Thinking Head MulSeMedia: A Storytelling Environment for Embodied Language Learning

ABSTRACT

With technological improvements and widespread computer usage, the need for user-friendly human-machine interfaces is steadily increasing. Computer graphics and audio provide a colourful world of visual and auditory experiences; nevertheless, tutoring systems have traditionally used keyboard and mouse almost exclusively, limiting significantly the input mechanisms available to content designers. Multisensory learning is beneficial for learning a new language (Birsh, 1999), and through increasing the affordances of interaction, learning scenarios can be significantly augmented.

DOI: 10.4018/978-1-60960-821-7.ch009
The Thinking Head is an embodied conversational agent that can be used in intelligent tutoring systems as the representation of a teacher and tutor for one-on-one computer learning and teaching. Although there are many applications that benefit from an embodied conversational agent and there has been a push in our research community towards improving embodied conversational agents, our main area of research lies in the creation of applications for second language learning (as introduced in Powers, 2007; Powers et al., 2008). We have explored speech synthesis and speech recognition to permit users to interact with computer systems with their voices. To facilitate language learning through conversation, we have incorporated these auditory capabilities into the Thinking Head; additionally, we seek to create a bridge between the virtual and physical worlds through physical manipulation, achieved primarily through visual input recognition with affordable devices such as webcams.

The hybrid world provides a framework for the creation of lessons that teach and test knowledge in the second language. In one of our example grammar and vocabulary lessons, the Thinking Head instructs the student to move a tiger to a lake—a student moving the toy tiger in the real world effectively moves a virtual tiger in the virtual arena. This type of interaction is beneficial for computer-based language learning especially because we are able to see a student has successfully understood the directions if the tiger is moved to the vicinity of the virtual lake. Physical movement helps the learner to internalise the novel relationships in the second language. We also provide for additional forms of communication, including dialogue with an embodied conversational agent and writing stories using markers on a whiteboard. In summary, our system provides a natural interface with which to communicate with the hybrid-reality learning environment.

INTRODUCTION

In this chapter, we explore interactions with the Thinking Head computer interface arising from a fusion of inputs of voice, objects, and writing—without keyboard and mouse. The Thinking Head system for language learning teaches through stories; however, the Thinking Head is not just the storyteller, but rather the medium in which the storyteller and story exist, and through which a learner learns to tell the stories. A learner engages in the multisensory storytelling activity by interacting with physical props within a room to learn the stories.

A user manipulates real-world objects within the Thinking Head virtual arena. In our implementation, objects are most commonly small graspable things; however, our use of the term object applies more generally to refer to things of all sizes, persons, places, and ideas. Indeed, humans have given names to objects for thousands of years, and the meanings of objects take form as the words we use as tools in communication. In SIMULA 67, arguably the first object-oriented computer programming language, objects were representations in a discrete event system (Dahl, 1968). As Lorenz (1993, p. 12) characterises objects in computer programming, “An object is anything that models ‘things’ in the real world. These ‘things’ may be physical entities such as cars, or events such as a concert, or abstractions such as a general-purpose account.” Thus, for our purposes, an object is any natural thing or combination of things that can be quantified, described, or elaborated by man or machine.

Traditionally, computer multimedia applications consist of audio and video, while MulSeMedia, or multiple sensorial media, enhances the user interaction experience. The use of realia (physical props such as apples, hats, and magazines) adds to the language learning experience because realia engages the senses. Object recognition determines