Chapter 68
Towards the Use of Dialog Systems to Facilitate Inclusive Education

David Griol Barres
Carlos III University of Madrid, Spain

José M. Molina López
Carlos III University of Madrid, Spain

Zoraida Callejas Carrión
University of Granada, Spain

Araceli Sanchis de Miguel
Carlos III University of Madrid, Spain

ABSTRACT

Continuous advances in the development of information technologies have currently led to the possibility of accessing learning contents from anywhere, at anytime, and almost instantaneously. However, accessibility is not always the main objective in the design of educative applications, specifically to facilitate their adoption by disabled people. Different technologies have recently emerged to foster the accessibility of computers and new mobile devices, favoring a more natural communication between the student and the developed educative systems. This chapter describes innovative uses of multimodal dialog systems in education, with special emphasis in the advantages that they provide for creating inclusive applications and learning activities.

INTRODUCTION

Technological advances currently reached by computers and mobile devices allow their use to access information and a number of services. In addition, users want to access these services anywhere and anytime in a natural, intuitive and efficient way. Speech-based interfaces have become one of the main options to facilitate this kind of communication as it is a good solution to the shrinking size of mobile devices, eases the communication in environments where this access is not possible using traditional input interfaces (e.g., keyboard and mouse), and facilitates information access for people with visual or motor disabilities.

With the advances of speech, image and video technology, human-computer interaction (HCI) has reached a new phase, in which multimodal information is a key point to enhance the communication between humans and machines. Unlike

DOI: 10.4018/978-1-4666-4422-9.ch068

traditional keyboard- and mouse-based interfaces, multimodal interfaces enable greater flexibility in the input and output, as they permit users to employ different input modalities as well as to obtain responses through different means, for example, speech, gestures, and facial expressions. This is especially important for users with special needs, for whom the traditional interfaces might not be suitable (McTear, 2004; López-Cózar and Araki, 2005; Wahlster, 2006).

In addition, the widespread use of mobile technology implementing wireless communications enables a new type of advanced applications to access information. As a result, users can effectively access huge amounts of information and services from almost everywhere and through different communication modalities.

There is a large variety of applications in which spoken dialog systems can be used. One of the most wide-spread is providing information on a specific topic, such as flight/railway and booking information, tourist and travel information, weather forecast, banking systems, or conference help (Glass et al., 1995; Zue et al., 2000; Bohus and Rudnicky, 2005; Andeani et al., 2006; Callejas and López-Cózar, 2008). In some cases, spoken interaction can be the only way to access information, as, for example when the screen is too small to display information (e.g. hand-held devices) or when the eyes of the user are busy in other tasks (e.g. driving) (Mattasoni et al., 2002; Jokinen et al., 2004; Weng et al., 2006). Spoken interaction is also useful for remote control devices and robots, especially in smart environments (Lemon et al., 2001; Montoro et al., 2006; Ábalos et al., 2006; Menezes et al., 2007; Augusto, 2009). Finally, one of the most demanding applications for fully natural and understandable dialogs are virtual agents and companions (Hubal et al., 2000; Catizone et al., 2003; Corradini et al., 2005).

With the growing maturity of speech technologies, the possibilities for integrating conversation and discourse in e-learning are receiving greater attention, including tutoring, question-answering, conversation practice for language learners, pedagogical agents and learning companions, and dialogs to promote reflection and metacognitive skills. This chapter focuses on some of the most important challenges that researchers have recently envisioned for future multimodal interfaces applied to educative purposes. It describes current efforts to develop intelligent, adaptive, proactive, portable and affective multimodal interfaces. All these concepts are not mutually exclusive, for example, the system's intelligence can be concerned with the system's adaptation enabling better portability to different environments.

DIALOG SYSTEMS: MODULAR ARCHITECTURE AND PROCESSES

The complexity of the interaction between the user and the dialog system can vary and some of the previously described components might not be used. For example, for a simple menu, semantic analysis is not necessary. However, for a conversational companion all the modules must be used in order to interpret the user input, take justified decisions on what the system will respond, and finally tailor the answer to user needs and expectations. This way, the implementation of multimodal dialog systems is a complex task in which a number of technologies are involved, including signal processing, phonetics, linguistics, natural language processing, affective computing, graphics and interface design, animation techniques, telecommunications, sociology and psychology. The complexity is usually addressed by dividing the implementation into simpler problems, each associated with a system's module that carries out specific functions. Usually, this division is based on the traditional architecture of spoken dialog systems: automatic speech recognition (ASR), spoken language understanding (SLU), dialog management (DM), natural language generation (NLG) and text-to-speech synthesis (TTS).