

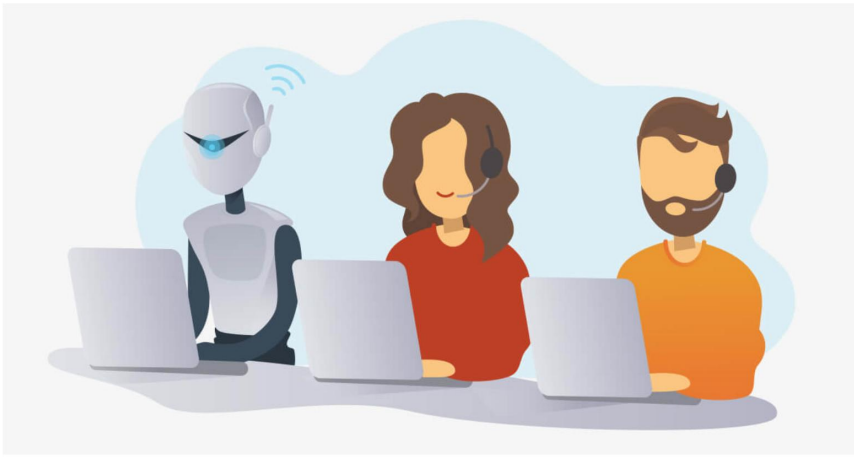


September 13, 2022
By **Rupal Patel** Vice president of voice and accessibility at Veritone
Industry Voices



Site Sponsors

AI Voice: Working with, Not Against, Humans



Selling artificial intelligence (AI) has always been a challenge since many think it's a replacement for humans. But it's our responsibility in the industry to educate people on how AI enhances human capabilities and removes the mundane. It's not a replacement for human creativity and ingenuity. That's especially true for new applications of the technology, such as AI voice.

It's similar to how people felt about autotune when it was first introduced in the music industry. Initially, people were put off by autotune and how it could sound tinny, but once it was compiled and integrated with other musical elements, people were able to discover a completely different kind of sound that could never be achieved with human voice.

Autotune is only as good as the input it's given: the artist still needs to be a good singer, allowing autotune to finesse the sounds, for the music to truly deliver. The same can be said for the training data required to create an authentic-sounding AI voice.

Today, autotune is a widely accepted tool inside a musician's toolbox, and if we play our cards correctly, AI voice could follow a similar path. Both are tools to enhance and augment human voices to help them reach newer potential.

We saw a similar sense of worry that text-to-speech (TTS) was trying to replace talent, when in reality that's not the case. AI voice cannot thrive without talent. It needs human talent to achieve realistic-sounding and convincing audio. This isn't competition between man and machine; this is a symbiotic relationship. AI needs human voice training data to create a voice clone; and at the same time, voice talent can open up new revenue streams and discover new ways of creating content.

Bev Standing, professional voice artist and the original TTS voice for TikTok, recently said, "Stock voices are great for many projects, like training or academic materials. When it comes to projects such as audio descriptions, scene narration, and radio imaging, a more humanistic voice should be used. And that can be the AI voice of a professional voice-over artist. It's a winning setup for modern voice-over artists and content creators: my AI voice provides additional revenue streams that can work for me while I'm away or doing other jobs. There's room for both."

Reframing how we refer to these voices can be key to easing trepidation by naysayers. As an industry, adopting new verbiage could help drive our point, but we might want to look toward alternative descriptors instead of synthetic or artificial, which can evoke a sense of something that's fake, inauthentic, or less desirable.

There's still a human in the loop. It's not entirely inorganic. Perhaps reframing the tech as amplified intelligence, accelerated intelligence, or automated intelligence would be the key to unlocking greater acceptance.

Achieving Humanlike Conversations

>Voice is such a powerful tool that goes beyond merely conveying information. It helps us connect with audiences and ensure that our messages can be felt, rather than simply heard. With the enhanced capability of AI, companies can create messaging that prioritizes personalization rather than merely broadcasting content, to push the boundaries of what is possible with this technology.

AI can accelerate production time and volume, allowing companies to churn out more audio content than ever without more studio time from the voice actor. And it's not just volume in terms of projects; with AI, the voice can adapt to different languages and dialects while staying on-brand.

Even more exciting is that given today's life-like quality, it is possible to create hybrid media from human and automated voices to craft more harmonious, authentic experiences for the people we're trying to reach. Our goal moving forward should be to keep the listener in the loop and come from the idea of speaking *with* them rather than speaking *to* them.

Calling back to our autotune analogy, many people have a dated notion of TTS as sounding choppy, inauthentic, and sterile, but the technology has come so far since the first iterations. Even the terminology has shown improvement. Using voice rather than speech imparts a feeling of more humanity and identity.

We're no longer in a place of trying to replicate the speaker. Instead, we can enhance the speaker. Rather than having 15 different AI voices, we can have 15 versions of one voice, each one for different audience preferences around language, accent, speed, tempo, and more.

A lot of the prominent examples of AI voice involve the media and entertainment industry, but it has its uses across all industries. Take for example the automated prescription pick-up reminder at big-name pharmacies. We've all heard how clunky and inflexible those reminders can sound, and they can at times be difficult to understand. The pharmacy's goal is to provide care, so why not extend that experience with a more human AI voice experience customized for customers' needs? If the call is going out to a senior citizen, the pharmacy could choose a voice that speaks slower and louder, for example. These simple ways in which to connect with the end user are key to improving the voice experience.

As the world continues to build acceptance for AI voice, it's important to continue to look for opportunities where AI and human voice can be used together and where TTS can be used in more creative ways than as an efficient means to broadcast information. From a speech technology perspective, we can continue to discuss its use with companies the way we'd address any other tool: as a way to present a new world of possibilities, both for employee assistance and better customer connectivity.

ALSO ON SPEECH TECHNOLOGY MAGAZINE

Speech Technology Has a Place in Ed ...

4 years ago · 1 comment

Technology advances in speech recognition, pronunciation error ...

What Does Scaling Mean for the ...

a year ago · 1 comment

Speech-to-text services have gained accuracy, but scaling will require ...

Call Center AI Market to Be Worth \$4.1 ...

2 years ago · 1 comment

MarketsandMarkets eyes 21.3 percent CAGR for contact center AI through ...

10 Tips for Maximizing Content for Voice

8 months ago · 1 comment

A quarter of the world population uses voice search, and the num...

0 Comments

Login



Start the discussion...

LOG IN WITH



OR SIGN UP WITH DISQUS

Share

Best Newest Oldest

Be the first to comment.

Subscribe Privacy Do Not Sell My Data

DISQUS

Resources

- Home
- Speech Technology Magazine
- Digital Editions (PDF Download)
- Speech Technology Newsletters
- Speech Technology Topic Centers
- Speech Technology Industry Solutions
- Speech Industry Awards
- Sponsored Videos
- Reference Guide
- About/Contacts

Other ITI Sites

- Database Trends and Applications
- DestinationCRM
- Faulkner Information Services
- InfoToday.com
- KMWorld
- Online Searcher
- Smart Customer Service
- Streaming Media
- Streaming Media Europe
- Streaming Media Producer
- Unisphere Research