English language learning activity using spoken language and intelligent computer-assisted technologies

Maria Fuentes, Meritxell González

TALP Research Center, Universitat Politècnica de Catalunya

mfuentes@lsi.upc.edu, mgonzalez@lsi.upc.edu

Abstract

This paper presents work in progress on language technologies applied to secondary school education. The application presented integrates several state-of-the-art technologies related to spoken language and intelligent computer-assisted language learning. We envision to show that the technology has reached a level of maturity that suggests that the time may be right to use it to second language learning. To achieve this objective, an activity was designed to be tested at several Spanish high schools. The aim was to carry out a proof of concept in real conditions and to obtain feedback from the students through a questionnaire as well as from the teachers by means of an interview. The activity was designed with the collaboration of some of the teachers at the secondary schools.

Index Terms: speech synthesis, speech recognition, natural language processing, dialog system

1. Introduction

This work is motivated by the Spanish education failure reported by the Program for International Student Assessment (PISA). This report is conducted by the Organization for Economic Co-operation and Development (OECD), which is responsible for standardized testing for 15-year-old students.

We think that the use of Information and Communication Technologies (ICT) can definitely help to improve some aspects of the educational process. On the one hand, the incorporation of ICT in the classroom may contribute to reduce the digital gap between teachers and students. On the other hand, our intuition is that embedded Artificial Intelligence components are probably crucial to create successful systems targeting educational aids and tools.

To encourage the use of ICT in educational environments several aspects of the learning process have to be taken into account. The human expertise is basic to correctly identify how to apply the technology to solve a specific problem in such environment. In addition, this information is necessary to codify the "know-how" of any Expert System. For that reason, the chair that Telefónica has at UPC[1] organized and sponsored, in December 2010, a meeting with all the agents involved in the educational system. The chair goal was to promote the use of speech and natural language processing technologies in the educative process.

The objective of the meeting was i) to detect and identify educative needs as well as technological restrictions and ii) to lay out a pilot to be carried out by the end of April. More specifically, the main efforts were focus on how to apply existent language technologies for second language learning.

The use of language technologies have been increasing in the last years. A good overview of the state of the art in speech technology is provided by [2]. In addition, [3] explores the relevance and uses of Natural Language Processing (NLP) methods in the context of language learning, focused on written language. Moreover, from the point of view of second language acquisition, [4] investigates how to increase the chances that second language readers look up and learn unfamiliar words during and after reading a text.

2. Application Framework

The challenge of the meeting was to find a good combination of some of the abilities of the existing tools to propose a demonstration in a real educational environment. We fixed 4-month for developing a demonstrative pilot. Having in mind this important time restriction, we were requested to design a virtual assistant having as much abilities as possible.

We considered and analyzed the state of the art of existing and available technologies and tools in order to identify which of them could be better integrated in the framework that we were designing. The opinion of the representatives for the four main agents involved in education (i.e. educators, technologists, publishing companies and administration) was considered essential. Following this objective, a panel discussion was organized during the meeting in which members of these agents shared their knowledge and experience in the field.

Table 1 shows the tools that were candidates to help to improve various aspects of oral and written communication, and which were presented to the audience. These tools were classified on the ground of five skills involved in communication and discourse: vision, hearing, cognitive decoding of language reading, comprehension and speech production. Additionally, we assigned to each tool a set of actions enabled for interaction: keyboard interpretation, avatar, status update and evaluation.

2.1. Pilot Framework

This section summarizes the information gathered during the debate and proposes an educational environment in which to apply the analyzed technologies and tools. Our general testing scenario is a virtual English Language Learning assistant.

The activity should allow the practice of as much aspects of the language as possible. When possible, this practice would be developing the real life tasks.

During the activity, the teacher role will be guiding and coaching individually, leveraging the best way to help each student. Some of the functionalities of the wizard will help the teachers with the task of assessing the evolution of the student. In this way, the system will provide personalized assistance tailored to each student needs.

The pilot test was held in several secondary schools within courses of about 20 or 30 students at the age of 12 or 13. The

Application	Author	Activity (ability, actions)
Vocaliza [5]	Zaragoza University	Fonetics (hearing, cognitive, evaluation)
Foundations To Literacy [6]	BLTEK, Colorado University	Fonetics (speech, keyboard interpretation, cognitive, evaluation)
Interactive Book [6]	BLTEK, Colorado University	Reading (Avatar, reading and evaluation)
FLORA [7]	BLTEK	Reading (hearing and evaluation)
Cuéntame [8]	Zaragoza University	Conversational (hearing, cognitive, speech and status update)
MyST [9]	BLTEK, Colorado University	Conversational (hearing, cognitive, speech and status update)
WERTi [10]	Tübingen University	Grammar (text converted in a grammar exercise, evaluation)
Interactive3DFramework [11]	Politècnica de Catalunya University	3D (cognitive, reading, vision)
Autolearn [12]	Barcelona Media, Pompeu Fabra University	Comprenhension (cognitive, reading, evaluation)

Table 1: Applications that assists the improvement of some aspect of oral or written communication.

log of the user interaction, answer contents and results were collected to be analyzed. In order to obtain feedback from the students, they were asked to answer a questionnaire about their opinion of the tools used the end of the activity. Moreover, the teachers were interviewed about their opinion with respect to the pedagogical capabilities of the set of tools and the global environment.

3. Integrated Technologies

The proposed framework provides a learning environment for training fluent and expressive reading with the objective of comprehension training and assessment.

The activity consists of several tasks; each demonstrates the use or the application of a specific technology for addressing the main goal of the activity.

Details about each specific technology are reported in the following subsections.

3.1. Virtual Learning Environment

The activity has been designed and integrated into the virtual learning environment Moodle 1.9. [1]

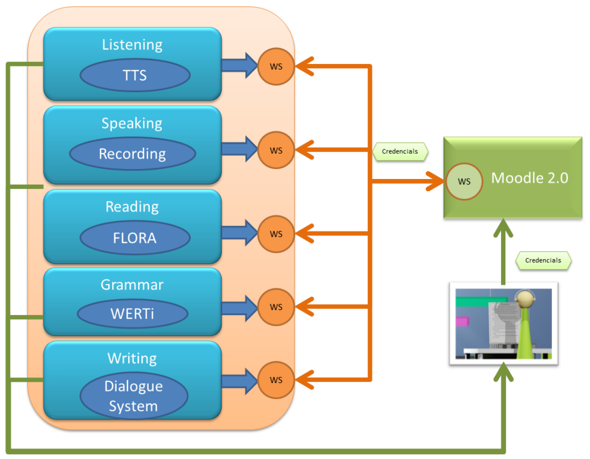

Figure 1: Architecture of the virtual environment

[1]The virtual environment can be accessed on-line at the following URL: http://nlp.lsi.upc.edu/alice/ .

The current implementation of the activity consists of five separated sections, each linked to the specific URL that contains the particular implementation of each task: listening, speaking, reading, grammar and writing (the architecture is shown in Figure 1). Although nowadays the virtual framework integrates five different tasks in a single activity, the number of sections is flexible and can be reduced or extended.

The tasks sections are presented in a predetermined order, but they can be reordered for each Moodle-course. At the beginning of the activity, only the first section is active. The rest of the sections can be enabled using a SOAP web service published by the activity environment. Then, the next section is activated once the current task has been completed. Additionally, the relevant information about the student interaction is also sent to the web service and properly recorded in a specific format into the Moodle platform. This information can be consulted by the teacher (for instance, for evaluation purposes), by the student (to know his/her progress) or by a program that can process this information and extract data about the student, the group/course or about the activity. Examples of this information are: the list of words that the student considers difficult to pronounce, the oral reading fluency score, the complete dialogue or the results of the grammar exercise.

3.2. Guidelines

The activity starts showing to the user the guidelines of the learning exercise. The guidelines are a flash movie that shows the control buttons of each task and explains in native language the goal of each task. The spoken messages have been produced by the speech synthesizer implemented by the Verbio Company and publicly accessed at their website[2].

Additionally, a short description of the task is played at the beginning of each task (when the specific section is activated). The language used for these messages is English and they have been recorded using the synthesizer described in section 3.3.

3.3. Listening Task

The listening task is the first touch of the student with the text chosen for the exercise[3]. The text is presented in separated paragraphs and, with the aim of helping in the comprehension pro-

[2]http://www.verbio.com/webverbio3/html/demos_ttsonline.php

[3]The English teachers chose a text from the DIBELS materials (http://dibels.uoregon.edu/)

cess, images were added near to the text of each paragraph.

The students are requested to select part of the text and listen the pronunciation of the words as spoken by the synthesizer. A punctuation score is given to the student at the end of the task, which is related to the number of listened words. Several words considered difficult to pronounce were chosen to give extra-points when any of them was listened.

We have integrated the Text-To-Speech Enabler synthesizer developed by Ericsson and published at the Ericsson Lab site[4]. The synthesizer provides a simple Java-based API that is accessible from the web page. Hence, dynamically, the task catch the text selected by the student in the web interface and sends the text to the synthesizer. In turn, the synthesizer returns an audio stream that is reproduced by the web browser.

3.4. Speaking Task

In the second task integrated in the virtual environment, the students are required to read a piece of the text, record their own speech and listen to their selves. The task consists of i) reading the same words that they were requested to listen in the previous task, and ii) comparing their pronunciation with the one produced by the synthesizer. The goal is that students can compare the pronunciation of any word and learn the phonetics of the language by means of practicing a set of examples. Ideally, the students read longer phrases or even the whole paragraph at a once and learn some prosody.

A Java applet, developed by Javasonics[5], has been integrated and used to record the students voice and upload the speech file to the web server. We used the evaluation version of the applet which allows recording one minute of speech. Nonetheless, the applet functions can be accessed repeatedly and then, the student can record different parts of the text as many times as required.

3.5. Reading Task

The goal of the Reading task is to estimate the student's oral reading fluency. For this purpose we have integrated the FLORA web-based system [7]. FLORA presents grade-level texts to children, who read out the text aloud, and computes the number of words correct per minute (WCPM), a standard measure of oral reading fluency. The researchers at BLTek have a large database of children's speech to train their ASR and read-tracking system.

The integration allows connecting the client with the web service at the BLTek offices at Boulder Colorado and controlling the interaction and logging data at our servers. The process is transparent to the student, which simply access to the corresponding task section in the Moodle framework. Figure 2 shows the architecture of the reading task, which includes the credentials from the Moodle framework and the connection to the FLORA Server at BLTek. The rest of the tasks integrated into the virtual environment follow a similar architecture.

3.6. Grammar Task

The main goal of the grammar task is to practice one or more aspects of the English grammar. In this task we have integrated the WERTi extension for the Firefox browser[6].

In our particular activity, the grammar exercise consists of seeking and clicking on target words (e.g. prepositions or determiners) in the text. The punctuation of the task is a F1-score of the hits and the fails.

[4]https://labs.ericsson.com/apis/text-to-speech/

[5]http://www.javasonics.com

[6]http://sifnos.sfs.uni-tuebingen.de/WERTi/

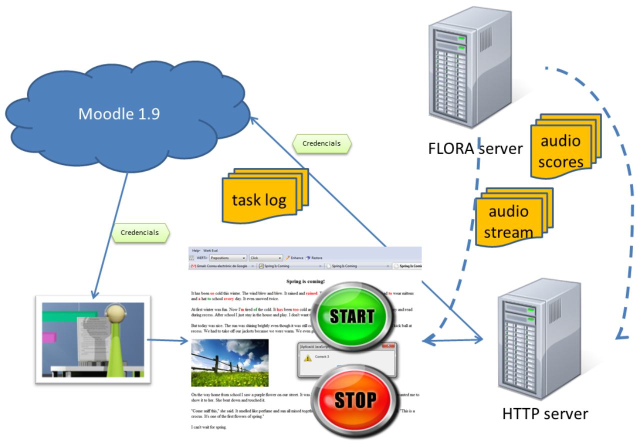

Figure 2: Architecture of the reading task

The flexibility and usability of the tool allow us to lay out personalized exercises for the varying abilities of the students. WERTi [10] allows enhancing any web page with morphological, syntactic, semantic and pragmatic features to learn English. This tool includes the most common exercises for learning English as second Language: determiner and proposition usage, uses of gerunds vs. infinitives, wh-questions form, phrasal verbs, and noun countability. The WERTi tool also displays different types of enhancements depending of the type of practice chose by the user (i.e. search for the target word and click, select from a list or cloze input).

3.7. Writing Task

The last task of the activity consists of chatting with a dialogue system. The main goal of the task is to assess the comprehension of the text.

The system asks to the student several questions related to the main theme of the text. However, the answer cannot be answered by taking a piece of text. The system analyzes the student's answer and performs different type of actions depending on its content (e.g. in case of correct response, the system gives some positive feedback and asks the next question). Figure 3 shows the chat interface and includes a piece of dialogue.

We have adapted a plan-based dialogue system[7] to drive a conversation with the user. The plan for the dialogue consists of several goals with different actions in a way that all the goals achieve the end of the plan. So, the conversation has several variations depending on the content of the student's answers.

For instance, the system asks the student what dishes with vegetables he/she knows. In case that the student forgets to mention the French Fries, the system explains that French Fries are also a fried vegetable.

The student is also required to answer the questions using correct spelling. Otherwise, incorrect spelling produces system misunderstanding of the answers. Consequently, the student is required to write strictly in English.

The plan-based dialogue system is described in [13]. It uses three separated knowledge bases to represent the different types of knowledge involved in communication: domain, dia-

[7]http:nlp.lsi.upc.edu/digui

logue and linguistic knowledge. In order to adapt the system, we developed new resources for the domain and the linguistic knowledge, which are application-dependent.

The domain resources consist of the communication plan (namely a set of dialogue actions and attributes) and the specification of the application tasks (i.e. the operations and parameters of the tasks, the constraints and conditions of the task, and a simple taxonomy of attributes involved in the task).

The linguistic resources are twofold. On the one hand, we created a domain-specific grammar and lexicon to process and interpret the user's intervention. This grammar feeds a domain- and language-independent syntatico-semantic parser that was adapted for the dialogue system. On the other hand, we created the linguistic resources to express the system's interventions. These resources, used by a domain- and language-independent language generator, consist of a set of phrases on the grounds of a set of dialogue acts, the concepts and the attributes of the domain.

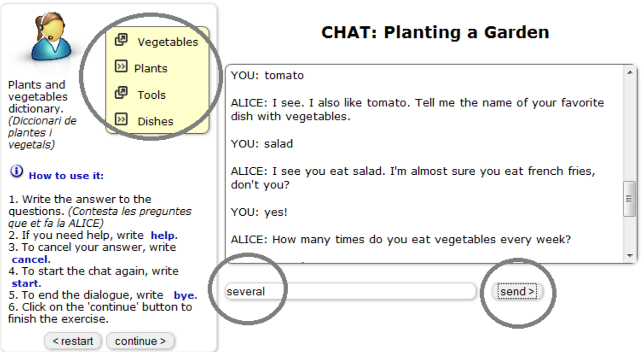

Figure 3: The Chat interface and an excerpt of dialogue

4. Conclusions and Future work

The application described integrates several state-of-the-art technologies related to spoken language and intelligent computer-assisted language learning.

In general, the demonstrative pilot carried out at secondary school was really positive. This first experiment has allowed us to empirically verify the performance of the proposed tools in a real environment. Even though the prototype has limited capabilities, what can be gathered from the questionnaires and the interviews is that both students and teachers are interested in using the proposed technologies.

Future directions depend on the abilities or the activities of the learning assistant to be promoted. We envision several evolutions of the current language learning assistant prototype:

- More cognitive and comprehensive abilities can be integrated in the dialogue system in order to evaluate the meaning of the student's answers according to the comprehension questions. In addition, the assistant could report lexical, syntactic, and semantic assessments.
- Multimodal interaction can be included. Instead of testing the comprehension level by answering questions, the student would realize different visual activities (e.g. relational or simulation based activities).
- Additional feedback can be included in each task. As an example, during the fluency test, the assistant can provide feedback about the pronunciation or the prosody.

5. Acknowledgements

This work has been partially funded by the Telefónica Chair at UPC. We want to thank Susanna and Enrique, English teachers at secondary school, for the important role they are playing for integrating ICT at the schools. We are sincerely grateful to Elena and the group of English teachers that participated in the brain storming session during the design process. We also express our gratitude to Ron Cole and his team and Detmar Meurers and his colleagues for their efforts in help us when integrating and adapting their tools. Finally, our work has benefited from the Moodle module developed by UPCnet.

6. References

[1] Fuentes, M., González, M., Guardiola, M., Jofre, L., Romeu, J., Vallverdú, F. "Cuaderno: Aprendizaje y Asistencia Virtual en Red". Cátedra Telefónica-UPC, 2011.

[2] Bejar, I., "Can Speech Technology Improve Assessment and Learning? New Capabilities May Facilitate Assessment Innovations", RD Connections Report No. RDC-15, 2010.

[3] Meurers, D.,"Natural Language Processing and Language Learning", in C. Chapelle [Ed], Encyclopedia of Applied Linguistics, to appear.

[4] Peters, E., Hulstijn, J.H., Sercu, L., Lutjeharms, M. "Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared". Language Learning, 59, 113-151,

[5] Vaquero, C., Saz, O., Lleida, E., Marcos, J.M., Canals, C. "Vocaliza: A Computer-Aided Application for Spanish Speech Therapy". 4as Jornadas en Tecnología del Habla, Zaragoza , 2006.

[6] Cole, R., Wise, B., Van Vuuren, S. How Marni teaches children to read. Educational Technology, 24:1, pp. 14-18, 2007.

[7] Bolaños, D., Cole, R., Ward, W., Borts, E., Svirsky, E. "The FLORA, Fluent Oral Reading Assessment for Children's Speech". Submitted in June 2010 to the ACM Transaction on Speech and Language Processing (Special Issue on Speech and Language Processing of Children's Speech for Child-Machine Interaction Applications).

[8] Rodríguez, W. R., Saz, O., Lleida, E., Vaquero, C., Escartín, A. "COMUNICA: Tools for Speech and Language Therapy". Workshop on Child, Computer and Interaction, Chania, Grecia, 2008.

[9] Ward, W., Cole, R., Bolaños, D., Buchenroth-Martin, C., Svirsky, E., Van Vuuren, S., Weston, T., Zheng, J. "My Science Tutor: A Conversational Multi-Media Virtual Tutor for Elementary School Science". Special Issue on Speech and Language Processing of Children's Speech for Child-machine Interaction Application, 2010

[10] Meurers, D., Ziai, R., Amaral, L., Boyd, A. Dimitrov, A., Metcalf, V., Ott, N. "Enhancing authentic web pages for language learners". In Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications (IUNLPBEA '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 10-18, 2010.

[11] Vázquez, P.-P., Hartmann, K., Gützelmann, T., Närnberger, A.. "An interactive 3D framework for anatomical education". International Journal of Computer Assisted Radiology and Surgery, Vol 3, Number 6, pp. 511-524. Ed. Elsevier, ISSN: 1861-6410, 2008.

[12] Estrada, M., Navarro, R., Quixal, M. A framework for the evaluation of language learning platforms. In Proceedings of the e-Society IADIS International Conference, vol. II. Ed. by Piet Kommers and Pedro Isaías. Pages 95-99, Barcelona, February, 25th-28th, 2009.

[13] González, M. "DIGUI: A flexible dialogue system for accessing web sevices". Phd Thesis, Universitat Politècnica de Catalunya, 2010