

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

SpeakEasy Pronunciation Trainer: Personalized Multimodal Pronunciation Training

Permalink

<https://escholarship.org/uc/item/0458r7rc>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Tyler, Jeramey
Braasch, Jonas
Si, Mei

Publication Date

2021

Peer reviewed

SpeakEasy Pronunciation Trainer: Personalized Multimodal Pronunciation Training

Jeramey Tyler

Rensselaer Polytechnic Institute, Troy, New York, United States

Jonas Braasch

Rensselaer Polytechnic Institute, Troy, New York, United States

Mei Si

Rensselaer Polytechnic Institute, Troy, New York, United States

Abstract

The primary goals of computer-assisted pronunciation training (CAPT) systems are to provide a personalized interactive environment and to accurately diagnose mispronunciations. Automatic speech recognition (ASR) systems have been shown to be an effective tool for diagnosing mispronunciations. While the data ASR systems output can be difficult for the layperson to understand, presenting it in a multimodal fashion can make it easier and feeding it into an automated narrative system can produce personalized feedback. In the absence of native speech examples, synthetic examples produced by text-to-speech (TTS) engines have proven to be an adequate substitute, making data collection easier and allowing for larger CAPT systems. In this work we present the SpeakEasy pronunciation trainer, a CAPT system that leverages ASR, TTS, automated narrative systems, and multimodal data representation to provide a personalized interactive environment that tracks a user's progress over time.