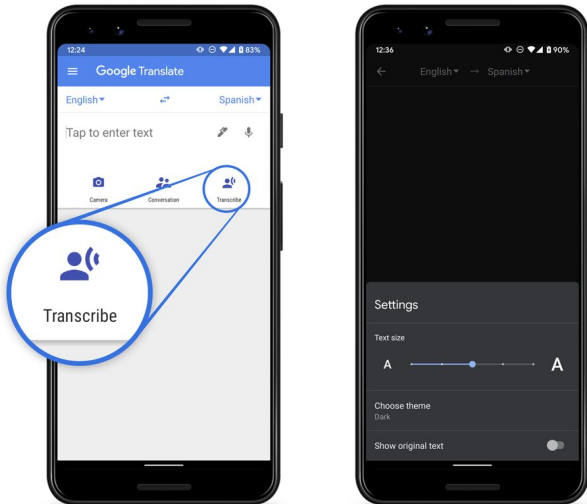


Google introduces real-time extended voice translation

18 March 2020, by Peter Grad



Credit: Google

Google has announced a new real-time transcription feature for its free Translate app for Android phones. An IOS version is planned for the future, the company says.

The feature will allow users to obtain instantaneous text translations of ongoing speeches, lectures or monologues into any of eight languages, including English.

Currently, Translate allows conversions of only relatively short snippets of [speech](#).

The only requirements are having only one speaker talking at a time in a quiet room (other voices or noises will diminish accuracy) and an Internet connection, necessary for interaction with Google's cloud-based Tensor Processing Units.

The rollout begins today (March 18) and should be available to all users by the end of the week at Google's Play Store.

In conversation mode, the app permits users to have a back-and-forth conversation with someone speaking a different language.

In addition to English, translations are available in French, German, Hindi, Portuguese, Russian, Spanish and Thai.

The app will also work with playbacks of prerecorded audio. But Google says direct digital translation from uploaded audio files is not yet available.

This week's announcement is a reminder of just how far we have come since the earliest days of digital voice [recognition](#). Bell Laboratories debuted its futuristic "Audrey" system in 1952 that recognized the spoken digits 0-9. A giant step was made a decade later when IBM displayed the "Shoebbox" at the 1962 World's Fair—it could recognize a whopping 16 words.

For five years in the 1970s, voice recognition got a huge boost from America's military. The Department of Defense underwrote massive research projects into speech recognition, including Carnegie-Mellon's "Harpy" Speech Understanding Research (SUR) initiative, which built a recognition vocabulary of more than 1,011 words. That program notably introduced the concept of pronunciation patterns and probability for the first time, greatly enhancing the ability to recognize distinct modes of speech.

The 1980s brought ever greater advances in word detection, with researchers applying probability theory to unknown sounds. Tech giant IBM's program expanded recognition to 5,000 words. But the decade may be best remembered for the introduction of the world's first talking doll, "Julie," that understood speech. An ad campaign stated: "Finally, the doll that understands you."

Dragon brought voice recognition to the masses in

the 1990s, with its first largely accurate though still buggy consumer product priced at "only" \$9,000. By the end of the decade, the vastly improved Dragon NaturallySpeaking program, which for the first time did not require pauses between each spoken word, was available to consumers for about \$700.

Today we have Siri and Alexa and other free and low-cost mobile apps that let us request driving directions, order food, buy household items and type out spoken text in emails and word processing documents, all of which have expanded speech recognition to points unimaginable not too many years ago.

With the latest advances available to millions of users with handheld devices, Harpy, Audrey, Julie would likely be left speechless.

More information:

www.blog.google/products/transcribe-speech/

© 2020 Science X Network

APA citation: Google introduces real-time extended voice translation (2020, March 18) retrieved 25 May 2022 from <https://techxplore.com/news/2020-03-google-real-time-voice.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.