Module 6

Unit 4: Text-to-speech

Core video transcript

Module 6

Unit 4: Text-to-speech

Transcript

Slide 1

Hello. I'm Anna Matamala, from UAB, and in this ADLAB PRO video I will talk about text-to-speech technologies. This is Unit 4 (Text-to-speech) in Module 6.

Slide 2

The voicing of an AD can be done by a describer, by a voice talent or by an automatic text-to-speech system. In this unit you will learn how a text-to-speech (or TTS) system can be applied to audio description (AD) and audio subtitling (AST).

Funded by the Erasmus+ Programme of the European Union

Slide 3

A TTS system is an engine that converts text to speech. There have been big advances in speech synthesis, in the artificial creation of human speech.

Slide 4

TTS AD has been researched by a team led by Agnieszka Szarkowska in Poland. They tested it in a monolingual feature film in Polish, in a dubbed educational TV series for children, in a foreign fiction film with voice-over, in a non-fictional film with AST, and in a dubbed film. Results indicate that most participants accept TTS although it is not the preferred solution.

Slide 5

Anna Fernández-Torné and I found similar results when testing artificial voices in Catalan, in a dubbed fictional content. Although human voices generally scored higher than synthetic voices, 94% of the participants considered TTS AD an acceptable solution.

Slide 6

Extensive research is also reported by Kobayashi and his team, who analyse the application of TTS AD in online videos.

Slide 7

115 adults with sight loss took part in a survey in Japan. A human voice, a standard TTS voice, and a prototype TTS voice were tested. The study continued in the USA with 236 participants. Results show that TTS ADs are generally accepted, especially for relatively short videos and informational content.

Funded by the Erasmus+ Programme of the European Union

Slide 8

In fact, Walczak and Fryer have shown that the genre has an impact on the reception of TTS. In their tests, human voices prompted significantly higher levels of presence for drama compared to TTS voices. They were assessed as more interesting and less confusing. However, in documentaries, results were similar, with no statistically significant differences.

Slide 9

It remains to be seen whether TTS AD will be transferred extensively into the market. But incorporating a TTS system in an app or a production system is a

good way to guarantee accessibility anytime anywhere for everybody. Check the proposal by Oncins and other researchers from UAB or the Scribit platform.

Slide 10

TTS systems have been successfully implemented in AST in countries such as Sweden, Denmark or the Netherlands. There is a paper by Verboom on the topic.

Slide 11

The ISO standard on AST also acknowledges this practice, and states that the narrator can be a person or a technology. However, the standard advices to select synthetic voices carefully because, and I quote, “quality, naturalness and reading speed vary depending on the synthetic voices”.

Funded by the Erasmus+ Programme of the European Union

Creation of these training materials was supported by
ADLAB PRO (Audio Description: A Laboratory
for the Development of a New Professional Profile),
financed by the European Union under the Erasmus+ Programme,
Key Action 2 – Strategic Partnerships,
Project number: 2016-1-IT02-KA203-024311.

Funded by the
Erasmus+ Programme
of the European Union