HEARING THE SIGHTS! SEEING THE SOUNDS!

Bringing 3-D Animation to the Classroom

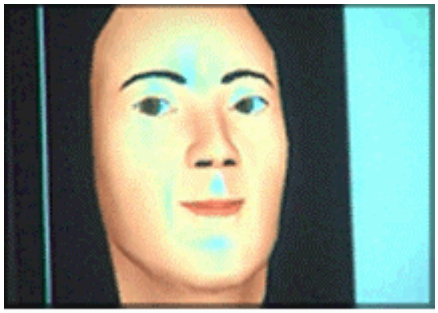
Tressa Steffen Gipe

Despite the shaky economy and gloomy industry forecasts, telecommunications and information technology chatter still focuses predominantly on tangible technological advancements such as building better fiber optic networks, creating greater bandwidth and rolling out new mobile devices. Far fewer headlines are made, however, by equally significant, albeit less dazzling applications of the latest technologies for the physically and learning impaired. In particular, remarkable progress has been made in the field of 3-D computer-generated animation, which can help people with autism, profound deafness, stroke-related aphasia and linguistic learning disabilities improve pronunciation, lip-reading, vocabulary, language fluency or other linguistic skills.

Three-D animation has become common in amusement parks, cartoons, video games and TV commercials, but until recently, expensive production costs kept the broader use of interactive 3-D animation to a minimum. Through research and development in both the private sector and academia, costs for 3-D animation programs have diminished with the introduction of PC-based programs that mimic the complex oral dynamics and subtle facial expressions of a human being. Many of these programs can be run on most 95/98/00 Windows and Linux operating systems, and are accessible to anyone with 64 MB RAM and over.

Bringing a New Dimension to the Classroom

At the Tucker-Maxon Oral School in Portland, Oregon, 3-D animation is being applied in the form of a computerized language tutor named Baldi. A wide-eyed, glossy-cheeked fellow with rosy lips, the unassuming Baldi holds the distinction of being the most accurate computer-generated copy of visible human speech in the world. Developed at the University of California, Santa Cruz, Baldi was designed on a principle that is intuitive to anyone who has begun a foreign language or taught a child to read: people learn language best by watching, hearing and doing.



Profoundly and partially deaf children at Tucker-Maxon apply this same principle by watching Baldi form words, which they mimic with the assistance of Baldi's real-time feedback. Baldi is also capable of showing internal teeth, tongue and jaw movements to bring the most natural approximation to human speech possible. When using Baldi, many children are assisted by headphones connected to an acoustic nerve implant that converts sounds into electrical stimuli, which are then relayed to parts of the brain responsible for decoding sound and language. Baldi can also be adapted to each student's needs depending on the level of development or severity of disability. Coupled with supplemental language training, Baldi is yielding impressive improvements in lip-reading, pronunciation and fluency skills for children at Tucker-Maxon and elsewhere.

Clearly interactive 3-D programs like Baldi have applications to other children with disabilities. Animated 3-D programs are also helpful to children with social disorders such as autism, which commonly manifests an inability to develop normal social relations, communicative abilities and, in some cases, visuospatial skills. In particular, many autistic children will use single words, phrase sentences as questions, engage in echolalia (speech imitation) and echopraxia (movement imitation), all of which create awkwardness for social interactions with others. However, 3-D programs can at least provide vocabulary-building, facial expression and pronunciation exercises to facilitate their interactions with the outside world. Three-D programs can also be used by people with

dyslexia, stroke-related aphasia and common speech impediments to improve language skills.

Student at Tucker-Maxon Oral School interacting with Baldi (http://www.tmos.org)

The Future is Now

The private sector originally generated most of the 3-D technology that is used in learning settings like Tucker-Maxon. There continues to be significant innovation in the 3-D animation world that will likely lead to more sophisticated educational applications in the future. Most notably, companies are beginning to market programs that blend speech recognition and synthesis with 3-D animation to create a highly interactive computer learning environment. Some speech synthesis and recognition programs are so sophisticated, in fact, that they can possess anywhere from 5,000 to 50,000 vocabulary word repertoires that do not require training for different human voices.*

A company called Sensory has developed a Fluent Animated Speech program that uses text and spoken words to elicit a response or simulation through 3-D animated characters. With a speech recognition and synthesis component, sound is usually transferred via microphone to the sound card, which then is processed into sonic frequencies using mathematical and statistical methods. The frequencies are checked at regular increments (say, every 1/100th of a second) and then analyzed by the speech recognizer for patterns that fit the smallest block of human speech called a phoneme. The phonemic components are compared with a word database and, ideally, the word is recognized or synthesized by the system. The 3-D animation component is created using a set of static pictures that can be combined in unlimited combinations to form the appropriate facial expressions, lip movements and sounds to accompany a replicated or generated stream of words.

Of course, many analysts foresee possibilities that extend far beyond the rather small market that currently exists for 3-D animated language technologies. Many predict that interactive 3-D animated speech recognition/synthesis technologies will become commonplace in personal computers, bank machines and drive-throughs.

As the technology develops, it will certainly have broader applications in educational and business settings where teaching and training are conducted. For instance, a business could use a 3-D tutor to train employees in basic business Japanese before they go abroad to present products. A school could train students to expand vocabulary through interactive drills with a 3-D image that gives feedback with synonyms, antonyms or examples of the word in a sentence. All the possibilities sound exciting, but it remains to be seen if people will be comfortable with non-human entities interacting with them on such a sophisticated level.

Predictions aside, the fact that complex technologies like interactive 3-D animation are being used to make people's lives quantitatively and qualitatively better speaks to the importance of continued innovation in even the most fledgling technological markets. Beyond the headlines, video games and theme parks, 3-D animation demonstrates that creative application of commercial technologies for people with (or without) disabilities can open up an entirely new dimension in computer-assisted learning: the human one.

*The large-vocabulary models are not necessarily available with 3-D animation, however.

http://more.abcnews.go.com/sections/primetime/2020/primetime_010315_baldi_feature.html#2
http://www.tmos.org
http://mambo.ucsc.edu/
http://www.indiatimes.com/infotech/help/software/voicereq.html
http://www.idg.net/english/crd_recognition_173310.html