Speech Technologies and Their Applications in Mobile Learning:

Slide 2

Good day everybody. My name is Aga Palalas, welcome to the webinar on speech technologies and their applications in mobile learning. At the end of this short webinar, you will be able to identify how selected advanced speech technologies can be utilized to enhance teaching and learning using mobile devices.

I will focus on the application of text-to-speech and automatic speech recognition tools.

Enjoy the webinar.

Slide 3

As the advancements in mobile technologies have been closing the gap between the computer and the phone, the usage of text and voice on mobile devices has been converging as well.

Voice-activated technologies offer both convenience and efficiency. Most people can speak faster than they can type and they can read faster than they can speak. For mobile learners, teachers, and instructional designers that might mean more efficiency in creation and consumption of learning content, feedback and ideas. Additionally, voice-to-text technology could allow learners and teachers to record feedback, record directions, or field notes and deliver them to learners faster and in the format of their choice. It could also allow them to interact more efficiently by commenting and sharing their insights hands-free and in real time.

With advanced text-to-speech and speech recognition technologies, smartphone users can now choose between sending a text message or a voice message depending on their preferences and circumstances.

Although many mobile phone users favour text messaging to voice calls, at least among the young-adult demographic, voice commands, messages and notes to yourself might soon become the medium of choice. As the speech recognition and voice-to-text tools have been improving the size of the keyboards of our mobile devices have been shrinking a bit. In addition, restrictions on texting while driving are introduced in many countries, hence many mobile phone users turn to using their voice to issue commands and/ or send messages.

Let's look at some mobile learning and performance support activities that can be optimized by speech technologies.

Slide 4

While there are still many limitations to the accuracy and usability of the speech technologies in our phones, those users who speak standard dialects of their language should be able to rely on their mobile audio search engine or voice

commands to access information, start a conversation, or simply to operate their mobile device. Command and control options allows for speech control of common functions, like answering or rejecting calls, but not creating or editing any content.

Users can dictate their ideas, questions, observations, and reflections as notes to themselves or as a language practice; they can replace traditional note-taking methods with voice-based note-taking.

Transcription of voicemails and voice typing options, which are implemented at a fundamental level on some mobile devices, offer additional tools for recording and exchanging information and ideas.

But most importantly, speech technologies can be utilized by handicapped smartphone users and those undergoing speech therapy. People with some physical disabilities or temporary injuries to the hands and forearms may use these tools as an input method in some voice-based browsers and to issue voice commands as well. At the same time, speech synthesis tools can help the visually-impaired with reading and with telephone inquiries.

As you can see, speech recognition tools, originally developed to help the hearing or speech impaired, are used on mobile phones for a number of purposes. Thanks to the advancements in the fields of linguistics and artificial intelligence, as well as voice technologies, we have applications like *Siri* on the iPhone. If you had a chance to "talk" to *Siri* you know that in most cases she can understand what you say and what you mean and then she even answers you back.

Mind you, *Siri* technology is derived from artificial intelligence rather than speech recognition. *Siri* technology goes beyond the voice recognition capabilities, more into recognition of the semantics of natural language. Also, *Siri* does not converse equally effectively in all languages and on all topics.

Although voice-to-text applications still have many limitations, including processing speed and accuracy, the technology has improved markedly in recent years.

Slide 5

Likewise, text-to-speech technologies have gone a long way and the resulting tools can optimize the mobile learning experience.

A learner could, for example, listen to a passage that was dictated by an expert or a peer or to a post that was shared via text in a blog or *Twitter*. He or she can then decide to respond to the story either by posting a text-based comment or by dictating a voice comment which can be subsequently transcribed by the speech recognition tool. The combination of the two speech technologies truly enriches the mobile communication channel which is vital to the success of mobile learning.

Leaners could also listen to alternative audio versions of any text whenever they prefer to listen than read, for instance, while driving or waiting at a bus stop.

They can employ text to-speech options for voice commands in similar situations, and while using audio dictionaries to hear the pronunciation of words, phrases, and sentences, they can improve their language skills as well.

Moreover, utilize those features for MALL and literacy training. As for instance in the international project *Mobile E-Learning for Africa* (MELFA), which integrated a text-to-speech (TTS) engine to support reading and literacy skills training, focusing mainly on the practice of South African English as well as an indigenous African language.

Slide 6

In terms of mobile-assisted language learning, speech technologies offer much more than just convenience. They are popular tools helping both in the acquisition of the native and of foreign languages.

Speech recognition (ASR) technologies are applied to analyze learner speech and provide feedback. They are also utilized to create auditory interactions between the learner and the application, as well as between learner and their peers or experts, for instance in audio narratives or stories created with the Audioboo app.

ASR is also the basis for commercial dictation systems such as Dragon NaturallySpeaking. Such systems tend to work better with native speakers, but we can increase accuracy further by training the system to recognize particular voices.

Nowadays, many of the speech tools can detect with rather high accuracy pronunciation mistakes, distortions in the speech, and difficulties in speaking and reading, thus, helping students with their pronunciation practice or their reading practice, for example. Speech Analysis applications and Language Learning Software feature spectrogram visualizations which display human speech in a graphic representation. The visual display shows a representation of a learner's utterance alongside that of a native speaker. Students listen closely to model speech, then generate their utterance themselves following that model. They receive feedback, often both visual and auditory.

Speech Analysis and Language Learning Software is available mainly for computers but increasingly commercial language learning programs, such as Tell Me More or Rosetta Stone, offer mobile versions of their tools.

Speech technologies can also support vocabulary practice as illustrated by the Kumar et al. study which explored the use of two speech recognition-enabled mobile games, which were used to help rural children in India to read and understand new words.

Computerized speech can be also used to present learning activities and provide feedback to the learner. All of that having been said, more research is needed to understand how such oral reinforcement affects the performance of the learner. More research into acoustic analysis techniques, into Automatic Speech Recognition (ASR) systems and pronunciation evaluation algorithms is definitely needed before these tools could provide a correct and accurate feedback to all users for the improvement of their oral accuracy and proficiency.

Slide 7

Other advantages of using speech technologies in mobile learning would be the inclusion of:

- Voice-based tools as "assistive technology" for the physically challenged

- The individualized attention achieved through the voice-based feedback, i.e., targeting the various preferences and needs of our students, thus, making mobile learning more learner-centered with more attention to the diverse needs and preferences

- Learners may proceed at their own pace and focus on specific areas (e.g. practicing the oral skills in the language practice)

- Opportunities for repetition and rehearsal that students can optimize by using speech technologies and the tools, [examples of which] I am going to share on the next slide.

Slide 8

Here are some examples of mobile apps incorporating voice technologies:

- Starting with **ChaCha Answers** -a question-and-answer service available for the Android Platform

- **Google Mobile Apps –** we all tend to know these

- **Vlingo** – it's like *Siri* on the iOS platform, it's your personal assistant using voice

- **DriveSafe.ly Pro -**This app reads your incoming text messages to you and lets you compose and send responses by voice

- **Dragon Downloadable Apps – which are** voice-recognition software apps offering a variety of voice-based tools

- **Jibbigo Voice Translation tool**

Slide 9

It is time to mention a few limitations of speech technologies in the mobile learning environment.

Those technologies have not been successful at interpreting spontaneous, natural foreign accented speech. There are quite a number of technology factors that can potentially influence the accuracy of speech recognition, such as the microphone being used or the speech recognition engine that you are using.

Other important considerations would be the context in which the speech-based interaction takes place. Is it a classroom environment, for example, where you're recording audio notes? That might not be convenient or might not meet with social acceptance. Maybe it would be more appropriate to complete this type of activity in a more private space. Also consider the data consumption when it comes to using voice tools in your mobile learning.

These are only a few of the limitations that we could think of. All in all, there are restrictions to the "hands free, eyes free" learning on the go. The technological performance of speech tools will continue improving though, leading to enhanced voice-based information documenting, retrieval, and sharing.

Slide 10

And now that we've reached the end of this webinar, I hope you've learned something new and, I hope you have enjoyed this video.

Have a good day.