ReadSpeaker ◄))

pioneering voice technology

The Promise of Synthetic Voices to Improve Learning Outcomes

In today's technical landscape, artificial intelligence, virtual humans and voice technology are taking on an increasingly important role in education technology. Historically synthetic, or computer-generated voices, have been seen as inferior to human voices for learning results. However, recent studies have shown that with the continual advances of voice technology, when paired with a virtual human, modern synthetic voices can actually produce better learning results than either human voices or old text-to-speech engines.

According to the study *'Reconsidering the voice effect when learning from a virtual human'* carried out by Scotty D. Craig from Arizona State University and Noah L. Schroeder from Wright State University, **"the modern voice engine produced significantly more learning on transfer outcomes, had greater training efficiency, and was rated at the same level as an agent with a human voice for facilitating learning and credibility while outperforming the older speech engine. These results call into question previous results using older voice engines and the claims of the voice effect."** [1]

As technological innovations are reaching the classroom, there is ever-growing research into effective design and implementation of educational technologies. And in general, it has been found that learning technologies become more effective when virtual humans, or on-screen, human-like characters, are used. [2]

Virtual humans are used in multimedia learning environments and intelligent tutoring systems as instructors, characters in educational video games or as pedagogical agents. These characters help in the learning process by signaling, motivating, role-playing, as a facilitator or by modeling learning strategies.

> "The modern voice engine produced significantly more learning on transfer outcomes, had greater training efficiency, and was rated at the same level as an agent with a human voice for facilitating learning and credibility while outperforming the older speech engine."

As can be reasonably deduced, researchers have shown that the design of the pedagogical agent, meaning its voice, speech patterns, or appearance, influences how effectively the agent facilitates learning [3]. These results highlight the importance of "purposeful, data-driven agent design." [4]

According to what is known as the 'voice effect', or 'voice principle', "learning will be improved when a "standard-accented" recorded human voice provides the narration during a multimedia learning situation rather than a computer-generated voice, or so-called "machine voice." [5]

Mayer found compelling evidence to support this conclusion. However four of these studies were published at least 10 years ago. Since then voice technology has been advancing rapidly, and text-to-speech software has greatly improved.

Craig's and Schroder's 2017 research, *'Reconsidering the voice effect when learning from a virtual human',* looks at the implications of the voice effect paired with a virtual human on learning outcomes, cognitive load, and perceptions of the agent.

Historically, researchers have thought that learning with an artificial voice put an additional cognitive load on the learner, as well as caused distractions. Early research by Mayer et al in 2003 and again in 2005 showed that human voices outperformed synthetic voices. However, the results of a similar study in 2012 by Mayer and DaPra using more advanced voice technology indicated no differences in learning between groups that had agents with human voices or those with synthetic voices, pointing to the fact that perhaps the voice effect no longer existed.

Craig & Schroder used Microsoft's speech engine as a classic engine as a baseline and **NeoSpeech's (now under the ReadSpeaker brand) 'Kate' voice** as the representative of a modern engine. A human voice was used as a high-end control. All three of these voices were given to a female virtual human.

A random selection of participants were evaluated on perception, cognitive load, multiple choice questions, and retention. "For the first (learning) and second (cognitive load) research questions, consistent results were found that either showed no differences between conditions or demonstrated that the presentation by the agent with a modern voice engine was more effective compared to the older voice engine or the human voice. This provides consistent evidence against the voice effect." [6] No statistically significant differences were seen on the multiple choice and retention learning measures and the other efficiency measures.

In can be concluded that the type of voice used when comparing modern text to speech or recorded human voices, is not as important for learning outcomes as once assumed, and modern voice engines may be just as effective as a recorded human voice. Similarly, no differences were seen in a participant's ratings of the agent's ability to facilitate learning and perceived credibility.

While Craig & Schroder's study using more advanced voice technology not only debunks the myth of human voices being superior in learning environments, it also points to the fact that modern synthetic voices can even produce better results than human voices.

It is possible that the long-standing idea for virtual humans to improve learning is currently possible, and will continue to a greater extent in the future. [7]

(1) - Craig & Schroeder, 2017

(2) - Dehn & Van Mulken, 2000; Graesser, McNamara, & VanLehn, 2005; Graesser & McNamara, 2010; Johnson & Lester, 2016

(3) - Baylor & Kim, 2004, 2009; Clark & Choi, 2005; Domagk, 2010; Kim & Wei, 2011; Moreno & Flowerday, 2006; Ozogul, Johnson, Atkinson, & Reisslein, 2013; Schroeder, Romine, & Craig, 2017; Veletsianos, 2010

(4) - Craig & Schroeder, 2017

(5) - Mayer, 2014b, p. 358

(6) - Craig & Schroeder, 2017

(7) - Johnson & Lester, 2016

ReadSpeaker 🔊

pioneering voice technology

ReadSpeaker is a global voice specialist providing dozens of languages and lifelike voices. Using its own industry-leading technology, the company delivers some of the most natural-sounding synthesized voices on the market. ReadSpeaker uses next-generation Deep Neural Network (DNN) technology to structurally improve voice quality at all levels. ReadSpeaker is a subsidiary of the Memory Disk Division (MD) of the HOYA Corporation, with offices in 15 countries, and over 10,000 customers in 65 countries, providing a complete text-to-speech (TTS) offering, both as Software-as-a-Service (SaaS) and as licensed solutions. A fully integrated TTS provider, ReadSpeaker encompasses all of HOYA's state-of-the-art technologies (NeoSpeech, Voiceware, VoiceText and rSpeak), providing a wide variety of applications for varying channels and devices in multiple industries. ReadSpeaker gives a voice to businesses and organizations for online, embedded, server or desktop needs, apps, speech production, custom voices and more. With more than 20 years' experience, the ReadSpeaker team of experts is leading the way in text to speech. ReadSpeaker is "Pioneering Voice Technology".

✉ contact@readspeaker.com

🌐 www.readspeaker.com

in company/readspeaker

f readspeaker

🐦 @readspeaker

To learn more about ReadSpeaker, contact us at www.readspeaker.com/contact-us