ogg Vorbis is a new audio compression format. It is roughly comparable to other formats used to store and play digital music, such as MP3, VQF, AAC, and other digital audio formats. It is different from these other formats because it is completely free, open, and unpatented. Currently most portable MP3 players will not play OGG files but most of the software MP3 players will also play OGG files. More information can be found at <http://www.vorbis.com/>

Speex

Unlike other MP3 and Ogg Vorbis, Speex is designed to compress voice at bitrates in the 2-45 kbps/channel range. Possible applications include VoIP (Voice over IP), internet audio streaming, archiving of speech data (e.g. voice mail), and audio books. In some sense, it is meant to be complementary to the Ogg Vorbis codec. It is currently in beta stage and the standard has not been finalized, but is something to keep in mind in the future. More information can be found at <http://www.speex.org>