Chapter 12
The Use of Synthetic Speech in Language Learning Tools: Review and a Case Study

Oscar Saz
University of Zaragoza, Spain

Eduardo Lleida
University of Zaragoza, Spain

Victoria Rodríguez
Vienna International School, Austria

W.-Ricardo Rodríguez
University of Zaragoza, Spain

Carlos Vaquero
University of Zaragoza, Spain

ABSTRACT

This chapter aims to bring up a discussion on the use of Computer Synthesized Speech (CSS) in the development of Computer-Aided Speech and Language Therapy (CASLT) tools for the improvement of the communication skills in handicapped individuals. CSS is strongly required in these tools for two reasons: Providing alternative communication to users with different impairments and reinforcing the correct pronunciation of words and sentences. Different possibilities have arisen for this goal, including pre-recorded audio, embedded Text-to-Speech (TTS) devices or talking faces. These possibilities are reviewed and the implications of their use with handicapped individuals are commented, showing the experience of the authors in the development of tools for Spanish speech therapy. Finally, a preliminary study in the use of computer-based tools for the teaching of Spanish to young children showed how the synthetic speech feature in the language learning tool was sufficient to maintain the possibilities of the tool as a valuable language teaching element in the absence of other visual elements.

DOI: 10.4018/978-1-61520-725-1.ch012
INTRODUCTION

Different development, sensorial or physical impairments like Down’s syndrome, hearing loss or cerebral palsy, among others, are also associated to mid-severe speech disorders like dysarthria or dysglossia. These disorders are characterized by affections on the central nervous system that prevent from a correct control of the articulation organs (dysarthria) or by morphological affections on those organs like left clip and palate (dysglossia). Other disorders at the speech and language level arise from functional or hearing disabilities that produce a delay in the normal process of language acquisition in the student. In other cases, traumatic situations like surgery can make the patient lose the phonation and articulation abilities and force for a language re-training.

The main effect of these disorders is the degradation of the acoustic and lexical properties of the patient’s speech compared to the normal healthy speech, creating a wide barrier to the communication of these individuals with the surrounding environment. These speakers produce a speech whose intelligibility is much lower than that of unimpaired speakers, in some severe cases of dysarthria leading to a totally unintelligible speech; or, in other cases, they change or delete phonemes in words during their speech production, leading to semantic and syntactic misunderstandings and inaccuracies.

Speech therapy allows, in many occasions, the reduction of the pernicious effects of these disorders and provides these patients with a more effective communication, favoring the social inclusion of these individuals. Unfortunately, it is usually the case that there are not sufficient resources to provide this therapy in the way in which speech therapists would like to. Speech therapy activities are usually very time-demanding for the therapists as they have been traditionally based in the direct interaction between patient and educator, limiting the possibilities of carrying out an extensive program with several patients in the same time period or for the patients to continue and extend the therapy at their homes.

The interest in fulfilling these necessities has produced, in the latest years, a great deal of research in speech technologies for the development of computer-based tools that can provide an effective support for the semi-automation of speech therapy oriented to the speech handicapped community. These Computer-Aided Speech and Language Therapy (CASLT) tools are part of the whole effort put in the development of Computer-Aided Language Learning (CALL) tools, where CASLT tools are included as well as tools oriented to other target users, like Second Language (L2) learning tools for non-native foreign students.

The bigger effort about these tools has been focused in studying and understanding how novel acoustic analysis techniques, Automatic Speech Recognition (ASR) systems and pronunciation evaluation algorithms could provide a correct and accurate feedback to the users for the improvement of their oral proficiency. The increase in the capabilities of these tools has been significant during this time, and now most of these tools can detect with high accuracy pronunciation mistakes of the speaker, difficulties in reading, distortion in the speech and problems in the acquisition of the native or a foreign language.

However, there is little information on the use of Computer Synthesized Speech (CSS) in these tools, as most of the authors take for granted that any kind of CSS can be the optimal solution for the presentation of the audio prompt to the user. While most of the CALL tools take advantage of the possibilities of computerized speech to present the activities or to provide feedback to the user, it is not well known how the presence of this oral reinforcement affects the performance of the students to improve their communication or how students perceive this oral output, especially in the cases of severely handicapped individuals, whose perception of CSS can be extremely different compared to the unimpaired users.
Related Content

Visual Speech and Gesture Coding Using the MPEG-4 Face and Body Animation Standard
[www.igi-global.com/chapter/visual-speech-gesture-coding-using/31066?camid=4v1a](www.igi-global.com/chapter/visual-speech-gesture-coding-using/31066?camid=4v1a)

Two Distinct Sequence Learning Mechanisms for Syntax Acquisition and Word Learning
[www.igi-global.com/chapter/two-distinct-sequence-learning-mechanisms-for-syntax-acquisition-and-word-learning/108737?camid=4v1a](www.igi-global.com/chapter/two-distinct-sequence-learning-mechanisms-for-syntax-acquisition-and-word-learning/108737?camid=4v1a)

Visual Speech Perception, Optical Phonetics, and Synthetic Speech
[www.igi-global.com/chapter/visual-speech-perception-optical-phonetics/31077?camid=4v1a](www.igi-global.com/chapter/visual-speech-perception-optical-phonetics/31077?camid=4v1a)

The Writing-Pal: Natural Language Algorithms to Support Intelligent Tutoring on Writing Strategies
[www.igi-global.com/chapter/writing-pal-natural-language-algorithms/61055?camid=4v1a](www.igi-global.com/chapter/writing-pal-natural-language-algorithms/61055?camid=4v1a)