readspeaker.ai

Selecting The Right Voice For Your Interactive Voice Response System

The 5-Step Process

A ReadSpeaker Guide

Introduction

There was a time when companies treated interactive voice response (IVR) systems as a necessary evil—something that didn't have to provide a great experience, because, frankly, traditional call centers rarely did. Things have changed.

Today, every interaction with your customers has the potential to affect your brand reputation, positively or negatively.

Now, all customer service is branding. If a caller has a bad experience with your IVR system, they can tell the world, amplifying their critique through the megaphone of social media.

Meanwhile, technology has improved the IVR experience considerably. Advanced conversational IVR systems use artificial intelligence (AI) to route calls, share information, and even perform customer service tasks. These platforms are more dynamic than ever. Text-to-speech (TTS) allows users to update content nearly instantly, without costly, time-consuming recording sessions. More importantly, synthetic speech puts the "voice" in "voicebots," creating the key connection between a caller and the system's AI core—systems that can create a customized experience, which leads to greater consumer satisfaction in automated self-service channels like IVR.

In fact, throughout any IVR call, the only point of contact between the customer and your brand is the TTS voice you install in your system. In other words, **TTS defines the brand experience with every call**.

This guide will help you choose the TTS voice that makes that brand experience as successful as possible.

Voice Branding and the Conjured Persona

Hans van Dam is co-founder and CEO of the [Conversation Design Institute](#), a training academy for conversational-interface design. Presenting at the [2019 Chatbot Conference](#), van Dam made a striking connection between the eye's tendency to recognize faces where none exist—*pareidolia*, it's called—and the ear's conjuration of personas when confronted with speaking voices.

"We all know it's not a face, but we see a face in it," van Dam said. "It's the same with voices. If we hear a voice, within one second, we create an image of who that voice belongs to."

It's a helpful comparison—and if the mind constructs characters when confronted with voices, the careful marketer will control the creation of that persona as much as possible. That requires choosing a voice and speaking style that guide the caller's vision of your voicebot—**a vision that is no less than a personification of your brand**. Put another way, your TTS voice is a literal brand voice, an asset that represents *who you are* just like your logo, color schemes, and audio watermarks like tones or jingles. That's why it's so important to get the speaking voice right.

But what exactly does *right* mean in this context? Follow these five steps to find the TTS voice that's ideal for your unique conversational IVR system.

5 Steps for Choosing a TTS Voice for Conversational IVR

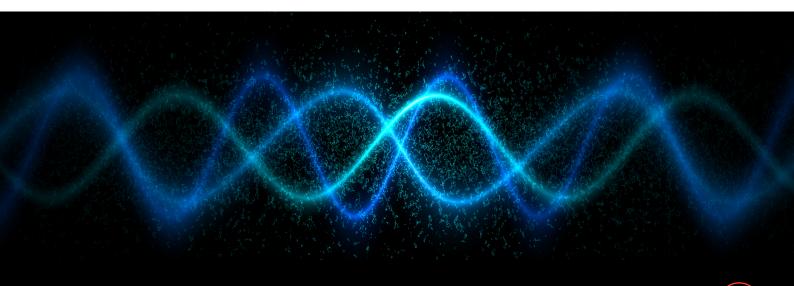**1**   Identify every possible use for TTS in your IVR system.

When choosing a TTS voice, it's important to be clear about the customer scenarios that voice will discuss. You wouldn't want a cheerful, cartoonish persona reporting a caller's chronic debt, for instance. And because even conversational IVR systems typically divide work between the AI and human agents, it's crucial to know who will be doing what once you deploy your system.
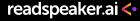
E-commerce platform bol.com worked with ReadSpeaker to develop a custom branded TTS voice for a Google Assistant implementation, but the lessons of this case study also apply to TTS for IVR systems. The bol.com voice assistant responds to customer questions about, for example, daily deals, gift ideas, and delivery status. These types of questions—and the bol.com brand identity—require a branded TTS voice that projects kindness, helpfulness, and trustworthiness.

Those are different needs than those of an energy drink brand. Say such a company operates an IVR system that only uses TTS to answer the line and collect standard customer information. In this scenario, you'd probably be better off with a hyped voice that reflects excitement first and foremost. Those are two very different examples of how use cases determine optimal voice quality for different brands. The point is, consider *what* your IVR voicebot will be discussing before making any decisions about *which* TTS voice to integrate. **Voice *form* and voicebot *function* must align.**

2 Consider your audience and their preferences.

What are the demographics of your most frequent callers? Many providers choose TTS voices that match the communities they serve. That includes matching languages and dialects, of course, but could also affect dimensions of voice that suggest gender, age, or even personality.

But choosing a TTS voice isn't always a question of matching customer demographics perfectly. One provider of pension funds has a customer base of mostly older adults. Rather than choosing a TTS voice that matches this age range, that company realized their callers would prefer a voice that reads as slightly younger. When you A/B test your voices (step 5 below), you can ask customers for their preferences directly—but having these discussions about caller expectations early can help identify the appropriate voice characteristics.

Gendering Voicebots: A Shifting Landscape

When voice AI technologies emerged in force in the 2010s, all the major U.S.-based voice assistants defaulted to a female voice: Amazon's Alexa, Microsoft's Cortana, Apple's Siri, and Google Assistant. Similarly, as many as 75% of IVR systems reportedly use female voices.
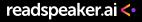
Why is this the trend? "We carried out research and found that a woman's voice is more 'sympathetic' and better received," Daniel Rausch, head of Amazon's Smart Home division, told Business Insider.*

But critics argue that *always* gendering voicebots as female reinforces sexist stereotypes of female servility. A 2019 United Nations special report warns that "because the speech of most voice assistants is female, it sends a signal that women are obliging, docile and eager-to-please helpers, available at the touch of a button or with a blunt voice command like 'hey' or 'OK'."

Some tech enterprises appear sympathetic to these arguments. In 2021, Apple stopped defaulting to the female Siri voice in English-speaking markets. Instead, users will choose between a range of voices, coded both male and female.

This debate is also leading to calls for more gender-neutral TTS voices, potentially even *clearly robotic* voices. Some consumers may prefer TTS voices they instantly recognize as machine speech, especially as conversational AI systems become indistinguishable from human speakers.

*In fact, the research on gendered voice preference is all over the place. One study suggests that both men and women prefer lower-pitched voices; another says people generally prefer voices of the opposite sex; still another found that listeners preferred female voices by 12.5% when listening to TTS speech, but preferred male voices by 2.3% when listening to recordings of voice actors. The safest assumption is that gender preference will vary by audience; your best bet is to know yours well.

3   List the brand traits your voice should convey.

Before you can find the right voice, you must know what you're looking for—
and what you're looking for is an audio expression of brand traits. If you haven't
officially spelled out your brand traits yet, look into Jennifer Aaker's seminal work
on the "Big 5" dimensions of brand personality. (Aaker defines brand personality
as "the set of human characteristics associated with a brand.") Each of Aaker's
five dimensions is associated with a handful of specific traits, as demonstrated
in this table (adapted from Aaker's 1997 *Journal of Marketing Research* article
"Dimensions of Brand Personality"):

Personality Dimension	Brand Traits
Sincerity	Cheerful, genuine, honest, domestic
Excitement	Up-to-date, imaginative, spirited, daring
Competence	Efficient, dependable, responsible, reliable
Sophistication	Romantic, charming, glamorous, pretentious
Ruggedness	Strong, tough, outdoorsy, rugged

You can use Aaker's concept to construct a model of your brand personality, which
may have some part of each dimension, drawing traits from them all. A brand like
Patagonia, for instance, may be described in terms of Sincerity (genuine, honest),
Competence (dependable), and Ruggedness (all four traits).

Here's what that means for choosing a TTS voice: These traits can map directly
onto acoustic and prosodic qualities of your speaking persona. For Patagonia, that
could mean a slightly rough, low-toned, male-coded voice.

How about you? Once you have your brand traits identified, you have an idea of the characteristics you'd like to hear, regardless of the perceived age or gender of the TTS voice speaker.

4 Listen to a wide variety of examples.

Once you have an idea of what you're looking for, the fun begins. Listen to off-the-shelf TTS voices to find a match. (Find samples of ReadSpeaker TTS voices here, browse this page for a fuller list of languages and voices, or contact us with your preferences for a curated list of potential TTS voices.) Gather stakeholders together and try to reach consensus on two or three TTS voices that express your brand traits and your audience's preferences.

You don't have to make a final decision yet, though. Once you have your choices narrowed down to two or three, A/B testing will help you make a data-driven final decision.

5 Test your choices.

For a conversational IVR system, A/B testing is simple to set up. ReadSpeaker's API plugs directly into your conversation management system (CMS), providing online access to the voices you're testing. Or you can use ReadSpeaker's web-based applications to convert text files into synthetic speech.

Either way, stakeholders can have actual conversations with the system, just as callers will. If possible, convene a panel of your customers to test the voices. After each conversation, ask testers to evaluate the voices to compare results. At the end of this process, you'll have a TTS voice that's ready for deployment—although eventually, most companies will move on from ready-made TTS voices to a custom TTS voice that represents their brand perfectly.

IVR and Beyond: Custom Brand Voices from ReadSpeaker

Conversational IVR is just one of the voice-first channels available for connecting with customers. You might find value in a branded AI assistant through your app or website; you may develop a custom Alexa skill; or you may run an interactive voice ad through Pandora or Spotify. Better yet, you'll take advantage of *all* these voice platforms—and more.

To get the most out of such a unified voice branding strategy, consistency is key. Just as you wouldn't use six contrasting versions of your logo in six different print ads, you mustn't allow your brand's literal voice to vary from one experience to the next.

The good news is that the above steps won't just help you pick a voice for a conversational IVR system. You can follow the same process to choose a TTS voice that represents your brand traits across all your voice touchpoints. Best of all, once you prove the cost savings associated with an off-the-shelf TTS voice, it will be easier to get across-the-board buy-in for a custom branded TTS voice that works across platforms.

The ReadSpeaker VoiceLab works with you to create a one-of-a-kind TTS voice, using our proprietary deep neural network technology. Because we retain original training data, all ReadSpeaker custom voices are future-proof; as we push the quality level of TTS ever higher, we can reprocess original data for consistency alongside continuous improvement.

Multilingual TTS allows your brand voice to speak in multiple languages, for cross-market consistency. Emotional TTS speaking styles allow your IVR AI to respond appropriately in different situations. And ReadSpeaker's global team of computational linguists provides ongoing support, ensuring perfect pronunciation of even industry jargon and proper nouns.

[Ready to find the ideal TTS voice for your conversational IVR system? Contact us to get the process started.](#)