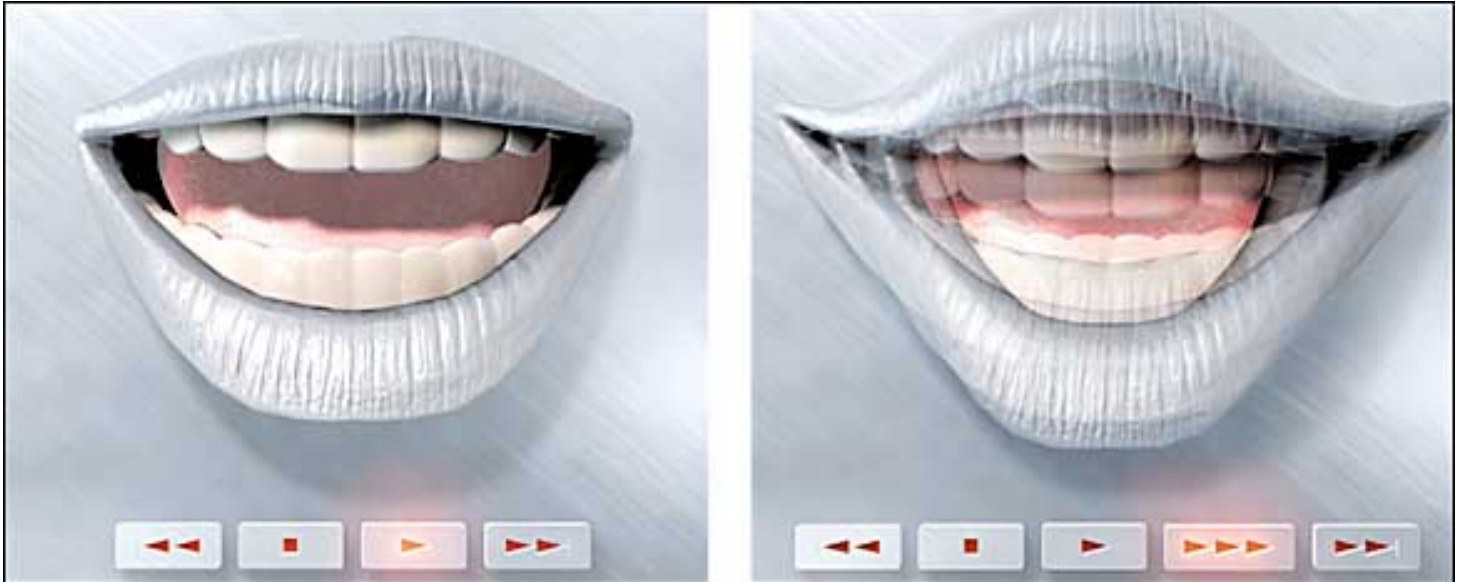# Now Hear This, Quickly

Mirko Illic

By DOUGLAS HEINGARTNER

Published: October 2, 2003

"WE call it the 66-second minute," Laura Gaines said.

Ms. Gaines is the vice president of Prime Image, a maker of devices like the Digital Time Machine that shorten audio and video recordings by up to 12 percent with "no discernible results." Micro-editing, as the process is called, created a stir last year when some broadcasters were reported to be using the technology to squeeze more advertisements into the same block of time.

As it turns out, it was hardly an isolated phenomenon. Creating more time is the impetus behind many new technologies that allow listeners to pick up the pace.

From call centers and intelligence agencies to radio stations and universities, such technology helps listeners try to keep up with the growing number of audio recordings piling up on the air, on the phone and on the Web. Wading though this mountain of words faster than it takes to say them not only saves companies money; it might help people absorb more knowledge.

The new software programs, DVD players and phone services rising to this challenge all take advantage of the human ability to comprehend speech much more quickly than the typical spoken rate of 140 to 180 words a minute. How many times as fast? "I've heard of instances where people go to 4X, and they still want it to go faster," said Blake Erickson of Telex Communications, which makes "talking book" audio players for the educational market.

Scientists have long known that people can understand speech at a rate of up to 400 words a minute and beyond. "Speech rate isn't limited by the listener," said Arthur Wingfield, a psychology professor at Brandeis University. "It's limited by the speaker."

In normal conversation, only a small part of the brain is taxed, leaving excess processing power to be used for listening for lurking predators, filtering out background noise or simply daydreaming.

But speeding up speech on analog equipment like cassette decks traditionally led to the dreaded chipmunk effect, making long-term listening untenable. Digital time compression, however, works by discarding tiny segments of repetitive audio (for example, 30 milliseconds of a vowel) and reconnecting the remaining bits, leaving the pitch unaltered.

Simple versions of digital time compression have been available for years in devices like answering machines and hand-held recorders but did not offer much in terms of user control. A confluence of smart software, wider Internet access and inexpensive hardware, however, now enables listeners to choose when to step on the gas.

Auxiliary programs, or plug-ins, that allow digital audio and video recordings to be played faster (or slowed down) at will have recently become available for popular software like RealOne and Windows Media Player. Perhaps the most popular is Enounce's 2XAV plug-in (which works with both Real and Windows players and costs $29.95); the latest version of Windows Media Player offers a proprietary version of this feature. Similar capabilities are finding their way into other hardware - for example, the latest DVD recorders from Panasonic.

"You can watch a two-hour movie on a one-hour flight," said Chris Binace, an Enounce software developer. Yet this kind of software is not generally intended for entertainment listening. So far most end-user applications have involved academia, for example, allowing students to listen to archived audio or video lectures.

Online, the amount of recorded audio is growing at an overwhelming rate, providing a new impetus for speed listening. A spokeswoman for National Public Radio said that demand for NPR audio on the Web was about 50 percent greater in June than it was a year earlier, and now averaged 5.5 to 7 million audio downloads a month.

"You just have oodles of data," said Ed Rucinski, a vice president of the Dictaphone Corporation, "and if you can only listen to it in a real-time fashion, that's your

bottleneck." Mr. Rucinski's company records "literally millions of hours" of audio every year: medical dictation, emergency calls to 911 centers, even financial transactions. "Any time you call your broker," he said, "that gets recorded."

One company addressing the deluge is Fast-Talk Communications, which makes software for large businesses that scours voice and audio data much the way search engines sift through text. Many Fast-Talk clients work in intelligence. "But there's a limited number of linguists," said Bob Crochetiere, a Fast-Talk sales engineer, so companies have to find ways of processing this material more efficiently. Mr. Crochetiere said clients would often listen to audio at speeds increased by as much as 50 percent, but only in bursts because after too much fast listening, "they start zoning."

Hannah Hawkins, transcription manager for CCBN, a company that records and archives hundreds of lengthy conference calls each week for the financial industry, said, that speed was crucial. Clients need the transcripts as soon as possible after the call is finished, so CCBN transcribers sometimes double the playback speed of familiar portions like introductory legal disclaimers.

"If they're speaking very slowly," Ms. Hawkins said, "you can understand them perfectly" at accelerated speeds.

Richard Brownrigg, a general manager at [RealNetworks,](RealNetworks) which makes the RealOne media player, said that fast playback was still in its early days,but that he could imagine its value expanding as voice technology crossed into new areas. Playing back long cellphone messages in half the time, for example, becomes attractive "when people don't want to chew up their minutes," Mr. Brownrigg said.

In advertising, where costly post-production of commercials can take longer than the production itself, the potential savings are vast. "To edit a 30-second spot can take half a day," said Ms. Gaines of Prime Image, but takes just minutes with the company's technology. (She hastened to point out that the compression was intended to enable advertisers to say more in the same period of time, not to let broadcasters shortchange the advertisers.)

Most research has indicated no loss of comprehension or intelligibility at playback speeds of two or even three times normal speed. Cameron Earle, who is helping to commercialize variable-speed playback applications developed by Brigham Young University, said that most students chose rates that were 80 to 120 percent faster than normal with no decrease in test scores. Although it does take some getting used to, Mr. Earle said, he estimates that "80 percent of acclimation is in the first hour."

Perhaps even more significant, the technology may have benefits beyond saving time and money. "People who are listening at accelerated speeds learn just as much, and there's some evidence they may learn even a bit more," said Kevin Harrigan, an associate professor at the Center for Learning and Teaching Through Technology of the University of Waterloo in Canada. The consensus is that the extra brainpower needed to follow

speedy speech enhances comprehension. "If you're listening at accelerated speeds," said Joel Galbraith, a researcher in Penn State's instructional systems program, "it forces you to not do anything else, so you're more focused on it."

Ray Juang, a University of California undergraduate who would often fall asleep in Berkeley's vast lecture halls, agrees. "On average, I understand the material better during playback than in the actual lecture room," Mr. Juang said. "The speed-up does force me to pay more attention."

Accelerated speech also piques interest. A quarter-century ago, Priscilla La Barbara, a marketing professor at New York University, found that time-compressed radio advertisements were perceived as more interesting and led to higher rates of recall.

But the days of those fast-talking radio announcers ("3.7 percent A.P.R.," "void where prohibited") may be numbered: Esther Janse, a post-doctoral researcher at the University of Utrecht, has found that digitally accelerated speech is more intelligible than the natural speech of a person talking rapidly. "When you try to speak faster and faster, speech gets very blurred," Ms. Janse said. The distinctions fade, she said, whereas digitally accelerated speech uniformly preserves all the crucial intonations and inflections.

There are other examples of how machine-altered speech may trump that of humans. Professor Wingfield of Brandeis said that airplane pilots had been shown to pay greater heed to warnings issued by computerized voices than natural human recordings. "When one of these hokey synthesized computer voices says to pull up," he said, "it's like, 'Oh, well, that's a computer. It must know better than I do.' "

Synthesized accelerated has many other devotees. "When I listen to the newspaper, I tend to go as high as 650" words per minute, said Gregory Rosmaita, a Web designer based in Jersey City. Because Mr. Rosmaita is blind, his interface with computers is audio-based, in the form of a synthesized voice that reads text aloud. He prefers British English to American in this regard. "With the more clipped British speech," he said, "I can increase the rate even faster."

He said he had become so accustomed to accelerated speech that normal rates could sound unnatural. "It's actually difficult to comprehend the speech when it becomes that slow," he said. "It's sort of like watching a marquee scrolling one letter at a time rather than one word at a time."

Some users compared it to going back to dial-up Internet access after experiencing broadband. "I cannot stand to listen at 1.0," said Mr. Earle of Brigham Young. Mr. Galbraith of Penn State agrees. "Once you go faster, you just can't go back to real time," he said.

There are some caveats: for example, the capacity to understand fast speech seems to fade with age. "The younger the person is, the faster they can go," said Mr. Earle, who

said he had noticed a drop-off around age 30. "Professors can never go as fast as the students. Students can crank it out."

Few question that rapid playback saves time. "There's no doubt, absolutely," said Patrick McClanahan, a Navy lieutenant commander who used variable-speed playback while earning his master's degree in business administration at the Wharton School. Commander McClanahan said he most appreciated the ability to find a crucial point in a recorded lecture. "It's virtually impossible to slide that little thing across and find exactly what you want," he said of the cursor in audio playback software. Variable-speed playback eliminates the need to do so.

Mr. Juang, who as a Berkeley undergraduate has sometimes watched six two-hour lectures a day, said that even with occasional buffering delays and the need to replay bits that went by too fast, "an hour takes 35 or 40 minutes at most."

So as fast listening becomes commonplace, will more people turn into fast talkers?

"We're used to hearing things faster, so it probably translates into our talking as well," Mr. Galbraith said. "We'll start conditioning ourselves to just expecting and needing it faster."

Professor Wingfield of Brandeis is not so sure. "Knights were jousting with the same brain that we're using today," he said. "The articulatory system, the physiology of speech has not changed."