Griol, D., Molina, J.M., Callejas, Z. (2014). Incorporating android conversational agents in m-learning apps. *Expert Systems*, 34(4), e12156.

DOI: 10.1111/exsy.12156

Incorporating android conversational agents in m-learning apps

David Griol,[1]* José Manuel Molina[1] and Zoraida Callejas[2]

(1) Department of Computer Science, Carlos III University of Madrid, Avda. de la Universidad, 30.28911Leganés, Spain
E-mail: david.griol@uc3m.es
(2) Deparment of Languages and Computer Systems, University of Granada, CITIC-UGR, C/ Pdta. Daniel Saucedo Aranda s/n.18071Granada, Spain
E-mail: zoraida@ugr.es

Abstract: *Smart mobile devices have fostered new learning scenarios that demand sophisticated interfaces. Multimodal conversational agents have became a strong alternative to develop human-machine interfaces that provide a more engaging and human-like relationship between students and the system. The main developers of operating systems for such devices have provided application programming interfaces for developers to implement their own applications, including different solutions for developing graphical interfaces, sensor control and voice interaction. Despite the usefulness of such resources, there are no strategies defined for coupling the multimodal interface with the possibilities that these devices offer to enhance mobile educative apps with intelligent communicative capabilities and adaptation to the user needs. In this paper, we present a practical m-learning application that integrates features of Android application programming interfaces on a modular architecture that emphasizes interaction management and context-awareness to foster user-adaptively, robustness and maintainability.*

Keywords: computer supported education, conversational agents, mobile learning, m-learning, adaptive learning, multimodal interaction, spoken interaction, mobile devices, android, intelligent assistants, university digital libraries

1. Introduction

As a generalized phenomenon, technology integration is an increasingly crucial element of teaching and learning (Blevins *et al*., 2014; Boe *et al*., 2015; Duta and Martínez-Rivera, 2015; Lakkala and Ilomaki, 2015). Devices such as smartphones, tablets and wearable's are becoming widespread, and the fact that increasingly more individuals have always with them a device with numerous displays, sensors and connectivity possibilities, opens new learning scenarios that demand more sophisticated interfaces with multimodal capabilities (O'Halloran, 2015; Salse *et al*., 2015; Callahan *et al*., 2014; Pérez-Marín and Pascual-Nieto, 2011).

Multimodal conversational agents (Pieraccini, 2012; McTear and Callejas, 2013; Griol *et al*., 2014) can be defined as computer programmes designed to emulate communication capabilities of a human being including natural language and several communication modalities (such as speech and visual information). These interfaces employ a variety of techniques to engage students in learning. Using natural language in educational software also allows students to spend their cognitive resources on the learning task, rather than in how to use the interface of the application (Beun *et al*., 2003). Also, as a complement to teacher–student communication in online scenarios and out of school studying hours, dialogue and anthropomorphic characteristics of pedagogical multimodal systems may help supporting the social dimension of the learning activities, and the social context has been argued to help the cultivation of, and motivation for, knowledge (Baylor and Kim, 2005; Fryer and Carpenter, 2006). One of the advantages of mobile apps at the educative level is the personalization and students involvement in their learning by means of the development of new activities and even new pedagogical models around mobile devices. However, because of this variability and the huge amount of factors that must be taken into account, these systems are difficult to develop and typically are developed ad-hoc, which usually implies a lack of scalability (Rouillard, 2007; Paek and Pieraccini, 2008). Multimodal conversational systems developed to provide these functionalities typically rely on a variety of components, such as speech recognition and synthesis engines, natural language processing components, dialogue management, databases management and graphical user interfaces. Laboratory systems usually include specific modules of the research teams that build them, which make portability difficult. Thus, it is a challenge to package up these components so that they can be easily installed by novice users with limited engineering resources.

To tackle the challenges discussed, it is necessary to perform careful user modeling and effective dialogue management. Statistical approaches are suited for this purpose, as the application of such techniques to user modeling and dialogue management makes it possible to consider a wider space of dialogue strategies in comparison to engineered rules (Young, 2002; Griol *et al*., 2014). The main reason is that statistical approaches can be trained from real dialogues, modeling the variability in user behaviours. Although the parameterization of the model depends on expert knowledge of the task, the resulting conversational agents have a more robust behaviour, better portability and are easier to adapt to different user profiles or tasks (Seneff *et al*., 2007; Kartakis, 2010). In addition, the combination of conversational interfaces and mobile devices has lead to a new paradigm in which they can collect information from the user pervasively. This can help to build more complex user and dialogue models to be employed not only to provide the system functionality, but also to boost its performance.

In our previous work (Griol and Molina, 2014), we proposed a general-purpose framework to develop conversational interfaces for mobile devices. It provides a modular and scalable architecture to develop such systems efficiently for Android-based mobile devices. Android, the most popular alternative among developers of mobile Apps, offers libraries to build interfaces including different resources for graphical layouts as well as speech recognition and text-to-speech synthesis. Currently more than 82% of smartphones and tablets operate with Android OS (Corporation, I. D., 2015). Also, there is an active community of developers who use the Android Open Source Project and have made possible to have more than one million applications currently available at the official Play Store, many of them completely free. For these reasons, our framework makes use of different facilities integrated in Android-based devices.

The proposed framework integrates the facilities of the Android application programming interfaces (API) in a modular architecture that emphasizes interaction management and context-awareness to build sophisticated, robust and maintainable educative applications. Our proposal for context awareness is based on the definition of a statistical methodology for user modeling that estimates the user intention during the dialogue. In our contribution, the information provided by the user model can also enriched by considering information related to the external context of the interaction.

The internal and external context, respectively related to the prediction of the user intention by the user model and the information provided by the sensors supported by mobile devices, make it possible to adapt the system dynamically taken into account these valuable information sources. To do this, a statistical dialogue model based on neural networks is generated taking into account the contextual information and the history of the dialogue up to the current moment. The next response of the conversational agent is selected by means of this model. The codification of the information and the definition of a data structure which takes into account the data supplied by the user throughout the dialogue makes the estimation of the dialogue model from the training data and practical domains manageable.

In this paper, we show how this framework can be easily integrated into mobile learning applications that provide a range of functionalities covering the provision of personalized information, access to social networks, resolution of different kinds of educative exercises, etc. As a proof of concept, we have designed the Your-University App, which demonstrates the feasibility of our approach with a variety of pedagogic objectives, restrictions and demands. In the app, the first set of functionalities is focused on the generation of practical educative exercises that promote autonomous learning and self-assessment. Our proposal to develop pedagogic conversational interfaces improves the interaction to solve the proposed exercises, personalize the selection of the different learning activities considering the student's preferences using the system and their previous uses, and allows the provision of immediate feedback by means of the automatic assessment of the learning activities.

The second set of functionalities is focused on the provision of user-adapted information and services. Among the different educational application domains, these functionalities have been selected to develop advanced digital libraries (Griol *et al*., 2015), which extend the concept of Library 2.0. To do this, we have relied on the recommendations for the development of University Digital Libraries (Vassilakaki and Garoufallou, 2013) compiled by recent studies (Chen and Albee, 2012; Moreno *et al*., 2014), which propose using new open-source technologies to build complex systems at less cost, increasing the efficiency and sharing of bibliographic records for all libraries, maximizing the use of these libraries, providing new types of multimedia contents and interaction possibilities and transferring library efforts into higher-value activities.

The remainder of the paper is as follows. Section 2 describes the motivation of our proposal and related work. Section 3 describes the framework used to develop adaptive multimodal conversational agents and how it has been adapted to mobile learning scenarios. Section 4 describes the development of the Your-University App based on the framework. Section 5 shows the results of the evaluation of this system using a variety of measures. Finally, Section 6 presents some conclusions and future research lines.

2. State of the art

Several recent studies describe the great potential of mobile devices to transform the traditional classroom to one that is more interactive and engaging (Ferdousi and Bari, 2015; Mouza and Barrett-Greenly, 2015; Gómez-Domingo and

Badia-Garganté, 2016; Zydney and Warner, 2016). Mobile devices offer a number of features especially important for the design of educative applications: technology-based scaffolding, location-aware functionalities, visual/audio representations, differentiated roles, digital knowledge-construction tools, digital knowledge-sharing mechanisms, portability, interactivity, context sensitivity, connectivity, individuality and social media.

The positive impact of mobile learning (m-learning) applications is related to new ways of learning (Furió *et al*., 2015), engaging students in the learning process (Lu *et al*., 2014), reinforcing students' learning (Churchill and Wang, 2014), increasing motivation (Adegbija and Bola, 2015), generating autonomous and personalized learning (Gerger, 2014), encouraging pupils to take control of their own learning (Boticki *et al*., 2015), helping learners to manage their self-directed learning (Lu *et al*., 2014), encouraging students to be supportive coaches for their classmates (Gerger, 2014), facilitating access to information (Yang *et al*., 2015) and promoting collaborative learning (Murphy, 2011).

Research on mobile learning has rapidly increased in different kinds of applications, such as mobile learning games (Avouris and Yiannoutsou, 2012), mobile computer-supported collaborative learning (Hsu and Ching, 2013), or mobile apps (Jeng *et al*., 2010). Taking into account the studies described (Gómez-Domingo and Badia-Garganté, 2016), three main types of educative Apps can be identified: Learning Skills Apps, Informational Management Apps and Content Learning Apps. Learning Skills Apps enable students to create their own knowledge. Informational Management Apps have the ability to work within the specific context and environment of the learning. Finally, Content Learning Apps are considered to proportionate students' different activities that allow them to rehearse, reinforce, practice and assess curricular content.

The requirement of usability and design of educational materials adapted to the teaching contexts are usually seen as key factors in order to succeed when learning via mobile devices (Klenner, 2015). Giménez *et al*. describe different methods of adapting educative digital content for its delivery via mobile devices (Giménez-López *et al*., 2009). Magal-Royo *et al*. extends this description by means of a proposal of new paradigms in education regarding the use of audiovisual contents adapted to mobile devices (Magal-Royo *et al*., 2010). According to their proposal, the main differences between e-learning and m-learning environments are related to technical validation, creation, development and analysis of students' results, as the most important components of an autonomous learning process.

Yue and Zin proposed a design model for the development of mobile application games adapted to hearing handicapped children (Yue and Zin, 2013). The proposed model, which integrates voice recognition capabilities, can be adapted using some strategies for increasing interest of learning, fostering motivation, self-efficacy and impacting handicapped students' attitude to study.

King *et al*. have recently investigated mobile app development as a means of engaging students in creating historical narratives (King *et al*., 2014). The MIT AppInventor constructivist learning tool[1] is proposed to help students with no programming experience to design apps through a drag-and-drop visual programming language. The main results of the study suggest that this kind of apps increase students' abilities to think, analyze, critique, cultivate their ideas and seek out opportunities to refine and share them with others.

Nordin *et al*. proposed a mobile learning framework for lifelong learning (Nordin *et al*., 2010). The set of general requirements that are described includes to be highly portable, unobtrusive, available, adaptable, persistent, useful, user-friendly and support individual learning. These aspects are also described in (Ferdousi and Bari, 2015) as the main outcomes that can be achieved by introducing m-learning into undergraduate courses. These aspects can be addressed by means of conversational multimodal systems that establish a more engaging and human-like relationship between the students and the m-learning apps.

Current possibilities to employ multimodal conversational agents for educative purposes include tutoring applications (Pon-Barry *et al*., 2006), question-answering (Wang *et al*., 2007), conversation practice for language learners (Fryer and Carpenter, 2006), pedagogical agents and learning companions (Cavazza *et al*., 2010), storytelling educative games and applications (Padilla-Zea *et al*., 2014), and dialogues to promote reflection and metacognitive skills (Kerly *et al*., 2008a). These agents may also be used as role-playing actors in immersive learning environments (Griol *et al*., 2012).

Conversational systems as personal coaches integrate information about the application domain. Systems of this kind are characterized by the possibility to represent and continuously update information that models the cognitive and social users' state. The main objective is to guide and monitor users in the learning process, providing suggestions and other interaction functionalities not only with the developed application but also with the rest of students. In order to achieve this goal, these applications usually integrate realistic and interactive interfaces.

For example, Grigoriadou *et al*. (2003) describe a system where the learner reads a text about a historical event before stating their position about the significance of an issue and their justification of this opinion. The answers are classified as scientific, towards-scientific or non-scientific, and a dialogue generator produces 'appropriate reflective diagnostic and learning dialog for the learner'. Similarly, in the CALM system (Conversational Agent for Learner Modeling) (Kerly *et al*., 2008b), the users answer questions on the domain and state their confidence in their ability to answer correctly. The system infers a knowledge level for the student based on their answers and encourages the learner to engage in a dialogue to reflect on their

[1]http://appinventor.mit.edu/explore/

self-assessment and any differences between their belief and that of the system about their knowledge levels.

Another example of natural language tutoring is the Geometry Explanation Tutor (Aleven *et al.*, 2004), where students explain their answers to geometry problems in their own words. Additionally, the Oscar conversational intelligent tutoring system (Latham *et al.*, 2012), aims to mimic a human tutor by implicitly modeling the learning style during tutoring, personalizing the tutorial to boost confidence and improving the effectiveness of the learning experience. The system uses natural language to provide communication about specific topics with the users and dynamically predicts and adapts to a student's learning style.

Other conversational systems provide a visual representation through an animated bot with gestures and emotional facial displays. These bots have shown to be a good interaction metaphor when acting in the role of counsellors (Gratch *et al.*, 2002; Marsella *et al.*, 2003), personal trainers (Bickmore, 2003), or healthy living advisors (de Rosis *et al.*, 2005; Callejas *et al.*, 2014); and have the potential to involve users in a human-like conversation using verbal and non-verbal (Cassell *et al.*, 2012).

The combination of gamification and conversational interfaces is one of the most innovative trends to develop educative applications (Seaborn and Fels, 2015; Urh *et al.*, 2015; Kuo and Chuang, 2016). Educational games benefit making decisions, promote problem-solving abilities and creativity, improve the ability to cooperate and respect for others, increase motivation and arouse interest in learning, strengthen knowledge and can also be used to propose roles and improve skills required by specific works (Pedreira *et al.*, 2015; Urh *et al.*, 2015).

Blemins *et al.* propose the use of these applications to enhance citizenship education and increase students' civic engagement with very positive results (Blevins *et al.*, 2014). Lee and Molebash have recently completed a study of the main design factors that must be considered to design this kind of applications (Lee and Molebash, 2014). Among them, we find original intent, student project models and the scaffolding story and delivery method.

To cover them, we have adapted our general framework for the development of mobile conversational agents to the m-learning scenario. As will be described in the following section, its capabilities to support comfort and user awareness, rich dialogue model and intent adaptation make it a very relevant choice for the development of mobile learning apps.

3. Application of our general framework for the development of mobile conversational agents to the m-learning domain

Figure 1 shows our framework for the development of adaptive multimodal conversational agents for mobile devices that was presented in (Griol and Molina, 2014). A multimodal conversational agent providing spoken interaction integrates five main tasks to deal with user's spoken utterances: automatic speech recognition, natural language understanding, dialogue management, natural language generation and text-to-speech synthesis (TTS).

Discarding the simplest case, these agents require a sequence of interactions between the user and the system to achieve their final purpose. Therefore, the user's goal is gradually reached during several dialogue turns. To do this, it is necessary to endow the system with the abilities to reference information that has appeared previously during the dialogue, take the initiative to recover the dialogue after a failure, request information that is necessary to fulfil the objective, or require clarification if it is not confident about the information provided by the user.

During the communication process, the system initially generates a message to welcome and inform the user about the features and functionalities of the system. Then, the system must perform a basic set of actions that are cyclically repeated after each user utterance: recognize the sequence of words mentioned by the user; extract the meaning of these words (i.e. understand the information that is useful for the system domain), access web services and databases to extract the information required by the user, adapt the interaction to the context features described earlier, decide what action or actions should be performed after each user request and play a spoken message to provide a response to the user.

Our proposal integrates the Google Speech API to include the speech recognition functionality in a multimodal agent. Speech recognition services are available on Android since Android 2.1. Android 4.0 introduced voice input capacity to transcribe the recognition results as the user speaks, without waiting until it has finished, and Android 4.1 Jelly Bean enabled offline voice recognition, making it possible to use speech recognition without the need for Internet access.

This makes our framework very well filled for a variety of m-learning scenarios that require user inputs of different complexity, from isolated commands (e.g. predefined options to respond to a question) to open speech (e.g. elaborated explanations).

Once the conversational agent has recognized what the user uttered, it is necessary to understand what he said. Natural language processing generally involves morphological, lexical, syntactical, semantic, discourse and pragmatical knowledge (Wu *et al.*, 2010). Lexical and morphological knowledge allow dividing the words in their constituents distinguishing lexemes and morphemes. We propose the use of grammars in order to perform the semantic interpretation of the user inputs (Wang and Acero, 2006; Kaufmann and Pfister, 2012).

In our framework, we model the context of the interaction as an additional valuable information source to be considered along with the semantic representation of the user input. In order to do so, we consider external and internal contextual information sources.

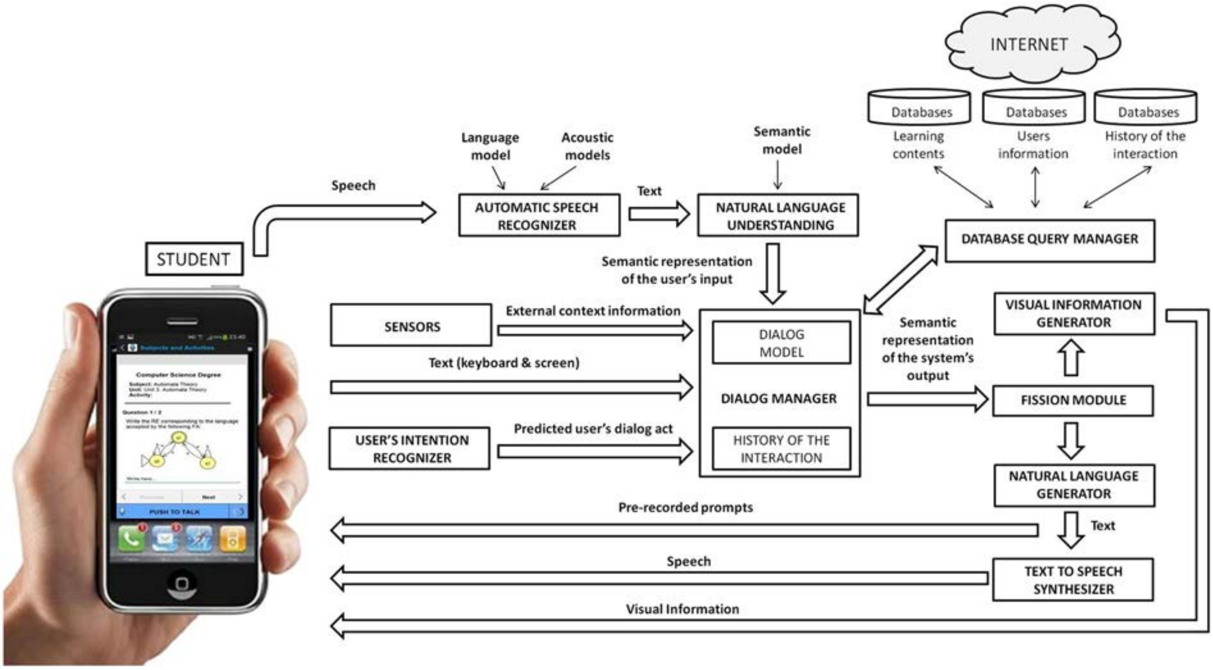

Figure 1: *The general-purpose framework for the generation of m-learning multimodal conversational agents.*

We acquire external context by using sensors currently supported by Android devices. Most Android-powered devices have built-in sensors that measure motion, orientation, position and various environmental conditions.

This rich approach for external context acquisition provides very interesting opportunities to foster the potential of mobile learning to move from isolated, silenced and immobilized learning spaces to new scenarios of conversation, freedom of movement and independence.

Regarding internal context, our framework is based on the traditional view of the dialogue act theory, in which communicative acts are defined as intentions or goals. Our technique is based on a statistical model to predict user's intention during the dialogue, which is automatically learned from a dialogue corpus. This model is used by the system to anticipate the user's needs by dynamically adopting their goals and also providing them with unsolicited comments and suggestions, as well as responding immediately to interruptions and provide clarification questions. The model takes into account the complete history of the interaction and also the information stored in user profiles.

The information in the user profile can be classified into three different groups:

- General user information. User's name and machine identifier, gender, preferred language, pathologies or speech disorders, age, current location, date and time.
- Skill level: This level is estimated by taking into account variables like the number of previous sessions, dialogues and dialogue turns, their durations, time that was necessary to access a specific web service, the date of the last interaction with the system, etc. A low, medium, high or expert level is assigned using these measures.
- Usage statistics: They store the count of each action over the system that a user performs. Users' preferences are automatically evaluated considering the user's most required services during the previous dialogues, date and hour of the previous interactions, most frequent objectives and restrictions, and preferred output modality.

As can be observed, the user profile is very focused on the user technical skills and system usage statistics. To make the internal context more meaningful in the pedagogic domain, it is also possible to include the initial level and intermediate and final results of the students, as illustrated in the app implemented (Section 4).

The dialogue manager decides the next action of the system (Traum and Larsson, 2003; Williams and Young, 2007), interpreting the incoming semantic representation of the user input in the context of the dialogue. This module deals with different sources of information such as the natural language understanding results, database queries results, application domain knowledge and knowledge about the users and the previous dialogue history to select the next system action. We use a statistical methodology that combines multimodal fusion and dialogue management functionalities. To do this, a data structure is introduced to store the information provided by the user's inputs, the user's intention model and the context of the interaction.

In the m-learning domain this dialogue management approach can be used at a 'local' level to control specific dialogues (e.g. series of system questions and students responses), and at a 'global' level to use conversation as a dialectic framework to construct knowledge within a subject.

Finally, the modality fission module receives abstract, modality independent presentation goals from the dialogue manager. The multimodal output depends on several constraints for the specific domain of the system, for example, the current scenario, the display size and user preferences like the currently applicable modality mix. This module applies presentation strategies that decompose the complex presentation goal into presentation tasks. It also decides whether an object description is to be uttered verbally or graphically. The result is a presentation script that is passed to the Visual Information and Natural Language generation modules.

The visual generation module creates the visual arrangement of the content using dynamically created and filled graphical layout elements, while a text-to-speech synthesizer (we use the built-in Android engine) is used to transmit the voice message.

The multimodal output is very important to generate a rich and engaging m-learning experience, and its flexibility is especially relevant to ensure that students with functional diversity have access to the app and learning materials.

4. Practical m-learning multimodal application: the your-university App

To verify that our framework has the real potential of being used in m-learning, we have applied it to develop the Your-University App, a practical mobile learning application covering the main functionalities of m-learning scenarios, which are related to the provision of user-adapted educative exercises, information and services. As a multimodal application, the system allows users to interact by means of the tactile mode, by voice or by combining both modes.

In addition to the technologies described in the previous section, the application uses different data repositories, several of them arranged in an external Web server in order to allow access to the application and user identification using mobile devices. A specific web service has been developed to transmit a request from the Android application to a web server, which accesses the MySQL database and returns the requested information to the Android application in JavaScript Object Notation format (JSON). A JSON element consists of two structures: a collection of pairs of name-value and an ordered list of values. The Android application extracts information from these elements using the org.json.JSONObject and org.json.JSONArray classes. The Android application also accesses an internal SQLite database to store locally specific user information.

The application also includes features that require extracting information from different web pages (e.g. to access the catalogue of books or check the regulatory requirements in a library). To perform this query, the Android application uses the JSoup Java library to connect the web page and extract the HTML contents.

The following subsections summarize the main functionalities of the application, which can be classified into the design of educative activities and the provision of user-adapted information and services related to a University Digital Library.

4.1. Design of educative activities

The first block of the application makes it possible to interact with students to propose them educative questions and practical exercises, obtain their answers, automatically process the responses provided, inform students whether the answers provided are correct or not and provide them an appropriate feedback depending on the responses provided. The design of this functionality of the application is oriented to promote autonomous learning and increase and self-assess the knowledge acquired by University students.

The users of the application must firstly register (see Figure 2, first image). There are different permissions and functionalities associated to students and teachers. In the registration process, students select their username and password and complete a questionnaire with questions related to their university studies and preferences when interacting with the application. This information is required to personalize the application and configure a wide range of options related to the main functionalities that are provided (see Figure 2, second and third images).

After providing a valid username and password, there appears a screen where students can select the educative activities corresponding to a specific subject and unit (see Figure 3, first and second images). To make this selection, students can use the visual elements on the screen, which are dynamically generated according to the information provided in the registration process and the set of units and activities associated to each one of the subjects, or by means of speech interaction using one or more steps. To do this, different grammars have been designed for the SLU module to recognize and understand sentences like 'I would like to carry out the Test 1 of Unit 1 for the subject Automata Theory', 'Show me the activities available for the Unit 2 of Artificial Intelligence' (step 1), 'Exercise 3' (step 2). Such grammars are not used during recognition because of the restrictions posed by the Android speech recognizer. Thus, recognition is performed openly (every phrase in the target language is recognized), and we employ the grammars during a post-processing phase in the language understanding unit to discard non-appropriate inputs and obtain the semantics of the valid utterances.

As a tool to foster autonomous learning, our app does not only provide access to the learning resources by demand, but also gives facilities to the students to establish weekly plans and thus help them to organize their work and keep engaged in their learning helping them to achieve regularity in their study (see Figure 3, third image). In addition, the

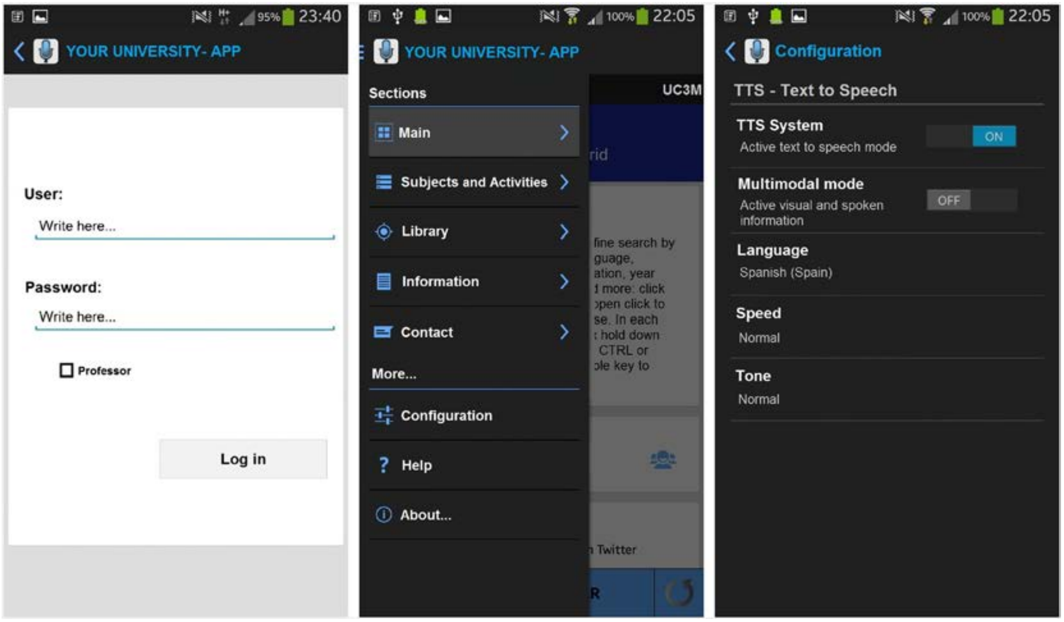

Figure 2: *Screenshots of the registering and configuration modules.*

application provides a list of recommended activities according to the previous uses of the application. The set of recommend exercises is dynamically generated to include the ones that the student has not completed yet according to the weekly planning and the exercises with which the student has experienced more difficulties during previous interactions with the application. This way, the system has an effective influence on the correction of errors and also considers the specific evolution of each student. This approach is being used in successful language learning apps, such as Duolingo,[2] in which the exercises in which the user has committed mistakes are more frequently posed.

The application allows to configure and complete three types of activities, which can be selected and alternated to design an educative exercise. The first activity is related to questions in which the only possible answers are true or false (see Figure 4, first image). Students can complete these questions by selecting the corresponding option on the screen or directly uttering ‘True’ or ‘False’. A specific feedback message can be configured for each one of the options in the question. This message, which has been especially designed for the cases in which a wrong option is selected, is provided to the students by means of a text-to-speech synthesized voice or an audio file. The next and previous spoken commands and buttons facilitate the interaction to complete an exercise including more than one activity.

The second type of activities allows defining multiple choice questions for which students must select the option (s) that are correct (see Figure 4, second image). As in the previous type of activity, students can select the option(s) using the visual elements on the screen or uttering the letters associated to each one of the options. A feedback message can also be specified to be provided to the students for each one of the options in each question.

The last type of activities corresponds to practical exercises for which the student must provide a valid solution. By means of these activities, it is possible to incorporate the wording of the exercise and also different kinds of image and video file formats, as Figure 4 - third image shows. This type of exercises incorporate a field to provide the result and allows to display the correct solution and also reproduce a text-to-speech synthesized voice or audio file with a feedback message each time a proposed result is provided.

Each time students complete an exercise containing the previous types of described educative activities, they can access a screen showing the mark obtained for the exercise (see Figure 5, first image). It is also possible to access the results for all the exercises already completed. It is possible to filter these results by selecting only a specific subject, each one of the units of a subject, or each one of the activities (see Figure 5, second image). This module of the application also allows accessing the different exercises as they were completed by the student (see Figure 5, third image). The information provided to the student includes the mark of the exercise, the option selected for each activity in the exercise, and the correct option for each one of them. It is also possible to access a list with the exercises corresponding to the selected criteria that have not completed yet.

4.2. Provision of user-adapted information and services related to a university digital library

The second block of the Android application consists of a multimodal virtual assistant that provides information and services related to a University Digital Library. The main

[2]https://en.duolingo.com/

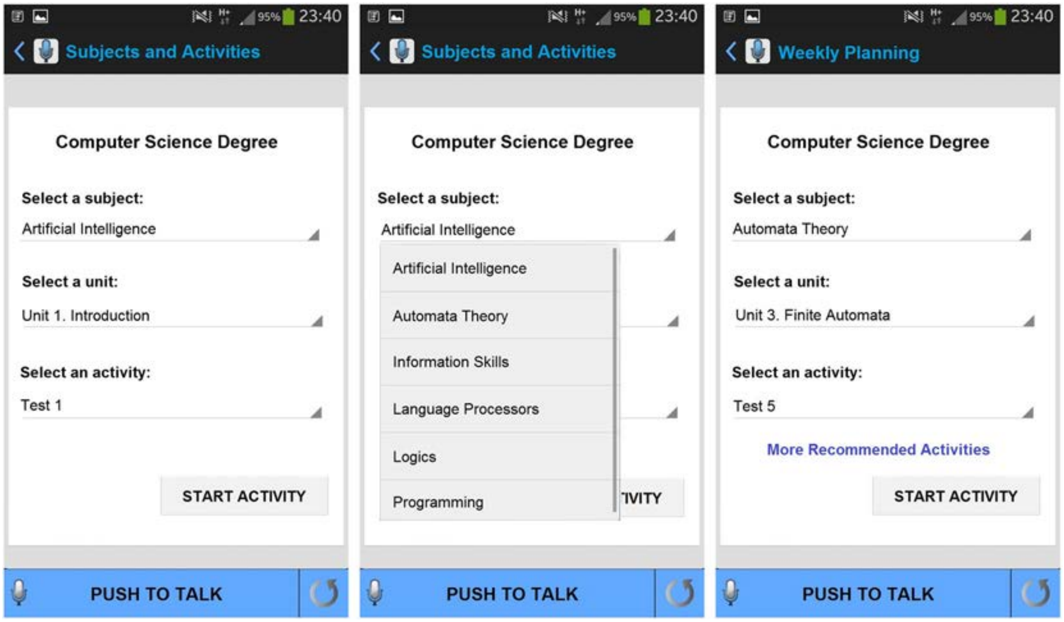

Figure 3: *Screenshots of the selection of educative activities and weekly planning.*

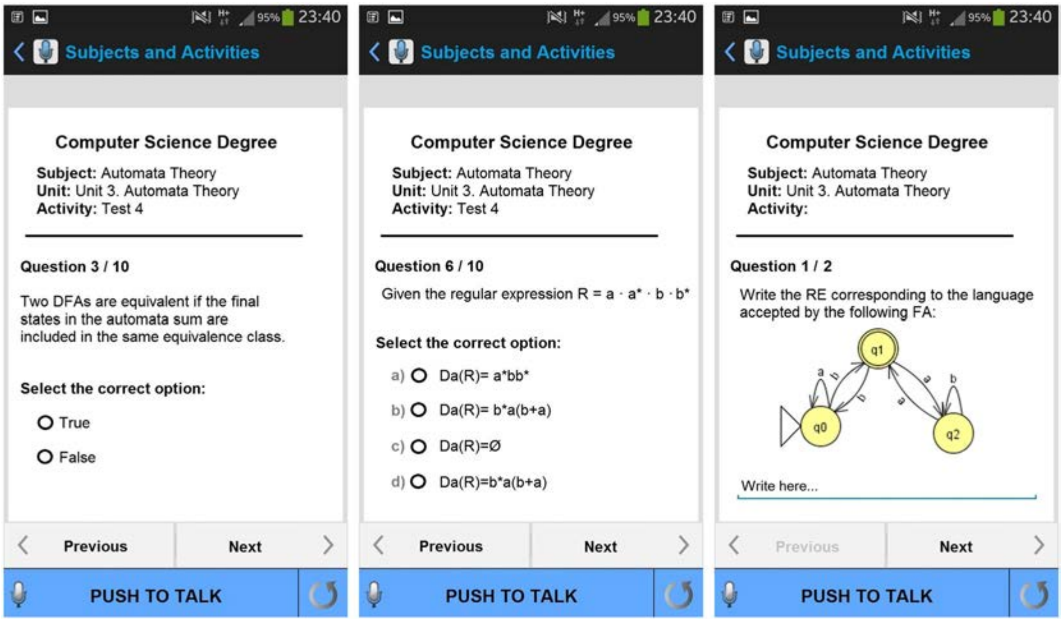

Figure 4: *Screenshots of the different types of educative activities.*

functionalities provided by the multimodal application can be classified into five main modules.

The first module offers access to the library catalogue stored in different databases. Users can provide their query orally or using the tactile elements. The application provides the completed information of the requested resource, their current state (available, borrowed or waiting list) and images with a visual description extracted from the web repository of the university library (see Figure 6).

The application can also guide the user through the UC3M Library at the Leganés Campus. To do this, the user can provide the name of a specific location, a service offered by the library, the title of a bibliographical resource, or subject in the catalogue. The application uses images; photographs and synthesized speech that convey specific step by step instructions to reach the required location starting from the main door of the library (see Figure 7, first and second images).

The third module of the application is related to the frequently asked questions. The set of questions that the application is able to answer has been extracted from the lists provided by the university libraries of the University of León, University of Alcalá de Henares, University of Granada, Carlos III University of Madrid, University of

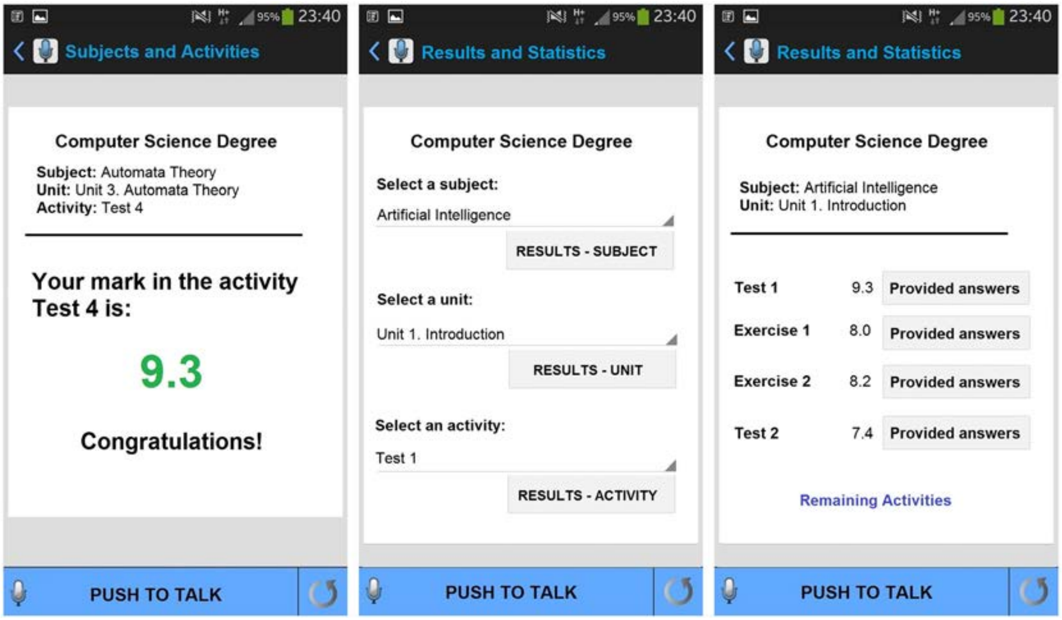

Figure 5: *Screenshots of the results and statistics of the educative activities.*

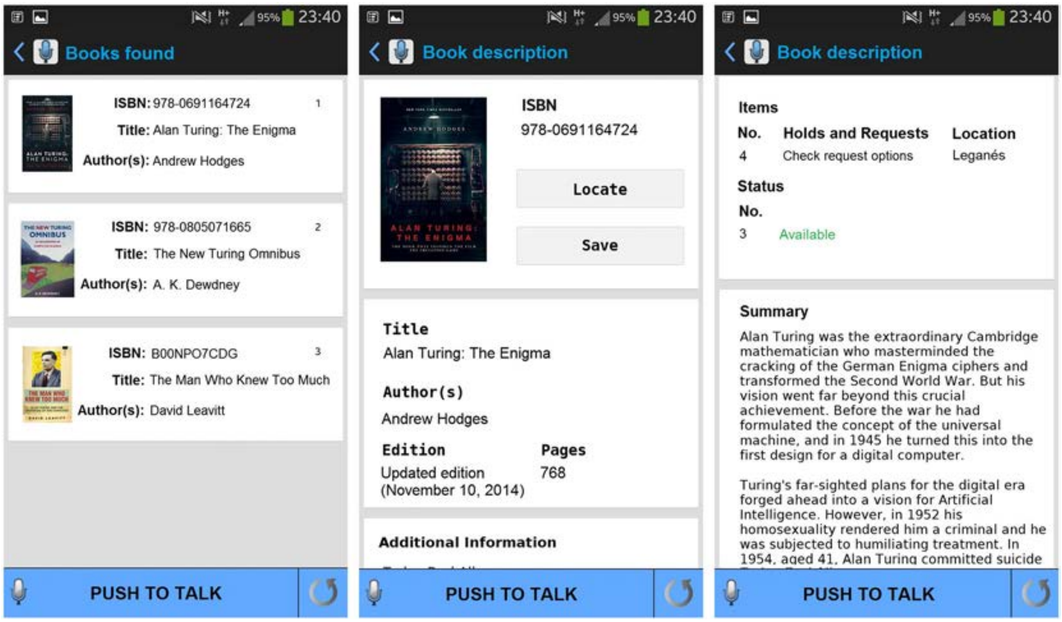

Figure 6: *Screenshots of the library catalogue module.*

Almería, University of Murcia, Autonomous University of Barcelona, University of Córdoba and University of Alicante.

This set of questions has been classified into 13 categories (General, University Identity Card, Services and Infrastructure, Books, Borrowing and Renewing, Interlibrary Loan, Finding Information and Access to Resources, Search, Electronic journals, eBooks, Institutional Repository, Training courses, Impact and quotations and Acquisitions), encompassing more than 250 frequently asked questions asked by users of these libraries. A specific response adapted to the infrastructure and regulations of the libraries at the Carlos III University of Madrid has been designed for each question (see Figure 7, third image). An example of a query and the corresponding response provided by the application is shown below:May I renew my books online, or by telephone?

Response: Yes, books and other items that you have checked out may be renewed online, with some exceptions. To renew online, use the My Library Account link and log in using your currently valid University Network ID and password. You can renew an item up to 10 times from the original check-out date unless another borrower has requested it, an instructor

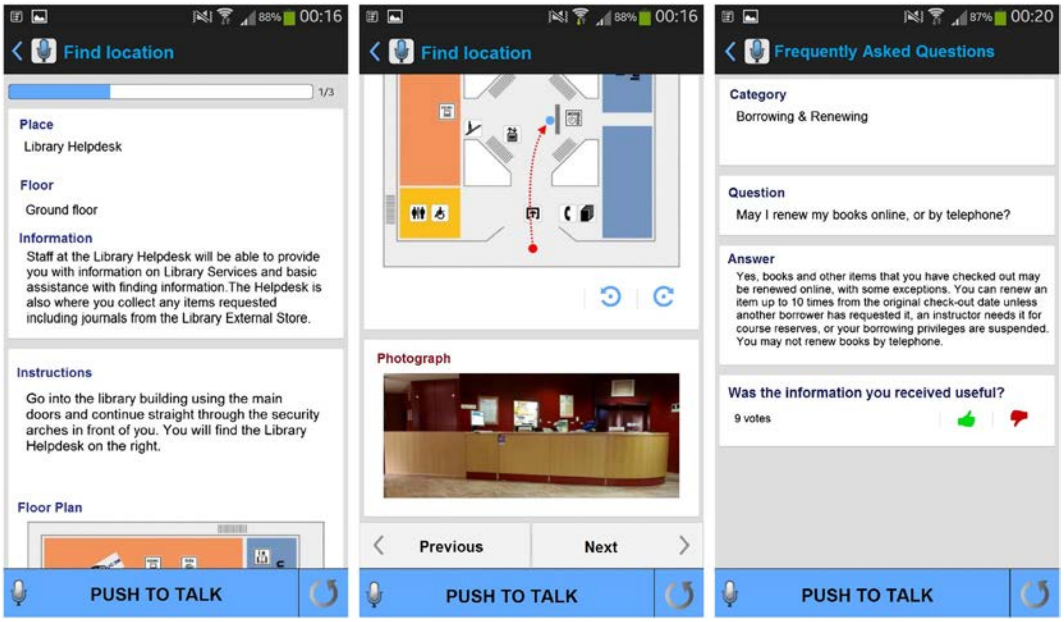

Figure 7: *Screenshots of the find location and frequently asked questions modules.*

> needs it for course reserves, or your borrowing privileges are suspended. You may not renew books by telephone.

The fourth module of the application provides information on how to contact the Library. This module provides access to a web form to request information directly to university staff, transmit suggestions, complaints, or compliments.

Finally, the Help & Quick Find module of the application provides a detailed help about the different functionalities of the application, as well as a shortcut to each of them. This information is provided by means of screenshots of the application, instructional videos, use cases, synthesized speech and practical examples explaining how to use the previously described functionalities. Also, users can access a wide range of Web 2.0 tools provided by the Carlos III University of Madrid (website of the UC3M Library, Twitter account, Facebook and institutional repository of news and weekly schedule).

5. Evaluation

Task-oriented multimodal conversational systems are usually evaluated in terms of interaction parameters and subjective judgments (Dybkjaer *et al.*, 2004; Kanto *et al.*, 2004; Moller *et al.*, 2006; Callejas and López-Cózar, 2008). Interaction parameters include the technical robustness and core functionality of the system components as well as system performance measures such as task completion rate; whereas subjective usability evaluations estimate features like naturalness and quality of the interactions, as well as user satisfaction reported in questionnaires and interviews.

We have followed this method to evaluate the previously described practical m-learning application with 42 students attending to the Formal Languages and Automata Theory course of the Computer Science Degree at the Carlos III University of Madrid over the three-month period between September and December 2015. This subject covers theoretical and practical contents about Finite Automata, Push-Down Automata, Turing Machines and the design and analysis of programming languages. Each student had at his/her disposal a mobile device and interacted with the application during the course. Before the lectures started, the tutors guided the students through the different functionalities of the m-learning application and provided them with a little training session on the available multimodal communication features.

A total of 125 questions and exercises were designed combining the teaching contents and proposed exercises provided by the Moodle server at our university (Aula Global[3]). For the generation of these educative activities for the subject, we considered the following types of exercises:

Test questions concerning theoretical contents as a review of methodologies and concepts (e.g. what is a grammar?, what is the difference between a language and a grammar?, which abstract machine recognizes Type-2 languages?, what is a token?, which are the main modules that make up a compiler? and which are the main methodologies studied to develop a syntax analyzer?);

Exercises with practical cases proposed to the student to obtain conclusions about the appropriate processes for resolve specific problems (e.g. detect the language recognized by a Finite Automaton, design a Push-Down Automaton to recognize a specific part of a programming

[3]https://aulaglobal.uc3m.es/

language, detect reserved words in a text by means of regular expressions and select a parsing grammar to detect valid if-then-else statements);

Use of the contents learned at the lab sessions and software programmes (like the JFLAP software[4] (Durham, NC, USA) for experimenting with formal languages topics). These activities were used to propose students to deal with practical implementations and detect which would be the results generated after code execution.

The completed evaluation of the practical m-learning multimodal application is focused on two main aspects. The first one is usability, measuring the benefits of multimodal interaction compared with unimodal interfaces (only spoken or tactile interaction). The second main aspect is the assessment of the technical quality and didactic potential.

5.1. Usability assessment

The methodology used to evaluate the multimodal interaction with the developed application is based on the work presented in (Metze *et al*., 2009), which points out that the usability assessment of a multimodal system requires the evaluation of the different interaction modalities. To do this, we have developed the assessment questionnaire shown in Table 1, which is based in standard questionnaires like AttrakDiff (Hassenzahl *et al*., 2003) and SASSI (Hone, 2014).

As it can be observed, the questionnaire consists of nine questions. Each question has five possible answers from which only one is selected. The main aspects that are evaluated are students previous experience using multimodal interfaces, the degree to which the student finds that the system understood them and they understood the system, the perceived interaction rate, the perceived difficulty level of the interaction with the system, the presence of errors, the certainty of the student of what to do at each time, and the global level of satisfaction with the system.

Firstly, the students evaluated the application using the tactile mode. Subsequently, they evaluated the application using the spoken interaction. Finally, they evaluated the multimodal mode combining the use of traditional interfaces (screen and keyboard) with spoken interaction. The students freely chose the actions to be taken and modules and sub-modules to be accessed during the evaluation.

Table 2 shows the results of the subjective evaluation using the described questionnaire. We can observe that the participants' previous experience using multimodal interfaces is very varied, which made possible to evaluate the system with users with different degrees of familiarity with these systems.

[4]http://www.jflap.org/

Table 1: *Results of the evaluation of the objective and subjective interaction parameters*

	Minimum/ maximum	Average	Standard. deviation
TQ01	3/5	4.18	0.67
TQ02	3/4	3.66	0.49
TQ03	4/5	4.81	0.41
TQ04	4/5	4.90	0.08
TQ05	4/5	4.63	0.46
TQ06	4/5	4.85	0.35
TQ07	4/5	4.82	0.39
TQ08	4/5	4.51	0.52
TQ09	4/5	4.81	0.39
TQ10	4/5	4.71	0.44
TQ11	3/5	4.52	0.75
DP01	4/5	4.88	0.09
DP02	4/5	4.62	0.47
DP03	4/5	4.86	0.37
DP04	4/5	4.86	0.07
DP05	4/5	4.64	0.43
DP06	4/5	4.83	0.35
LC01	4/5	4.46	0.22
LC02	3/5	4.13	0.68
LC03	4/5	4.77	0.31
LC04	4/5	4.81	0.38
C01	3/5	4.01	0.41
C02	3/5	3.73	0.75
A01	4/5	4.67	0.51
A02	4/5	4.93	0.12
A03	4/5	4.78	0.27
P01	4/5	4.17	0.54
P02	4/5	4.13	0.31
P03	4/5	4.07	0.56
SR	CR	ECR	
96.62%	12.80%	93.21%	

CR, confirmation rate; ECR, error correction rate; SR, success rate.

With respect to the extent to which the students feel that the system understood them, it is possible to see that the recognizer had a very good performance. Tactile mode was perceived as more accurate than oral as expected, but the oral mode was punctuated very high by the users, with 4.64 over 5 respectively. As can be observed, students felt that the system understood them better with the multimodal than with the tactile mode.

The students also felt that the system responses were comprehensible, especially with the tactile and multimodal modes. The results of the oral mode were lower probably because of the quality of the synthesized voice. We used the Standard English voice provided by Android. The results were higher when using better synthesizers like the voices of IVONA TTS. Also with the library catalogue functionality, sometimes the data of the bibliographical resources were in languages that differed from the one selected for TTS which reduced the intelligibility of the results in these particular cases.

Regarding the interaction rate, it was found adequate in most cases, though in some cases the participants reported they were expecting barge-in mechanisms in the oral mode. The multimodal mode was found very useful to for the

Table 2: *Questionnaire employed for the assessment of the technical quality, didactic potential, collaboration, authenticity and personalization functionalities*

Technical quality
TQ01. The application offers enough interactivity.
TQ02. The application is easy to use.
TQ03. It is easy to know what to do at each moment.
TQ04. The amount of information that is displayed on the screen is adequate.
TQ05. The arrangement of information on the screen is logical.
TQ06. The application is helpful.
TQ07. The application is attractive.
TQ08. The application reacts in a consistent way.
TQ09. The application complements the activities without distracting or interfering with them.
TQ010. The application provides adequate verbal feedback.
Didactic potential
DP01. The application fulfils the objective of assimilating the learning contents.
DP02. The contents worked in the activities are relevant for this objective.
DP03. The design of the activities was adequate for the contents of the subject.
DP04. The activities support significant learning.
DP05. The feedback provided by the agent improves learning.
DP06. The system encourages continuing learning after errors.
Learning contents
LC01. The activities were easy to understand.
LC02. The activities were easy to answer.
LC03. The application help me to learn new things.
LC04. The activities are in agreement with the contents learned at class.
Collaboration
C01. The application encourages me face-to-face discussion using my mobile device.
C02. The application encourages online discussion using my mobile device (e.g., via email, SMS, Skype, social media such as a Twitter or Facebook 'conversation').
Authenticity
A01. The mobile device is used in a suitably realistic learning setting.
A02. The application improves the resolution of exercises and information access.
A03. The application allows to use your mobile device in a better way.
Personalization
P01. The application can be customized according to your preferences.
P02. The proposed activities can be tailored for you.
P03. The application allows you to control the content and learning goals of the activities.

students to interact with the system at an adequate pace, as they could switch between interaction modes. This is in consonance with the fact that the multimodal mode has reported the maximum perceived easiness of use.

Generally, the students have not perceived errors during their interactions in the tactile mode, while in the oral and multimodal modes more errors were detected though they did not imply a fail in the interaction and in every case they could complete the task. In particular, the participants reported that the multimodal mode was more useful than the oral mode for reporting errors to the users.

About the certainty of what to do at each moment of the interaction, participants felt more security in the multimodal mode, which also received the better punctuation in the overall satisfaction as it brings more flexibility to the user.

5.2. *Technical quality and didactic potential*

The second part of the evaluation was focused on the assessment of the naturalness, pedagogical potential and technical quality of the system. The opinion survey shown in Table 3 was defined for the evaluation taking into account recent studies for the evaluation of these important aspects in educative e-learning and m-learning initiatives (Lucia *et al.*, 2009; Kearney *et al.*, 2015; O'Bannon and Thomas, 2015; Schmitz *et al.*, 2015). The responses to the questionnaire were measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The students were also asked to rate the system from 0 (minimum) to 10 (maximum), and there was an additional open question to write comments or remarks.

Also, from the interactions of the students with the system, we completed an objective evaluation of the application considering the following interaction parameters:

Question success rate. This is the percentage of successfully completed questions: system asks – user answers – system provides appropriate feedback about the answer;
Confirmation rate. It was computed as the ratio between the number of explicit confirmations turns and the total of turns;
Error correction rate. The percentage of corrected errors.

The results of the questionnaire are summarized in Table 4. As can be observed from the responses to the questionnaire, the satisfaction with technical aspects was high, as well as the perceived didactic potential. The m-learning application was considered attractive and adequate and the teachers felt that the system is appropriate and the activities relevant. The global rate for the system was 8.7 (in the scale from 0 to 10).

From the results of the evaluation, it can be observed that students positively evaluated the facility of obtaining the data necessary to complete the exercises and found the interaction rate suitable. The suggestions that they mentioned for the improvement of the system include the correction of system errors and a better clarification of the set of actions expected by the application at each time. Another interesting consideration concerns the correlation between the student background and the rest of scores. We verified that the questionnaire results are not influenced by the sample characteristics: user impressions are positive also when students did not have a previous experience with the use of m-learning applications.

The students highlighted the fact that the use of the application helped them to easier understand the contents

Table 3: *Results of the usability assessment*

Previous experience using multimodal interfaces				
Question	Avgerage value	Maximum value	Minimum value	Standard deviation
Q1	3.41	5	1	1.32
Q2	3.27	5	1	1.53
User understood by the system				
Question	Avgerage value	Maximum value	Minimum value	Standard deviation
Q3 (tactile mode)	4.81	4	5	0.38
Q3 (oral mode)	4.67	4	5	0.48
Q3 (multimodal mode)	4.85	4	5	0.36
System understood by the users				
Question	Avgerage value	Maximum value	Minimum value	Standard deviation
Q4 (tactile mode)	4.93	4	5	0.28
Q4 (oral mode)	3.91	3	5	0.71
Q4 (multimodal mode)	4.51	4	5	0.49
Interaction rate				
Question	Avgerage value	Maximum value	Minimumvalue	Standard deviation
Q5 (tactile mode)	3.65	3	5	0.64
Q5 (oral mode)	2.87	2	5	1.06
Q5 (multimodal mode)	4.79	3	5	0.52
Easiness of the interaction				
Question	Avgerage value	Maximum value	Minimum value	Standard deviation
Q6 (tactile mode)	4.34	3	5	0.56
Q6 (oral mode)	3.65	3	5	0.69
Q6 (multimodal mode)	4.84	3	5	0.37
Absence of errors				
Question	Average value	Maximum value	Minimum value	Standard deviation
Q7 (tactile mode)	4.82	4	5	0.47
Q7 (oral mode)	3.81	2	5	1.02
Q7 (multimodal mode)	4.46	4	5	0.52
Certainty of what to do at each moment				
Question	Average value	Maximum value	Minimum value	Standard deviation
Q8 (tactile mode)	4.02	3	5	0.46
Q8 (oral mode)	4.40	3	5	0.63
Q8 (multimodal mode)	4.84	4	5	0.37
Global satisfaction with the system				
Question	Average value	Maximum value	Mininum value	Standard deviation
Q9 (tactile mode)	4.43	4	5	0.49
Q9 (oral mode)	3.67	3	5	0.76
Q9 (multimodal mode)	4.85	4	5	0.38

1, worst; 5, best evaluation

of the subject. Students also positively evaluated the personalization functionalities of the application by means of the possibility of not only adapting the configuration options related to the interaction, but also for the customization of content and proposed activities based on the previous uses of the application. This point was found as a very important aspect related to the students' satisfaction and involvement in the educative experience. The students were very satisfied with the experience, not only because it facilitated learning but also because it was amusing for them. Additionally, the marks obtained in the partial assessments improved with respect to previous three years in which the m-learning application was not used.

Although the results were very positive, in the open question the students also pointed out desirable improvements. One of them was to make the system listen constantly instead of using the push-to-talk interface. However, we believe that this would cause many recognition problems, taking into account the unpredictability of environments in which the application can be used. Although the students liked the amount of options and tools that they had at their disposal to work in the application, for some of them it had a negative effect as sometimes they felt disoriented. Despite these minor negative impressions, a majority of students said that they would repeat the experience for other courses.

The results of the objective evaluation for the described interactions show that the developed application could interact correctly with the users in most cases, achieving a question success rate of 96.62%. Additionally, the approaches for error correction by means of confirming or re-asking for data were successful in 93.02% of the times when the speech recognizer did not provide the correct answer.

6. Conclusions and future work

The ubiquity, flexibility and increasingly diverse capabilities of mobile devices have created considerable interest from

Table 4: *Questionnaire designed for the usability assessment of the application*

Previous experience using multimodal interfaces	
Q1. Assess on a scale of 1–5 your previous experience using voice interfaces.	(1 = 'Low', 5 = 'High')
Q2. Assess on a scale of 1–5 previous experience using multimodal interfaces.	(1 = 'Low', 5 = 'High')
Understanding of user responses	
Q3. How well did the system understand you?	Never, Seldom, Sometimes, Usually, and Always
Understanding of system responses	
Q4. How well did you understand the system messages?	Never, Seldom, Sometimes, Usually, and Always
Interaction rate	
Q5. In your opinion, the interaction rate was…	Very slow, Slow, Suitable, Fast, Very fast
Difficulty level using the system	
Q6. Indicate the difficulty level of the system.	Very difficult, Difficult, Normal, Easy, Very easy
Presence of errors	
Q7. Have you noticed errors during the interaction?	Never, Seldom, Sometimes, Usually, and Always
Certainty of what to do at each moment	
Q8. Was it easy to decide what to do after each system turn?	Never, Seldom, Sometimes, Usually, and Always
Global satisfaction with the system	
Q9. In general, are you satisfied with the system?	Very dissatisfied; Dissatisfied; Satisfied; Quite satisfied; Very satisfied

educators in using them to enhance pedagogy with m-learning initiatives. Multimodal interactive systems offer users combinations of input and output modalities for interacting with these devices, taking advantage of the naturalness of speech. Different vendors offer APIs for the development of applications that use speech as a possible input and output modality, but developers have to design ad-hoc solutions to implement the interaction management.

In this paper, we have demonstrated the feasibility to use a general-purpose conversational framework to develop context-aware m-learning multimodal systems that can be easily integrated in hand-held Android mobile devices. The framework is comprised of an architecture in which different systems and modules cooperate to provide adapted services, and a representation mode for knowledge sharing between the components of the architecture. By means of our framework it is possible to develop multimodal interfaces that are a useful alternative to graphic user interfaces for mobile devices, allowing the use of other communication as an alternative to tapping through different menus.

To show the pertinence of our proposal, we have implemented an evaluated an Android m-learning app that provides practical educative exercises to promote independent learning and self-assessment. The application improves the interaction to solve the proposed exercises, personalizes the selection of the different learning activities considering the student's preferences using the system and their previous uses, and allows the provision of immediate feedback by means of the automatic assessment of the learning activities. The application also allows students to provide a first-person view of their activities to their classmates, record all learning interactions and instruction for replay and reflective reprocessing, and for professors collecting assessment data and monitoring and assessing student performance. We also provide context-aware information and services related to University Digital Libraries. To develop this system we have defined the complete requirements for the task and developed the different modules, and the necessary information to be incorporated for user's adaptation.

We have completed an evaluation of the developed application to assess the benefits of the multimodal interaction. To do this, the users employed the system with visual only, voice only and multimodal modes. The results show that the maximum satisfaction rates were achieved by the multimodal mode, as the users were able to switch between modalities and found this flexibility very useful. We have also assessed the naturalness, pedagogical potential and technical quality of the application. The results of this evaluation show a high satisfaction with technical aspects, as well as the perceived didactic potential.

We are currently undergoing the next phases in the deployment of the application. First, we want to conduct a more comprehensive evaluation of the system's functionalities during the next academic courses. We also plan to study the effect of using different pedagogical approaches that are adapted to different learning styles and use the student responses to generate models of frequent mistakes to provide more detailed automatic feedback to students and teachers. In addition, we are interested in studying the benefits that could be derived from including additional features in the user intention recognition model related to the user emotional state. With the results of these activities, we will optimize the system and make it available in Google Play.

Acknowledgements

This work was supported in part by Projects MINECO TEC2012-37832-C02-01, CICYT TEC2011-28626-C02-02, CAM CONTEXTS (S2009/TIC-1485).

References

Adegbija, M. and O. Bola (2015) Perception of undergraduates on the adoption of mobile technologies for learning in selected universities in Kwara state, Nigeria, *Procedia Social and Behavioral Sciences*, **176**, 352–356.

Aleven, V., Ogan, A., Popescu, O., Torrey, C. & Koedinger, K. (2004) Evaluating the Effectiveness of a Tutorial Dialog System for Self-Explanation, *in* Proc. of 7th Int. Conference on Intelligent Tutoring Systems (ITS'04), pp. 443–454.

Avouris, N. and N. Yiannoutsou (2012) A review of mobile location-based games for learning across physical and virtual spaces, *Journal of Universal Computer Science*, **18**, 2120–2142.

Baylor, A. and Y. Kim (2005) Simulating instructional roles through pedagogical agents, *International Journal of Artificial Intelligence in Education*, **15**, 95–115.

Beun, R., de Vos, E. & Witteman, C. (2003) Embodied Dialog systems: Effects on Memory Performance and Anthropomorphisation, *in* Proc. of Int. Conference on Intelligent Virtual Agents (IVA'03), pp. 315–319.

Bickmore, T. (2003) Relational Agents: Effecting Change through Human-Computer Relationships, PhD thesis, Media Arts and Sciences, Massachusetts Institute of Technology.

Blevins, B., K. Le Compte and S. Wells (2014) Citizenship education goes digital, *The Journal of Social Studies Research*, **38**, 33–44.

Boe, T., B. Gulbrandsen and O. Sorebo (2015) How to stimulate the continued use of ict in higher education: Integrating information systems continuance theory and agency theory, *Computers in Human Behavior*, **50**, 375–384.

Boticki, I., J. Baksa, P. Seow and C.-K. Looi (2015) Usage of a mobile social learning platform with virtual badges in a primary school, *Computers and Education*, **86**, 120–136.

Callahan, C., J. Saye and T. Brush (2014) Social studies teachers' interactions with second generation web-based educative curriculum, *The Journal of Social Studies Research*, **38**, 129–141.

Callejas, Z., Griol, D., McTear, M. & López-Cózar, R. (2014) A Virtual Coach for Active Ageing Based on Sentient Computing and m-health, *in* Proc. of 6th Int. Work-Conference on Ambient Assisted Living and Daily Activities (IWAAL'14), pp. 59–66.

Callejas, Z. and R. López-Cózar (2008) Relations between *de-facto* criteria in the evaluation of a spoken dialogue system, *Speech Communication*, **50**, 646–665.

Cassell, J., J. Sullivan, S. Prevost and E. Churchill (2012) Embodied Dialog Systems, The MIT Press.

Cavazza, M., de la Camara, R.-S. & Turunen, M. (2010), How Was Your Day? A Companion ECA, *in* Proc. of Int. Conference on Autonomous Agents and Multiagent Systems (AAMAS10), pp. 1629–1630.

Chen, H. and B. Albee (2012) An open source library system and public library users: finding and using library collections, *Library and Information Science Research*, **34**, 220–227.

Churchill, D. and T. Wang (2014) Teacher's use of ipads in higher education, *Educational Media International*, **51**, 214–225.

IDC Corporation (2015) Worldwide Quarterly Mobile Phone Q7 Tracker, IDC. Framingham: MA, USA.

de Rosis, F., Cavalluzzi, A., Mazzotta, I. & Novielli, N. (2005), Can embodied dialog systems induce empathy in users?, *in* Proc. of AISB'05 Virtual Social Characters Symposium, pp. 1–8.

Duta, N. and O. Martínez-Rivera (2015) Between theory and practice: the importance of ICT in higher education as a tool for collaborative learning, *Procedia - Social and Behavioral Sciences*, **180**, 1466–1473.

Dybkjaer, L., N. Bernsen and W. Minker (2004) Evaluation and usability of multimodal spoken language dialogue systems, *Speech Communication*, **43**, 33–54.

Ferdousi, B. and J. Bari (2015) Infusing mobile technology into undergraduate courses for effective learning, *Procedia Social and Behavioral Sciences*, **176**, 307–311.

Fryer, L. and R. Carpenter (2006) Bots as language learning tools, *Language, Learning and Technology*, **10**, 8–14.

Furió, D., M. Juan, I. Seguí and R. Vivó (2015) Mobile learning vs. traditional classroom lessons: a comparative study, *Journal of Computer Assisted Learning*, **31**, 189–201.

Gerger, K. (2014) 1:1 tablet technology implementation in the Manhattan Beach Unified School District: A case study, PhD thesis, California State University, Long Beach, California, United States.

Giménez-López, J., T. Magal-Royo, J. García-Laborda and F. Garde-Calvo (2009) Methods of adapting digital content for the learning process via mobile devices, *Procedia Social and Behavioral Sciences*, **1**, 2673–2677.

Gómez-Domingo, M. and A. Badia-Gargantè (2016) Exploring the use of educational technology in primary education: teachers' perception of mobile technology learning impacts and applications' use in the classroom', *Computers in Human Behavior*, **56**, 21–28.

Gratch, J., Rickel, J., Andre, J., Badler, N., Cassell, J. & Petajan, E. (2002) Creating Interactive Virtual Humans: Some Assembly Required, *in* Proc. of IEEE Conference on Intelligent Systems, pp. 54–63.

Grigoriadou, M., Tsaganou, G. & Cavoura, T. (2003) Dialog-Based Reflective System for Historical Text Comprehension, *in* Proc. of Workshop on Learner Modelling for Reflection at Artificial Intelligence in Education, pp. 238–247.

Griol, D., Z. Callejas, R. López-Cózar and G. Riccardi (2014) A domain-independent statistical methodology for dialog management in spoken dialog systems, *Computer, Speech and Language*, **28**, 743–768.

Griol, D. & Molina, J. (2014) A Framework to Develop Adaptive Multimodal Dialog Systems for Android-Based Mobile Devices, *in* Proc. of Int. Work-Conference on Hybrid Artificial Intelligence Systems (HAIS'14), pp. 25–36.

Griol, D., J. Molina, A. Sanchis and Z. Callejas (2012) A proposal to create learning environments in virtual worlds integrating advanced educative resources, *Journal of Universal Computer Science*, **18**, 2516–2541.

Griol, D., Patricio, M. & Molina, J. (2015) The CALIMACO Multimodal System: Providing Enhanced Library Services Using Mobile Devices, *in* Proc. of Int. Work-Conference on the Interplay Between Natural and Artificial Computation, Bioinspired Computation in Artificial Systems (IWINAC'15), pp. 339–348.

Hassenzahl, M., M. Burmester and F. Koller (2003) Mensch & Computer 2003. Interaktion in Bewegung, Vieweg and Teubner Verlag chapter AttrakDiff: A questionnaire for measuring perceived hedonic and pragmatic quality, 187–196.

Hone, K. (2014) Subjective assessment of speech-system interface usability, *in* Proc. of CHI'14 Workshop on Designing Speech and Language Interaction, pp. 69–72.

Hsu, Y. and Y. Ching (2013) Mobile computer-supported collaborative learning: a review of experimental research, *British Journal of Educational Technology*, **44**, 2120–2142.

Corporation, I. D(2015.)Worldwide Quarterly Mobile Phone Tracker, IDC.

Jeng, Y., T. Wu, Y. Huang, Q. Tan and S. Yang (2010) The add-on impact of mobile applications in learning strategies: a review study, *Journal of Educational Technology and Society*, **13**, 3–11.

Kanto, K. J. K., Kerminen, A. & Rissanen, J. (2004) Evaluation of Adaptivity and User Expertise in a Speech-Based E-Mail System, *in* Proc. of the 20th ACL Int. Conference on Computational Linguistic, pp. 44–52.

KARTAKIS, S. (2010) A design-and-play approach to accesible user interface development in ambient intelligence environments, *Journal Computers in Industry*, **61**, 318–328.

KAUFMANN, T. and B. PFISTER (2012) Syntactic language modeling with formal grammars, *Speech Communication*, **54**, 715–731.

KEARNEY, M., K. BURDEN and T. RAI (2015) Investigating teachers' adoption of signature mobile pedagogies, *Computers and Education*, **80**, 48–57.

KERLY, A., R. ELLIS and S. BULL (2008a) CALM system: a dialog system for learner modelling, *Knowledge Based Systems*, **21**, 238–246.

KERLY, A., ELLIS, R. & BULL, S. (2008b) Dialog systems in E-Learning, *in* Proc. of AI'08, pp. 169–182.

KING, L.J., C. GARDNER-MCCUNE, P. VARGAS and Y. JIMENEZ (2014) Re-discovering and re-creating African American historical accounts through mobile apps: the role of mobile technology in history education, *The Journal of Social Studies Research*, **38**, 173–188.

KLENNER, M. (2015) A technological approach to creating and maintaining media-specific educational materials for multiple teaching contexts, *Procedia Social and Behavioral Sciences*, **176**, 312–318.

KUO, M. and T. CHUANG (2016) How gamification motivates visits and engagement for online academic dissemination - an empirical study, *Computers in Human Behavior*, **55**, 16–27.

LAKKALA, M. and L. ILOMAKI (2015) A case study of developing ict-supported pedagogy through a collegial practice transfer process, *Computers and Education*, **90**, 1–12.

LATHAM, A., K. CROCKETT, D. MCLEAN and B. EDMONDS (2012) A conversational intelligent tutoring system to automatically predict learning styles, *Computers and Education*, **59**, 95–109.

LEE, J.K. and P.E. MOLEBASH (2014) Becoming digital: using personal digital histories to engage teachers in contemporary understandings of teaching social studies, *The Journal of Social Studies Research*, **38**, 159–172.

LU, J., S. MENG and V. TAM (2014) Learning Chinese characters via mobile technology in a primary school classroom, *Educational Media International*, **51**, 166–184.

LUCIA, A.D., R. FRANCESE, I. PASSERO and G. TORTORA (2009) Development and evaluation of a virtual campus on Second Life: The case of SecondDMI, *Computers & Education*, **52**, 220–233.

MAGAL-ROYO, T., I. TORTAJADA-MONTAÑANA, J. GIMÉNEZ-LÓPEZ and F. GIMÉNEZ-ALCALDE (2010) New educative methods in the usage of audiovisual content in mobiles, *Procedia Social and Behavioral Sciences*, **2**, 4492–4496.

MARSELLA, S., W. JOHNSON and C. LABORE (2003)Interactive Pedagogical Drama for Health Interventions, IOS Press, Amsterdam, The Netherlands, 341–348.

MCTEAR, M. and Z. CALLEJAS (2013) Voice Application Development for Android, Packt Publishing: Birmingham, UK.

METZE, F., I. WECHSUNG, S. SCHAFFER, J. SEEBODE and S. MOLLER (2009) Human-Computer Interaction. Novel Interaction Methods and Techniques, Springer-Verlag, chapter Reliable Evaluation of Multimodal Dialogue Systems, 75–83.

MOLLER, S., ENGLERT, R., ENGELBRECHT, K., HAFNER, V., JAMESON, A., OULASVIRTA, A., RAAKEAND, A. & REITHINGER, N. (2006) MeMo: towards automatic usability evaluation of spoken dialogue services by user error simulations, *in* Proc. of Annual Conference of the International Speech Communication Association (Interspeech'06), pp. 1786–1789.

MORENO, M., B. CASTAÑO, D. BARRERO and A. HELLÍN (2014) Efficient services management in libraries using AI and wireless techniques, *Expert Systems with Applications*, **41**, 7904–7913.

MOUZA, C. and T. BARRETT-GREENLY (2015) Bridging the app gap: an examination of a professional development initiative on mobile learning in urban schools, *Computers and Education*, **88**, 1–14.

MURPHY, G. (2011) Post-PC devices: a summary of early iPad technology adoption in tertiary environments, *E-Journal of Business Education and Scholarship of Teaching*, **5**, 18–32.

NORDIN, N., M. AMIN-EMBI and M.M. YUNUS (2010) Mobile learning framework for lifelong learning, *Procedia Social and Behavioral Sciences*, **7**, 130–138.

O'BANNON, B.W. and K.M. THOMAS (2015) Mobile phones in the classroom: preservice teachers answer the call, *Computers and Education*, **85**, 110–122.

O'HALLORAN, K. (2015) The language of learning mathematics: a multimodal perspective, *The Journal of Mathematical Behavior*, **40**, 63–74.

PADILLA-ZEA, N., F. GUTIÉRREZ, J. LÓPEZ-ARCOS, A. ABAD-ARRANZ and P. PADEREWSKI (2014) Modeling storytelling to be used in educational video games, *Computers in Human Behavior*, **31**, 461–474.

PAEK, T. and R. PIERACCINI (2008) Automating spoken dialogue management design using machine learning: an industry perspective, *Speech Communication*, **50**, 716–729.

PEDREIRA, O., F. GARCÍA, N. BRISABOA and M. PIATTINI (2015) Gamification in software engineering - a systematic mapping, *Information and Software Technology*, **57**, 157–168.

PÉREZ-MARÍN, D. and I. PASCUAL-NIETO (2011) Conversational Agents and Natural Language Interaction: Techniques and Effective Practices, IGI Global: Hershey, PA, USA.

PIERACCINI, R. (2012) The Voice in the Machine: Building Computers that Understand Speech, The MIT Press: Cambridge, MA, USA.

PON-BARRY, H., K. SCHULTZ, E.-O. BRATT, B. CLARK and S. PETERS (2006) Responding to student uncertainty in spoken tutorial dialog systems, *Int. Journal of Artificial Intelligence in Education*, **16**, 171–194.

ROUILLARD, J. (2007) Web services and speech-based applications around Voice XML, *Journal of Networks*, **2**, 27–35.

SALSE, M., M. RIBERA, R. SATORRAS and M. CENTELLES (2015) Multimodal campus project: pilot test of voice supported reading, *Procedia - Social and Behavioral Sciences*, **196**, 190–197.

SCHMITZ, B., R. KLEMKE, J. WALHOUT and M. SPECHT (2015) Attuning a mobile simulation game for school children using a design-based research approach, *Computers and Education*, **81**, 35–48.

SEABORN, K. and D. FELS (2015) Gamification in theory and action: a survey, *International Journal of Human-Computer Studies*, **74**, 14–31.

SENEFF, S., ADLER, M., GLASS, J., SHERRY, B., HAZEN, T., WANG, C. & WU, T. (2007) Exploiting Context Information in Spoken Dialogue Interaction with Mobile Devices, *in* Proc. of Int. Workshop on Improved Mobile User Experience (IMUx'07), pp. 1–11.

TRAUM, D. and S. LARSSON (2003) The Information State Approach to Dialogue Management, Kluwer, chapter Current and New Directions in Discourse and Dialogue: Berlin, Germany, 325–353.

URH, M., G. VUKOVIC, E. JEREB and R. PINTAR (2015) The model for introduction of gamification into e-learning in higher education, *Procedia Social and Behavioral Sciences*, **197**, 388–397.

VASSILAKAKI, E. and E. GAROUFALLOU (2013) Multilingual Digital Libraries: A review of issues in system-centered and user-centered studies, information retrieval and user behavior, *Information and Library Review*, **45**, 3–19.

WANG, Y. and A. ACERO (2006) Rapid development of spoken language understanding grammars, *Speech Communication*, **48**, 390–416.

WANG, Y., WANG, W. & HUANG, C. (2007) Enhanced Semantic Question Answering System for e-Learning Environment, *in* Proc. of the 21st Int. Conference on Advanced Information Networking and Applications (AINAW'07), pp.1023–1028.

Williams, J. and S. Young (2007) Partially observable Markov decision processes for spoken dialog systems, *Computer Speech and Language*, **21**, 393–422.
Wu, W.-L., R.-Z. Lu, J.-Y. Duan, H. Liu, F. Gao and Y.-Q. Chen (2010) Spoken language understanding using weakly supervised learning, *Computer Speech & Language*, **24**, 358–382.
Yang, X., X. Li and T. Lu (2015) Using mobile phones in college classroom settings: effects of presentation mode and interest on concentration and achievement, *Computers and Education*, **88**, 292–302.
Young, S. (2002) The Statistical Approach to the Design of Spoken Dialogue Systems, Technical report, Cambridge University Engineering Department.
Yue, W. and N. Zin (2013) Voice recognition and visualization mobile apps game for training and teaching hearing handicaps children, *Procedia Technology*, **11**, 479–486.
Zydney, J.M. and Z. Warner (2016) Mobile apps for science learning: review of research, *Computers and Education*, **94**, 1–17.

The authors

David Griol

Dr. David Griol obtained his Ph.D. degree in Computer Science from the Technical University of Valencia (Spain) in 2007. He has also a B.S. in Telecommunication Science from this University. He is currently visiting lecturer at the Department of Computer Science in the Carlos III University of Madrid (Spain). He has participated in several European and Spanish projects related to natural language processing and dialog systems. His research activities at the Applied Artificial Intelligence Group are mostly related to the development of statistical methodologies for the design of spoken dialog systems. His research interests include dialog management and optimization, corpus-based methodologies, user modeling and simulation, adaptation and evaluation of spoken dialog systems and machine learning approaches.

José Manuel Molina

Dr. José Manuel Molina is Full Professor at Universidad Carlos III de Madrid, Spain. He joined the Computer Science Department of the same university in 1993. Currently, he coordinates the Applied Artificial Intelligence Group (GIAA). His current research focuses on the application of soft computing techniques (NN, evolutionary computation, fuzzy logic, and multiagent systems) to radar data 1processing, air traffic management, and ehyphen; commerce. He (co)authored up to 50 journal papers and 200 conference papers. He received a degree in telecommunications engineering from the Technical University of Madrid in 1993 and the Ph.D. degree from the same university in 1997.

Zoraida Callejas

Dra. Zoraida Callejas is Assistant Professor in the Department of Languages and Computer Systems at the Technical School of Computer Science and Telecommunications of the University of Granada (Spain). She completed a PhD in Computer Science at University of Granada in 2008 and has been a visiting researcher in University of Ulster (Belfast, UK), Technical University of Liberec (Liberec, Czech Republic), University of Trento (Trento, Italy), University of Ulm (Ulm, Germany), Technical University of Berlin (Berlin, Germany) and Telecom ParisTech (Paris, France). Her research activities have been mostly related to speech technologies and in particular to the investigation of affective dialogue systems. She has participated in numerous research projects, and is a member of several research associations focused on speech processing and human-computer interaction.