![Cerence logo]

Diversity in Voice Assistants: Adjusting Tone and Timbre to Deliver a Genderless Voice

1月 31, 2020

As voice assistants have become more prevalent in recent years, it's become clear that female-toned and -named voices are the most commonplace among their developers – think Alexa, Siri and Cortana, to name a few. When asked about the reason for this choice, developers often mention studies stating that female voices sound more pleasant and helpful than male voices. Apart from the difficulty to prove such a claim, there has been criticism that the choice of female voices for virtual assistants reinforces gender stereotypes and holds up a standard that women should live up to or can be compared against. It also makes it harder for companies to enter the market with a male voice assistant, almost like being the first male employee in a female-dominated workforce.

Another reason why companies hesitate to introduce male voice assistants is that it requires substantial effort to develop a high-quality computer voice and perfect it for the intended application. Companies want to avoid doing this work twice, or even more times if they want to offer a variety of voices with different ages, accents, personalities, etc. There is also a branding aspect, where the company wants the default voice to be recognizable and consistent with the product positioning.

In our 20 years building voice assistants for the world's leading automotive brands, we have developed hundreds of both female and male voices, ranging from portfolio voices available to multiple customers to highly customized, exclusive voices for individual brands. Recently, we developed a new text to speech feature that allows an application developer to control the timbre of a computer voice, changing the perceived age or physical size of the speaker. With this feature, it is also possible to make a female voice sound more masculine and a male voice more feminine. If the feature is dialed to a moderate setting, the voice sounds between female and male, or uniquely "genderless." This enables us to support our customers by giving them more choices for their voice assistant products and the ability to consider alternatives to the traditional choice of male or female.

We ran listening experiments with crowd-sourced listeners to determine the likeability and perceived gender of two female computer voices with different settings of the timbre parameter. As the parameter value increases, the number of listeners that perceives the originally female voice as male increases to 80%, while the remaining 20% of listeners characterize the voice as genderless. The pleasantness of the voice drops slightly as the amount of signal modification increases. One effect that does not happen, is that the pleasantness first drops as the voice becomes less clearly feminine and then increases again as it becomes more male sounding. The crowd-sourced listeners do not prefer voices that are clearly gendered over genderless voices.

The timbre parameter can be used by developers to offer a different default than the friendly female voice. The modification is performed at runtime and does not involve substantial cost, computation, or footprint. The timbre parameter can also be used by end users to tune the voice to their liking, making it sound younger or older, more feminine or more masculine, or genderless.

The human experience is only becoming more dynamic, futuristic and ripe with creative opportunities. In today's world of increasingly prevalent voice assistants, we believe that options are key. Check out the samples below, and see for yourself how these different voices bring different experiences to life.

Original	Genderless	Female to Male
Ava Original	Ava Genderless	Ava Female to Male
Zoe Original	Zoe Genderless	Zoe Female to Male