

## Speech Technologies Applied to eHealth and eLearning

Isabel Trancoso, Alberto Abad, Thomas Pellegrini  
INESC-ID. Lisbon, Portugal

Spoken language technologies have reached enough maturity to be integrated in many applications in eHealth and eLearning. The challenges and the potential are enormous. There are many other areas in which this claim could be equally made, but these two areas share many technical issues and, of course, they also share a huge significance from a social point of view.

This was the driving force for our recent efforts at the Spoken Language Systems Lab of INESC-ID in terms of eHealth and eLearning. This talk tries to give an overview of these efforts and, in spite of the fact that they will be demonstrated for the Portuguese language, it will also try to emphasize how easily they can be extended to new languages.

Our most recent eHealth project focus on aphasia patients. The Virtual Therapist platform (Vithea) has two key features: personalization and modularity. The first one takes into account the importance of being able to create new exercises for the patients, adapted to their hobbies or their favorite memories, thus adding to their motivation for completing the exercises. The second one will hopefully facilitate the adaptation of the platform to other diseases such as Alzheimer's or Parkinson's. The main speech module is currently based on keyword spotting, but by integrating other speech analysis modules, many different therapy / diagnosis tools can be developed, targeting pathologies such as dysarthria, sigmatism, cleft lip and palate, removed larynx, cancer of the oral cavity, etc.

Although the talk focuses mostly on therapy tools, speech and language technologies are also of paramount importance to the areas of active ageing and independent living. In this context, our most recent work is in the context of the DIRHA European project (Distant speech interaction for robust home applications) which aims at integrating speech technologies in an automated home equipped with digital microphone arrays, and evaluated by motor-impaired end-users.

In terms of eLearning, and in particular in terms of CALL (Computer Assisted Language learning), our efforts started by the development of a Portuguese version of a tutoring system from Carnegie Mellon University, focused in vocabulary learning. In the baseline version, students can learn from real texts selected from an open corpus such as the Web, on topics for which they previously marked their preference. Although the REAP (Reading Practice) platform provided the framework for several interesting thesis, dealing with readability measures, generation of distractors for cloze questions, etc., the work rapidly extended beyond the original goal of vocabulary learning.

One of the two main directions was the area of serious games. Our continuously growing set of games targets totally different goals such as learning grammar, practicing vocabulary, or improving its perception, just to name a few. Practically every NLP or speech technology module available at our lab has found an application in these games, from statistical machine translation to speech synthesis and recognition, integrating 3D technologies as well to make the games more appealing. Sometimes, the games may even use side information provided

by these modules – for instance, the confusion matrix of a speech recognition module may be used to generate distractors in a listening comprehension game. The other direction was a multimedia version of REAP. The students may now learn vocabulary from other documents beyond text, such as automatically aligned audiobooks or automatically recognized TV documentaries. In fact, our Daily REAP version is updated every day to allow students to learn from the written or broadcast news of the last 7-days, on the topics they choose. This version does in fact use all the different technologies integrated in our long broadcast news processing chain, starting with audio segmentation and speech recognition (marking the words recognized with lower confidence), and including as well capitalization, punctuation, story segmentation, and topic indexation.

ELearning, however, is far from being restricted to Language Learning. Tutoring systems are now being developed for many curricular activities, and the existence of on-line video courses such as the Khan Academy opens fascinating possibilities of using, for instance, question answering in these courses. Soft skills (such as presentation skills) can be also the target of eLearning tools, and here as well speech and language technologies may play a dominant role.

In conclusion, I hope that the examples that will be shown in this talk will contribute to illustrate my take home message: we have hardly begun to explore the full potential of speech and language technologies in these two fields.