Speech-Enabled Tools for Augmented Interaction in E-Learning Applications

Sid-Ahmed Selouani, Université de Moncton, Campus de Shippagan, Canada
Tang-Hô Lê, Université de Moncton, Canada
Yacine Benahmed, Université de Moncton, Campus de Shippagan, Canada
Douglas O’Shaughnessy, Institut National de Recherche Scientifique-Énergie-Matériaux-Télécommunications, Canada

ABSTRACT

This article presents systems that use speech technology, to emulate the one-on-one interaction a student can get from a virtual instructor. A web-based learning tool, the Learn IN Context (LINC+) system, designed and used in a real mixed-mode learning context for a computer (C++ language) programming course taught at the Université de Moncton (Canada) is described here. It integrates an Internet Voice Searching and Navigating (IVSN) system that helps learners to search and navigate both the web and their desktop environment through voice commands and dictation. LINC+ also incorporates an Automatic User Profile Building and Training (AUPB&T) module that allows users to increase speech recognition performance without having to go through the long and fastidious manual training process. The findings show that the majority of learners seem to be satisfied with this new media, and confirm that it does not negatively affect their cognitive load.

Keywords: computer programming course; mixed-mode learning; speech technology, user profile; virtual laboratory

INTRODUCTION

Recently, in the context of rapidly growing network applications, an abiding vision consists of providing computer-based media support where no sophisticated training is required. Among these applications, e-learning systems are rapidly gaining in popularity. In fact, easier-to-use development tools, lower costs, availability of broadband channels, and potentially higher returns in the form of better learner productivity have made e-
learning technology attractive to a wider variety of institutional and individual users. Numerous studies, including those of Najjar (1996) and Alty (2002), confirm that the type of computer-based media incorporated in e-learning materials can have a significant impact on the amount of information retained, understood, and recalled by learners.

Several Web-based techniques are used to develop online Collaborative Learning (CL) systems. These systems may integrate a form of chat window or forum through a public or private communication channel. To some extent, these features switch the system to an interactive and communication system, which may be separated from the underlying learning context. This justifies the development of the proposed “in context” system described below. Moreover, to design an effective e-learning tool for CL, we must avoid some common issues rising from applying or developing them. For example: (1) teachers fear to apply them in the classroom because of the apparent loss of control in the classroom; (2) students resist collaborating together because of the lack of familiarity with CL techniques and class management; (3) students are accustomed to working competitively, not cooperatively (Bosworth, 1994). These issues motivated us to develop a Web-based learning tool called Learn IN Context (LINC).

The ideal user environment has not yet been found, but individual interface technologies are sufficiently advanced to allow the design of systems capable of making a positive impact on the e-learning experience. Central to such systems is a conversational interaction using speech recognition and text-to-speech synthesis. Deng and Huang (2004) state that in recent years, Automatic Speech Recognition (ASR) and Text-To-Speech (TTS) have become sufficiently mature technologies which allow their inclusion as effective modalities in both telephony and other multimodal interfaces and platforms. In this article, we propose to include such technologies into a virtual laboratory dedicated to the mixed-mode learning of C++ language programming. This mode combines face-to-face and distance approaches to education. In this context, an instructor meets with students in the classroom, and a resource-base of content material is made available to students through the Web.

This article is further organized as follows. Section 2 is concerned with both TTS and ASR background. An overview on the relevance of including multimedia files into an e-learning environment is given in Section 3. Section 4 describes the main components and features of the LINC platform and its speech-enabled enhanced versions: LINC+ and LINC Desktop. This section also describes the spoken query system for navigating and searching, and the user profile and automatic training system. Section 5 presents the results of experiments carried out to evaluate how learners deal with the LINC (voiceless) platform, and reports objective and subjective evaluations of the ASR and TTS modules of LINC+. Finally, in Section 6, we conclude and discuss future perspectives of this work.

**SPEECH TECHNOLOGY BACKGROUND**

The general architecture of a TTS system has three components: text/linguistic analysis, prosodic generation and synthetic speech generation. Text analysis aims to analyze the input text. It involves orthographic/phonetic transcription which
Open Book Examinations in a Distance (Teacher) Education Programme: South African Teacher-learners’ Experiences
Vuyisile Msila (2011). *International Journal of Distance Education Technologies* (pp. 1-12).
www.igi-global.com/article/open-book-examinations-distance-teacher/55795?camid=4v1a