Toward Interactive Context-Aware Affective Educational Recommendations in Computer Assisted Language Learning

Olga C. Santos, Mar Saneiro, Jesus G. Boticario, M. C. Rodriguez-Sanchez

Abstract. This work explores the benefits of supporting learners affectively in a context-aware learning situation. This features a new challenge in related literature in terms of providing affective educational recommendations that take advantage of ambient intelligence and are delivered through actuators available in the environment, thus going beyond previous approaches which provided computer-based recommendation that present some text or tell aloud the learner what to do. To address this open issue, we have applied TORMES elicitation methodology, which has been used to investigate the potential of ambient intelligence for making more interactive recommendations in an emotionally challenging scenario (i.e., preparing for the oral examination of a second language learning course). Arduino open source electronics prototyping platform is used both to sense changes in the learners’ affective state and to deliver the recommendation in a more interactive way through different complementary sensory communication channels (sight, hearing, touch) to cope with a universal design. An Ambient Intelligence Context-aware Affective Recommender Platform (AICARP) has been built to support the whole experience, which represents a progress in the state of the art. In particular, we have come up with what is most likely the first interactive context-aware affective educational recommendation. The value of this contribution lies in discussing methodological and practical issues involved.

Keywords: context-aware learning, Arduino, TORMES methodology, computer-assisted language learning (CALL), interactive recommendations, affective computing, ambient intelligence, universal design

1. Introduction

Context-aware environments are sensitive and responsive to the presence of people and can provide what it is called ambient intelligence (Aarts et al., 2001). Applying ambient intelligence in educational scenarios allows learners to be immersed in a digital environment that is aware of their presence and context, and which also suits their needs through personalizing and adapting the learning environment to enable natural interactions (Santana-Mancilla et al., 2013), including the adaptations that are needed to cater for people with disabilities (Ntoa et al., 2013). Within this context, benefits in delivering personalized feedback based on student’s emotional states have been detected (Cooper et al., 2009).

In a previous work (Santos & Boticario, 2014) some educational emotion-aware recommendations rules were identified in the literature (Shen et al., 2009; Boff & Reategui, 2012; D’Mello & Graesser, 2013; Kaklauskas et al., 2013; Leony et al., 2013). These rules do not take advantage of ambient intelligence in their delivery, and consist of textual messages or speech dialogues provided through computers. However, in an intelligent environment the learner can receive recommendations through a more natural interaction in terms of visual, audio or tactile feedback. For instance, while
solving mathematical problems in an intelligent tutoring system, if the learner shows affective changes that may affect their performance, the chair where the learner is sat can vibrate to raise the learner awareness on her affective state, and for instance, produce some specific signal that the learner can recognize as a recommendation to make her slow down and think more deeply on the current task. With this framework in mind, we are facing the challenge of identifying recommendation opportunities in emerging context-aware learning environments. Thus, in (Santos & Boticario, 2014) we explored the potential of Arduino¹, together with appropriate sensors and actuators, to elicit affective educational context-aware recommendations that can be delivered in alternative interactive ways by making the environment more intelligent. Specifically, we discussed the human factors involved in the recommendation process, both in the input (i.e., detecting contextual information that reflects learners’ affective states) and the output (i.e., delivering affective recommendations) within the context in which the learner is placed. To the best of our knowledge –supported by a review of 82 recommender systems for technology enhanced learning scenarios (Draschler et al., 2015)–, to date there have not been other outcomes reported in the literature that reflect the use of Arduino to make educational recommender systems that interact with the environment. Further clarifications in this respect are provided in Section 2.

In this paper we further explore the potential of ambient intelligence (provided by an Arduino based infrastructure) to deliver more interactive educationally oriented recommendations. In particular, we have elicited an educational scenario in which the intelligence given to the environment is used to provide added value to the recommendation support. This scenario is focused on helping the learner when preparing for the oral examination of a second language learning course. In it, the learner talks aloud for five minutes about a given topic in a foreign language to practice for the oral exam. When the system detects that the learner is getting nervous, the system interacts with the learner in a way that does not interrupt her cognitive task and allows her to keep talking while she decides whether or not to follow the recommendation offered, which in this case, consists in suggesting the learner to reduce her breathing speed in order to calm her down. Providing a supportive textual description in the screen could be possible, but would make the learner stop talking to read. Thus, the idea is to code the recommendation in a sensorial way, which the learner understands as she has previously been told about its meaning. The idea behind this way of providing a recommendation is to interact with the learner in a more natural way when problematic affective situations are detected.

The rest of the paper is organized as follows. First, Section 2 reviews relevant literature. Next, in Section 3, we provide the details of applying TORMES elicitation methodology. Here we show how TORMES has served to develop the aforementioned educational scenario where an interactive context-aware affective educational recommendation has been elicited to be delivered in runtime taking advantage of ambient intelligence. This way of providing recommendations, to the best of our knowledge, represents the first instance in related literature where ambient intelligence features are used to enrich the feedback provided to learners in order to help them in dealing with their affective states. In this section, it is described how the TORMES methodological approach has been complemented with the implementation of an

¹ Arduino is an open source electronics prototyping platform, which is based on easy to use hardware and software intended for anyone making interactive projects (Banzi, 2009).
Arduino based infrastructure that can sense changes in the learners’ affective state and deliver recommendations in a more interactive way. After that, in Section 4 the experimental design and the results from a pilot study with users of the interactive recommendation proposed in the scenario are detailed. In Section 5, we focus the discussion on the limitations of our contribution, open issues and on-going works in relation to: i) how to deliver interactive recommendations, ii) when recommendations are to be provided, iii) what learners’ features seem to be of relevance, and iv) how to deal with social aspects. Finally, in Section 6 we conclude and outline future work.

2. State of the Art

In (Santos & Boticario, 2014) we researched the available literature that has reported the use of Arduino in context-aware recommender systems, covering both i) gathering context information, and ii) delivering context-based feedback. As to the former, we found 10 works that reported the usage of Arduino to gather physiological measures. Although data collected in those works is not used to detect the affective state of the user (but mainly for health monitoring), the same data can be used for this other purpose (Novak et al., 2013). Regarding the delivery of context-based feedback, we found 2 works that used audiovisual adaptations to reflect or evoke certain affective state in the user by providing feedback in terms of light (García-Perate et al., 2013; Guerreiro, 2013). In particular, the second work (Guerreiro, 2013) made use of the BITalino hardware platform (Silva et al., 2014), which although it is not built on top of Arduino, it follows the same low-cost prototypical approach and it is aimed to support rapid prototyping of end-user applications in the field of physiological computing. We did not find related research which is specific for the educational domain. Nevertheless, these works suggested that the utilization of Arduino in our research would be of interest both: i) to gather contextual information about the learner’s affective state, and ii) to use this information to deliver in a more interactive way recommendations that take into account changes in learners’ affective state.

After that review, works that deal with emotional information gathering with Arduino have been reported. In particular, Iwasaki et al. (2014) use the e-Health Sensor Platform\(^2\) to record electrocardiography signals when emotions are elicited with standardized movie clips from the Emotional Movie Database (Carvalho et al. 2012). Jaimovich et al. (2014) measure electrodermal activity and heart rate signals while music, chosen randomly from a pool of 53 songs with different valence and arousal values, is played. Other works, such as the HIPOP platform, focus on the multimodal acquisition of physiological signals, eye gaze, video and audio to perform an integrated affective and behavioral analysis (Lazzeri et al., 2014).

In turn, regarding delivery of context-based emotional feedback with Arduino, works found keep their focus on using light as the main feedback mechanism, although some of them also explore vibro-tactile feedback. Negru (2010) collects the user emotions from a set of questions and displays the collected emotional information to the user through a set of red, green and blue light-emitting diodes (LEDs) in different combinations (depending on the input) in order to express a diverse range of feelings.

---

Along with these colored LEDs, this work also uses an alpha-numeric liquid-crystal display (LCD) to show a predefined message based on the feeling identified. In addition, it is also suggested for future work that the system can set the light intensity or turn on the music in terms of the emotions gathered. Hao et al. (2014) have developed a wristband prototype for emotion regulation with visual feedback, which incorporates a series of colored LEDs that map the user’s affective state while watching video clips that evoke an agitated state. Lights are used to calm the user. In turn, Pradana et al. (2014) explore the emotional augmentation benefits of vibro-tactile stimulation, color lighting and the simultaneous transmission of both signals to accompany text messages. They have developed Ring U, a ring-shaped wearable system aimed at promoting emotional communications between people. They find that touch and color stimuli is effective to invoke and change the emotional perception to a text message, and it can be driven into the direction of the emotional characteristic of the stimuli. Another wearable system is WearBREATH (Min and Nam, 2014), which is defined as a breathing movement sharing device for affective communication and connectedness. It measures user’s abdominal breathing movements and translate them to squeezing movements of a bracelet that is worn by an intimate partner (e.g., father and child). In turn, SWARM is an intelligent scarf that can provide six different actuations (i.e., lights, vibration, music, heat, cooling, weights) for personal emotion awareness (Williams et al., 2015). In Hug Me (Fontes et al., 2014) a mannequin with sensors detects when a participant hugs it and triggers a blinking red LED in the mannequin’s chest (the place of the heart) along with the sound of a heartbeat to provides real-time immediate feedback to the participant. When the system recognizes a given feeling, it triggers glowing patterns of different colors related to specific feelings (love, happiness, longing, discomfort) and also changes the color of the walls outside the installation location. In this line, an anthropomorphic lamp has been proposed for the communication of emotions that is able to represent and collect users’ emotional states through a multimodal interaction based on tangible gestures aimed to make the interaction more natural (Angelini et al., 2014). Finally, visual-to-tactile mapping of facial movements have been proposed to build a sensory substation system for blind people by adding vibro-tactile stimulation on a chair that codes facial movements according to the Facial Action Coding System in terms of corresponding vibration patterns that can be felt through the users’ back (Bala et al., 2014).

As in our previous review, there have not been found educational scenarios involved in reported research. Therefore, there is a challenge and opportunity to research if the ambient intelligence support that can be deployed with a recommender system extended with an Arduino infrastructure is of value in educational settings.

3. Application of TORMES methodology

In order to explore the potential of ambient intelligence with Arduino to make more interactive recommendations in learning environments, we have applied TORMES methodology (Santos & Boticario, 2015). TORMES methodology follows the standard ISO 9241-210 (ISO, 2010) in order to guide the involvement of educators in the elicitation process (so their experience can be automated to provide adaptive personalized support with a recommender system) and is based on: 1) user centered design methods to gather educators’ tacit experiences, such as those acquired when supporting learners emotionally in e-learning scenarios, and 2) data mining techniques to detect relevant learners’ behavior from previous interactions with the system (when these data are available), such as specific situations that can influence their affective...
states. TORMES has successfully been used to elicit recommendation opportunities in diverse scenarios, namely: i) learning by doing in dotLRN learning management system, ii) reviewing concepts in Willow free-text adaptive computer assisted assessment system, iii) structured collaborative tasks as the Logical Framework Approach, and iv) math problems solving with affective support.

TORMES follows the four user centered design activities defined by ISO 9241-210 in an iterative manner: 1) Understanding and specifying the context of use: identifying the people who will use the system, what they will use it for, and under what conditions they will use it; 2) Specifying the user requirements: identifying any requirements or user goals that must be met for the system to be successful, considering the variety of different viewpoints and individuality; 3) Producing design solutions to meet user requirements, which can be made in stages to encourage creativity, from an initial rough concept to a final full-blown design; and 4) Carrying out the user-based evaluation of the design against the requirements.

As anticipated in Santos & Boticario (2014), TORMES could also be used to identify recommendation opportunities in educational scenarios that take advantage of available contextual information, such as environmental information (temperature, humidity, luminosity, etc.), the state of mind of the learner (bored, frustrated, motivated, etc.) and physiological information of the learner (heart rate, skin conductance, body temperature, etc.), among others. From the analysis carried out in that work (and in line with more recent works reported in Section 2), in our view Arduino and available sensor kits, such as the e-Health sensor platform to register physiological information, can facilitate rapid prototyping of low cost infrastructure for sensing the learning context (and thus, building the infrastructure required for affective contextual information gathering), as well as for delivering appropriate affective feedback with corresponding actuators (and thus, taking advantage of ambient intelligence to provide a more natural and context-aware interaction).

In order to identify recommendation opportunities from educational practice that take advantage of ambient intelligence, TORMES guided the recommendations’ elicitation process. As a result, one interactive recommendation aimed to provide affective support at runtime in a specific learning scenario has been elicited. Afore-mentioned design activities correspond to the identification, definition, implementation and evaluation of the elicited scenario where the recommendation opportunity appeared, as described below.

**Activity 1. Understanding context of use (scenario identification)**

The scenario “preparing for the oral examination of a second language learning course” was proposed as the starting point to understand a context of use where adding intelligence could enrich the recommendation process. This scenario requires the competence of speaking in public, which is widely used in the literature to induce stress (Plarre et al., 2011), because it produces substantial variations in physiological responses (Novak et al., 2013). Stressing situations may occur in several educational contexts, such as the one reported here. In particular, the ability to speak with others in a foreign language is critical, and when it is realized it turns into one of the most challenging moments for learners, as the anxiety that may occur while speaking has an impact on the speaking performance (Woodrow, 2006). Successfully talking to an audience requires that the speaker draw on her emotional resources (Dewaele et al., 2008).

A review of the literature regarding context-aware personalized support through recommendations in computer-assisted language learning shows that research has not
faced the challenges of oral communication, but it has focused on other skills, such as providing resources for learning vocabulary and expressions within the learner’s context and taking into account time information, learning activity and device status (Li et al., 2013). In any case, up to our knowledge, there have not been previous instances where emotional or affective issues during the language learning process have been considered. From a technological viewpoint, online videoconferencing environments are appropriate in language teaching (Hampel & Stickler, 2012).

Having this in mind, we decided to explore the recommendations opportunities that an ambient intelligent scenario can provide for preparing an oral examination of a second language learning course using a videoconferencing system. The corresponding requirements for this context of use were defined in the next TORMES activity, which serve to define the scenario to be considered.

**Activity 2. Specifying requirements (scenario definition)**

In order to identify any requirements or user goals that the system should meet taking into account the context of use elicited in the previous activity, the approach proposed by Rosson and Carroll (2001), and adopted in TORMES methodology, was followed. In it, a problem scenario based on the elicited context of use that describes situations demanding affective support was firstly defined. After that, the scenario was rewritten and turned into the solution scenario, where the problems identified in the first scenario were tackled or avoided by delivering interactive recommendations (that is, recommendations that take advantage of the intelligence of the environment to interact with the learner).

The complete reporting of both scenarios (problem and solution) is included in the Annex. Here, we identify the relevant issues from them that have guided the requirements specification activity. In summary, George needs to prepare for his oral examination of English as a second language using a videoconferencing system. For that, he decides to talk for 5 minutes playing a given situation. Previous to talking, he takes 1 minute to prepare for it. The problem scenario has identified four self-blocking situations that may reflect affective changes: i) when preparing the talk, ii) once finished the introductory speech learnt by heart that he uses to get into the talk, iii) when he speaks haltingly and do not use time to think in advance what to say, and iv) when he is stuck and do not react for a period of time because he is nervous and gets blocked. In these situations, some information related to his affective state can be gathered, such as physiological information (i.e., pulse, skin conductance and temperature), his voice status or his facial and body movements.

The resulting solution scenario has identified two kinds of affective support managed in terms of context interaction feedback: 1) playing some relaxing music to calm George down, and 2) sending some signals (using a light and a buzzer) to recommend him to breathe deeply. The first one entails a direct adaptation in the environment, as the system controls the learner mobile device and plays a song for him. The second is driven through the light and the buzzer, which are meant to remind him a recommendation that consists in suggesting an action to do (which George can decide to follow or not). Here the purpose is to provide an affective support to the learner for a better learning experience, which is expected to turn into better learning outcomes.

To cover the following step of the elicitation process, next we describe the design solution for the recommendation “to breathe deeply” elicited in the scenario, that is, we describe how we have implemented the scenario.
**Activity 3. Producing design solutions (scenario implementation)**

This activity deals with modelling and instantiating recommendations. In particular, the modelling of the recommendation elicited in the previous activity was done using TORMES recommendation model. This model serves to identify the recommended action (what), the recommendation rules (when and who), the justification of the recommendation (why), the recommendation format (how and where) and the recommendation attributes (which). The resulting modelling of the recommendation aimed to make the learner breathe deeply is reported in Table 1.

To implemente this scenario, we have designed the Ambient Intelligence Context-aware Affective Recommender Platform (AICARP). Its goal is to provide the required ambient intelligence features. AICARP is shown in Figure 1, which follows a modular design controlled by an Arduino board with a microcontroller Atmel. This platform interacts with the computer that the learner uses for training oral exams (Oral Exam PC) through the communication module.

The sensors module receives information from the learner (physiological signals) and the environment (ambient information) through internal and/or external sensors. Thus, the internal sensors are connected directly to the Arduino board such as the e-Health sensors platform. The external sensors are connected to Arduino using a short-length wireless communication (e.g., Wi-Fi or Bluetooth). This configuration allows the communication with smartphone-based electrocardiograms (such as the one described in Torrado-Carvajal et al. (2012)) or wearable devices for quantified-self monitoring (such as the Samsung Gear smart watch). In any case, the sensor module sends the gathered information to the control module. This module in turn is in charge of distributing the information to the storage module and the recommendation module. The storage module can record the interaction data gathered through the sensors and feed the recommendation process by modelling the learner and her context. In this way, the recommendation module can analyze the information and take the decision to deliver the recommendation to the learner that fits with her current needs and context.

The recommendation module offers interactive personalized support to the learner using the sensory communication channels available (visual, audio or tactile). This module can send information or queries to the Oral Exam PC (for instance, recommendation messages or images to be shown on the screen). Moreover, it can deliver the recommendations in a more interactive way that does not interrupt the learning activity. This can be done by using different signals representing different recommendations, each of them suggesting the learner how to behave, and whose meaning the learner already knows. To do this, the recommendation module takes advantage of the intelligence to the environment using both internal and/or external actuators. Internal actuators (connected directly to Arduino) can be a display (LCD), a group of light emitting diodes (LEDs) with different colors, a buzzer that plays musical tones and vibrates slightly, a vibration element that moves so the user can sense it, etc. External actuators (connected via a wireless communication) can be a smartphone, a wearable device such as the Google glass, etc. The recommendation channel configuration functionality within the recommendation module is meant to account for users’ interaction preferences (including accessibility needs) as it allows selecting the most appropriate sensory channel to deliver the recommendation.

The evaluation of the recommendation design proposed in Table 1 for the given scenario implemented with the AICARP system is performed in the fourth and last TORMES activity.
Activity 4. Evaluating design against requirements (scenario evaluation)

The evaluation of the scenario was carried out applying the user centered design method called Wizard of Oz (Dahlbäck et al., 1993). This method is of use to clarify the logic behind a proposed design as it enables unimplemented technology to be evaluated by using a human to simulate the response of a system. In this sense, we have implemented a prototype of the AICARP infrastructure, which instantiates the recommendation modelled in Table 1. This includes the implementation of the electronics for the detection of the physiological signals that provide affective information (i.e., heart rate, pulse, skin conductance and skin temperature) as well as electronics to deliver the sensory oriented feedback to the learner through the appropriate actuators (in this case, a green LED that lights and a buzzer that plays a pure tone and vibrates slightly when the learner is recommended to breathe deeply). The detection of the participants’ facial and body movements as well as voice status would require a significant effort to be implemented in the system at this point, but can be readily performed by a human. Thus, in this experiment, we rather focused on supporting the rest of the infrastructure involved. For this reason, the detection of the user behavior (movements and voice) was done by the ‘wizard’ (from the Wizard of Oz aforementioned method) assigned to the experiment who was a psycho-educational expert with wide experience in supporting learners (with and without disabilities) both in online and face to face settings.

In order to decide when to deliver the recommendation to the participants when training for the oral exam, the wizard was provided with a live video in a Samsung Galaxy S5 smartphone of the electrocardiogram signals measured in real time with a previously developed wireless Arduino-based system (Torrado-Navajas et al., 2012). In addition, a script was developed to allow the e-Health signals corresponding to the pulse, skin temperature, skin resistance and skin conductance to be visualized in a computer using UC Matlab in real time. The actuators used to deliver the recommendations are internal (i.e., the green LED and the buzzer playing a pure tone) and were controlled by the wizard by pressing a button when the recommendation was to be delivered.

With the above infrastructure, the impact of the elicited interactive recommendation on the learner was evaluated at the end of the experiment by means of a questionnaire and an interview. The questionnaire was the System Usability Scale (SUS) (Brooke, 1996) adapted for the experience (i.e., particularizing questions to refer to the AICARP system). The SUS is a ten-item attitude Likert scale that offers a global view of subjective assessments of usability on a scale of 0–100. The larger the score is, the higher usability level measures. Bangor et al. (2009) have divided the numeric scale into 5 levels, being A the highest one. In our case, the goal was to measure the usability level (i.e., ease of use and learnability) experienced by the participant when using the technological environment developed in this research to provide the interactive recommendation (i.e., through the AICARP infrastructure). The interview consisted of the open questions compiled in Table 2. The goal of these questions is to understand the participants’ opinions of their interaction with the system regarding perception (i.e., what the learner felt, how was the learner impacted), intrusiveness (i.e., to what extend the learner felt uncomfortable with how the data was gathered and the feedback provided), and utility (i.e., if the provided recommendation was timely given and appropriate for the learners’ needs). Evaluation results are commented in the following section.
4. Experimental design and evaluation with users

As commented in Section 1, recommender systems in educational scenarios have not taken advantage of ambient intelligence to provide an interactive context-aware affective support to learners. For this reason, the experiment carried out in this research was designed as a proof of concept involving a small number of participants aimed to gather useful information to be used later in larger-scale experiments. As a consequence, and taking into account the small sample size, the objective was not to obtain significant statistical results. Here the goal was to explore in a pilot study the suitability of designing interactive recommendations in an educational context that takes advantage of ambient intelligence features (thus, going beyond the state of the art) and evaluate their usage with learners (thus, dealing with an educational setting).

To that end this research has followed a qualitative methodological approach, which consists of applying a flexible design widely used in the educational field (Olson, 1995) and following an observation protocol where different data sources are defined and registered. Hence, any event, change or variation that is considered relevant to the investigation has been gathered during the experiment. In this way, the analysis of the results of the pilot study is based on the data observed by the researchers on the learners’ interactions (behavioral and physiological changes), as well as the outcomes of the SUS questionnaire and the interview carried out with each participant.

In order to identify the main issues involved in the current research, data were grouped in descriptive categories that served as a preliminary analytical framework that will enable further research. Data gathered have been analyzed following inductive and inferential methods. Specifically, a chi-squared test (or $\chi^2$-distribution) has been used for testing the goodness of fit of the observed distribution and independence among different random variables classified according to several qualitative criteria. This test allows establishing relations between the effectiveness of the recommendation provided and how this recommendation could be related to different attributes such as system usability, learners’ affective state, as well as recommendation impact, suitability and format.

Next, we describe the experimental design followed in terms of its procedure, participants and task, as well as the outcomes from the evaluation with users of the interactive recommendation elicited for the scenario “preparing for the oral examination of a second language learning course”.

4.1. Experimental design

Before starting the experiment, each participant was asked to read and sign an informed consent to assure participants’ understanding of the facts and implications of their participation in the experiment. They were also explained in a very general manner (not to influence the experiment) the activity to be carried out in the pilot study. Demographic information was gathered, namely: gender, age, studies, occupation, illness related to heart and brain, physical activity and level of English. Participants were informed about how (i.e., through a LED and the associated buzzer) and what (i.e., breath slowly) they were to be recommended in order to face some situations that may come up during the learning activity. In addition, they were trained during a short period of time on how to deal with this situation by managing their breathing rhythm, which is meant to make them more relaxed and focused on the learning tasks. Before the activity started and after it ended, participants were asked to relax in order to measure the baseline for the physiological signals.
The experiment was carried out by 6 participants (4 male, 2 female; age ranging 22-24 years old), one of them being visually impaired (blind). The sample was homogeneously distributed in relation to their level of English (participant 2 and 3 had low level; participant 4 and 6 had medium level; participant 1 and 5 had high level). Participants’ personality traits were gathered with the following questionnaires: 1) the General Self-Efficacy Scale (GSE) (Schwarzer, 1993) to assess the self-beliefs of participants to cope with a variety of difficult demands in life, 2) the Big Five Inventory (BFI) (John et al., 1991) to reveal the main five structural dimensions of personality (extraversion, agreeableness, conscientiousness, neuroticism and openness), and 3) the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988) that consists of twenty adjectives describing different emotional states which the learner considers herself to be in after carrying out the activity.

Two tasks were given to the participants to be carried out individually. Each task consisted in speaking aloud for 5 minutes, while being recorded with the webcam, about one of the following topics: “Healthy living environments” and “Trips and foreign destinations”. They were delivered in random order to each participant. Previous to talking, the participant had 1 minute to think about what to say.

In each task, the psycho-educational expert (and wizard) was in charge of supporting the participant along the session by delivering the interactive recommendation (in terms of sensory oriented feedback) when needed with the criteria specified in Table 1, as follows. If the learner’ pulse increases in relation to the values obtained in the baseline before starting the task, the learner’s movements rate (in face and body) increases, the learner stops talking for several seconds or the learner’s speaking rate increases and impacts on the speech intelligibility thus resulting in an unstructured message, then the wizard presses the button to deliver the recommendation.

During the experience, participants’ interactions were registered by means of the AICARP system (physiological signals), the webcam (facial expressions and voice), and a video-camera (body movements). Time-stamped notes were also written down by an observer.

4.2. Results

The main results from the experiment carried out in the pilot study are summarized in Table 3. In particular, for each participant and for the two tasks, it shows the average heart rate (HR) in terms of heart beats per minute during the baseline measurement period at the beginning of the experiment as well as at the time when the interactive recommendation was delivered (RTD). It also shows the criteria (or conditions) that triggered the delivery of the recommendation. In addition, average values of the physiological signals gathered (heart beat per minute, skin temperature, skin resistance and skin conductance) along the different phases of the experiment (for each task: preparation, speech and end) as well as the observed status of the participant (relaxed/nervous) from the videos recorded are compiled in Table 4. In it can be seen that an increase in skin temperature (ranging from a few tenths to a Celsius degree) and an increase in skin conductance (corresponding to a decrease in skin resistance) were also detected. These variations have been noticed in most of the participants and the variability (increase and decrease) of obtained values is specifically related to the interval of time while the learners are carrying out the given tasks. This information is to be taken into account in future studies by applying data mining on these data to enrich the recommendations elicitation process (as considered in TORMES).
Data collected in Table 3 shows that the largest increases in heart rate occurred during the execution of Tasks 1 and 2, reaching in some cases a variation of 60-70% compared to the values obtained in the baseline. It also shows that most of the participants received more recommendations during Task 1 than during Tasks 2. This finding could be related to two different aspects. On the one hand, participants might have become familiarized with the task approach after carrying out Task 1. On the other hand, the training and the recommendations provided in Task 1 might have allowed them to detect and manage by their own in Task 2 those affective states characterized by anxiety or nervousness that could be negatively impacting on their performance, hence not requiring the delivery of so many recommendations. Further studies are needed to clarify the cause of this difference in the number of recommendations received per task.

The usability of the AICARP system was evaluated using the SUS scale, obtaining values from 52.5 to 75 and a majority of D levels according to Bangor et al. (2009) (see Table 5 for individual scores). If we considered Sauro’s (2011) threshold to asses a system as usable (i.e., 68), which was abstracted from a set of 500 studies, only one participant considered the technological environment as usable. This finding is not surprising given the degree of intrusiveness presented by the current prototype. Further discussion on this issue is included in the next section. In addition a chi-square test was applied in order to determine if there was independence between the usability of the system (Table 5, last column, ‘SUS level’) and the recommendation effectiveness perceived by the participant (Table 6, last column, ‘Recommendation effectiveness’). The value obtained was $\chi^2 = 6$ being slightly higher to the critical value of the distribution $\chi^2 (0.95, 2) = 5.99$. This might suggest that there is no independence between usability (i.e., the SUS level) and recommendation effectiveness. However this result cannot be considered conclusive given the size of the sample.

The delivery of the interactive recommendation was evaluated with the five open questions compiled in Table 2. The answers provided by participants are summarized in Table 6. This table shows the results related to questions Q1, Q2, Q3, Q4 and Q5. For each question we analyzed the response (yes or no) and the explanation given, which was coded with a categorical attribute, as follows. For Q1 the Relax column reflects to the level of relaxation of the participants. Q2 shows the influence (or impact) in the participant. Q3 refers to the suitability of the recommendation. Q4 is related to the proposal of alternative formats to offer the recommended message by the system. Q5 refers to the effectiveness of the recommendation in improving the performance while training for the oral exam. A detailed discussion on the qualitative information gathered from the participants’ responses is given next.

Regarding Q1, it confirms that speaking a foreign language in a public scenario can be considered a stressful situation in which the participant undergoes changes in their affective state that influence their relaxation. Thus, they involve substantial variations in physiological variables such as those discussed before and compiled in Table 4. The feedback provided by participants show that most of them felt nervous especially during Tasks 1 and 2 due to some issues such as lack of vocabulary, lack of opportunities to practice English with other people, difficulties to structure the speech and control their emotions in front of a real audience, etc.

With respect to Q2, most participants reported a positive impact on their performance. They also commented that sound and light associated with the recommendation let them be aware of their affective state (i.e., nervousness) and encouraged them to get self-control by breathing more slowly, thus allowing them focus on their speech. Notice here that the feedback provided by one of the participants remarked that the recommendation had impacted on his/her performance in a negative
way because the sound associated with the recommendation was unpleasant and it startled him/her (see also Q4). In addition, another participant commented that the recommendation should provide information about what the learner is doing wrong so it can really impact on the performance.

According to Q3, most participants reported that the recommendation was provided at the right moment in time and always matched those situations where they were more nervous and blocked because they did not know how to restart the speech.

In response to Q4, participants in general considered that the recommendation was provided in an appropriate format. Nevertheless, they also commented several issues that are to be taken into account to improve the system, and thus, are discussed in the next section: 1) the light should be more visible because when they are nervous, they listen to sounds but sometimes they do not notice the light, 2) the sound associated with the recommendation should be more pleasant, 3) alternative formats they suggested were text messages and images on the screen (as they are used to them). One of the participants did not like receiving the recommendation through two complementary sensorial channels.

Finally, in Q5 all participants reported that recommendations were effective in modifying or improving their performance (even the participant who valued low the recommendation impact, suitability and did not like the format used to deliver the recommendation).

In order to verify if Q1 (relax), Q2 (impact), Q3 (suitability) and Q4 (format) could have dependence with recommendation effectiveness (Q5), corresponding chi-square tests were applied. The values gathered for Q1, Q2, Q3 and Q4 were $\chi^2 = 6$ being slightly higher than the critical value of the distribution $\chi^2 (0.95,2)= 5.99$. These results might suggest that there is no independence observed between relax, impact, suitability and format of the recommendations and the recommendation effectiveness as perceived by the participants. These results are not conclusive given the sample size, but should be considered as factors that need to be further explored in future research due the impact that could have on the recommendation effectiveness.

Taking into account the information collected in the set of questions previously described as well as the results from the SUS questionnaire, we could conclude that the recommendation mechanism proposed in this research is able to detect a problematic affective situation and react according in a rather efficient way but it is lacking as far as usability is concerned.

5. Discussion on contributions, open issues and related work

The goal of the research reported in this paper is to explore the potential of ambient intelligence to provide sensory oriented feedback that improves the personalized support that can be provided in educational settings through recommender systems. While designing, implementing and formative evaluating for the first time in the literature an interactive context-aware affective recommendation in a pilot study, we have identified some methodological and practical issues that are to be considered in further studies to be carried out in large-scale settings. This progress is of special value since there are no previous instances of related research available and through these findings we hope to pave the way for further advances. Taking into account the small sample size of the experiment reported here, findings cannot be considered conclusive and should be revisited in some future work. In addition, from these early clarifications we are in a much better position to keep eliciting other recommendations of this kind. Thus, in this section we discuss them in relation to four main issues involved, namely: i)
how to deliver interactive recommendations, ii) when recommendations are to be provided, iii) what learners’ features seem to be of relevance, and iv) how to deal with social aspects.

5.1. How to deliver interactive recommendations

Issues involved in delivering interactive recommendations include the preferred sensory channels available, the format to display the recommendation, the support to understand the purpose of the recommendation and the intrusion level. In addition, as commented in section 4.2, it needs to be further explored if there are dependences between the recommendation effectiveness and the system usability, the learner effective state, the recommendation impact, the recommendation suitability and the recommendation format, among others.

5.1.1 Selection of channels to provide the recommendation

The state of the art reviewed in Section 2 shows that context based feedback is already being provided in non-educational scenarios mainly by using lights and/or vibrotactile stimuli. In the same line, we have built an open infrastructure (with Arduino) that detects affective information from physiological signals and can deliver interactive context-aware educational affective support in terms of recommendations through complementary sensory channels: i) sight (by lighting a LED), 2) ear (by playing a pure tone with a buzzer), and 3) touch (the buzzer also moves slightly the table when vibrating). This is provided in a complementary way to follow a universal design approach in the sense that all channels are used as different ways to communicate the same output (and thus, suggest the same recommendation), so the learner can select the most adequate channel (or channels) for her preferences and context. This is useful for all learners due to their inherent functional diversity, but especially for those with some accessibility needs.

For instance, whereas a non-visually impaired person can read information on the LCD or be aware of a change of the light from the LEDs, a blind person (or a person who is in a context that does not allow her to look at those devices, for instance, because she is driving) can define audio and touch as her preference channel to receive information. In this case, the configuration functionality of AICARP can set the buzzer (audio channel) and the vibration (touch channel) elements as available communication channels to this learner. In turn, for a learner with hearing loss, the system can be configured to show messages using sign languages through the Smartphone or the Google glass without using the resounding characteristics of the buzzer module. For learners with cognitive impairments, the text information on the display or wearable devices can be adapted to improve comprehension. Further, there exist related works in the literature that discuss ways to adapt the user interface to people with disabilities (e.g., Postolache et al., 2012; Subbu and Gnanaraj, 2012). If no redundant information wants to be received, as suggested by one of the participants, then the AICARP system has been designed so the learner can configure her unique preferred channel.

In a future improvement, the system will be able to learn from the learners’ choices, so it can automatically use the most appropriate channel for each learner in each situation.
5.1.2 Format for displaying the recommendation

From the evaluation outcomes follows that there are some open issues with respect to the format used to display the recommendation that worth being investigated.

Participants’ answers in the interview showed that the recommendation displayed by the system was hardly perceived through the sight channel since the current implementation consists in lighting a unique LED, and this was too small to be perceived from the learners’ location. Light has been the most commonly used channel to provide feedback in ambient intelligence non-educational scenarios and thereof other ways to interact with the learner through light using the visual sensory channel should be investigated. As to the factors that may impinge on this, there are also cultural aspects that have to be taken into account (Laganier and van der Pol, 2009).

In addition, vibro-tactile feedback was not noticed by any of the participants, thus, this format of delivering the recommendation also needs to be explored in more detail, going further than making the table swing slightly. As pointed in (Foottit et al., 2014), vibro-tactile feedback can be delivered in terms of varied body parts, including the torso, arms and head, and can be used for instance to restore users attention (Garcia et al., 2013), alert about high priority events when working in visually overloaded environments (Salminen et al., 2008) or to portray visual information to visual impaired users (Palmer et al., 2012).

It can also be of interest to consider the smell sense as a way to interact with the learner. Borromeo et al. (2010) demonstrate the possibility of managing this capability objectively and there are works that show their potential value for affect modelling in education scenarios (Garcia-Ruiz and Santana, 2013).

5.1.3 Recommendation purpose

Regarding the recommendation purpose, participants commented that they would like to be shown what they are doing wrong and/or what they should do instead. We propose a couple of options to address this.

On the one hand, a blended approach can be followed at the beginning of the learners’ interaction with the system. This would consist on a textual message that is shown (on the computer, the LCD display, etc.) at the same time that the corresponding signal (or signals) are delivered to the learner to offer a specific recommendation. In this way, the text can clarify the recommendation meaning, but would interrupt the learner activity, so it has to be applied with care. Delivering this textual message can be managed from the control module, so it is delivered only when the learner still is not aware of the recommendation meaning.

On the other hand, a more implicit way can be used to suggest the learner what to do instead of what she is currently doing (provided that she already knows the meaning of the recommendation when delivered through the corresponding sensory channel). This consists in making the recommendation interact in a way that can serve as a model of what the learner is recommended to do instead. In particular, for the recommendation elicited here, this proposal would mean that both the light and the vibrations swing in the same rhythm that the learner should breathe to get to a relaxed situation that facilitates the retaking of the speech. In this way, the learner can mirror the signal movements with her breathing.

5.1.4 Intrusion level

The evaluation showed that participants perceived the system rather intrusive since uncomfortable sensors were attached to them during the pilot study to collect the
physiological signals (i.e., electrodes were stuck in ankles, wrists and fingers). Thanks to the advance of current wearable devices, these variables can be registered in a much less intrusive way, such as smart watches (Mukhopadhyay, 2015), and even by means of clothes with sensing devices (Koo et al., 2014).

In addition, in the MAMIPEC project, aimed to integrate cognition with user’s emotions in order to provide inclusive adaptive learning (Santos et al., 2013), we are improving the detection capabilities of the system by considering new low intrusive signals (e.g., eye tracking on the participants’ gaze on the screen, and body movements on an intelligent chair). In particular, a universal design approach is followed to provide personalized affective recommendations according to the user’s functional diversity (that account for the learners’ accessibility preferences).

5.2 When recommendations are to be provided

The timing to deliver the recommendation is influenced by the detection of physiological and behavioral changes, and should avoid interfering with the learning task. As to when to deliver the recommendation, in our pilot study, the wizard (from the Wizard of Oz user centered design technique) was in charge of deciding the appropriate moment to offer it, and the participants found this appropriate. The challenge now is to make the system identify these situations automatically from the learning context, while considering both physiological and behavioral changes. Here, as commented in section 4.2, data mining techniques can be explored to automatically identify the criteria that characterize the appropriate moment to deliver the recommendation (Salmeron-Majadas et al., 2015).

5.2.1 Physiological changes

As first step, deciding when to recommend requires improving the detection approach by selecting the most adequate physiological signals to detect affective states changes. In this respect, the results compiled in Table 4 are similar to those reported in the literature (Novak et al., 2012) and show that there are physiological signals such as skin temperature and conductance/resistance that show significant variations when participants are carrying out the tasks.

As it can be seen from the results included in Table 4, heart beat, skin temperature, skin resistance and skin conductance show some relation among them. When the observation of the participants suggested that they were nervous, their heartbeat was higher, their temperature was a little higher than in a normal condition, and their skin conductance was inversely proportional to their skin resistance. All of these attributes were inversely altered when the user was relaxed. Therefore, considering the physiological changes detected, it can be argued that these changes could be related to different affective states experienced by participants in tasks that can cause higher levels of stress, anxiety, nervousness and thus, they could be used as indicators to guide the moment when the recommendation should be provided. However, further experiments are needed.

In addition, all this requires clarifications of how accurate the sensors can be considered in terms of capturing affective states and how the measures collected through them can contribute to the recommendation delivery process (i.e., do recommendations get perceptibly better?). This possibility needs to be researched in more depth by evaluating to what extent the system could accurately detect the information and act accordingly.
5.2.2 Behavioral changes

Behavioral changes from facial and body movements, as well as in the voice status can relate to positive and negative affective states (Schindler et al., 2008; Pantic and Rothkrantz, 2000; Gnjatovic and Rosner, 2010; Kleinsmith and Bianchi-Berthouze, 2013) influencing engagement and learning achievement (D’ Mello and Graesser, 2010). Data compiled in Table 3 agree with findings reported in the literature and show that when the user is carrying out Tasks 1 and 2, the body and facial movements rate is higher than during the preparation stage (when the participant is more relaxed). In addition, the speech speed seem is another criterion to deliver the recommendation. Considering this information, it could be said that these changes could be related to different affective states and thus, they could be used as indicators for guiding when the recommendations should be provided.

Having this evidence in mind, it is necessary to research in more depth how affective states recognition techniques based on facial and body movements detection as well as speech processing can contribute to the recommendation delivery process so as to allow the system to act in accordance with them.

5.2.3 Interference of the recommendation in the task

Another issue related to when to recommend is whether to interrupt or not the learner. Interruption management has been frequently addressed in human-computer interaction research. There have been efforts showing the effects of interruptions on web-based systems and they stress the necessity of properly managing them (Puerta et al., 2008; Bailey et al., 2008).

The recommendation delivery process can suppose a cost on the cognitive processes (i.e., attention, memory, etc.) involved in solving a task. Thereby, the incorporation of new information in the task may affect the cognitive control in a negative way on the learner performance. There is evidence that the brain is sensitive to external interferences such as disruptions, performing two activities at once, etc. The negative impact can be higher in those who have lower cognitive control such as children, elderly or people with attention deficit disorders.

In the current approach, recommendations can be delivered simultaneously with the learner activity, so learners may not be interrupted to follow the recommendation (e.g., while they keep talking, they are recommended to breathe deeply to calm down). The results compiled in Table 6 about “suitability of the recommendations” (Q3) show that most of the participants considered that this issue was adequate. However it is also possible that there were moments during those talks in which delivering recommendation turns into an interruption rather than an aid or just the learner might prefer not to receive recommendations. Taking into account these facts, different factors (such as type of task, learner motivation and preferences, type of recommendation, time to task completion, etc.) should be further considered in future research.

5.3. Learners’ features

Two types of learners’ features have been identified of potential relevance in order to design other recommendations. These are the English level and the personality traits.

5.3.1 English level
The outcomes from the pilot study suggest the need of considering different types of recommendations aimed at addressing the different English level of the learners. For the pilot, we designed the same recommendation to all the learners, but participants who had a medium and high level of English received more recommendations than those who had low level. A possible explanation to this (which has to be explored in further research) is that participants with low level of English might have difficulties to produce a fluent speech that did not relate to poor control of their emotional state (since the situation was not accompanied by a high pulse and/or body movement rates, etc.) but, probably, to other issues such as a lack of vocabulary and/or ability to produce sentences. If that is the case, participants with low level of English might need other kind of support, focused on providing a vocabulary related to the topic, facilitating hints about how to produce short sentences with the proper grammatical structure, etc.

5.3.2. Personality traits

As to personality traits, they allow to detect user characteristics such as level of emotional control, strategies used to manage affective states, etc. These can impact on learner’s performance and thus, can influence how the learner is facing a problem situation and what kind of strategies can be implemented to solve it. Considering that our the work reported here is a pilot study, data is very preliminary, but from its analysis it is suggested the need to revise further if there are some dependence between the personality traits, the level of English and the number of recommendations received. Thus, future works should research if personality traits need to be taken into account when designing the recommendation (and in that case, how that has to be done). For this research, previous studies that associate personality traits to study goals, such as (Zimmerman et al., 2006; Ludford and Terveen, 2003; Harrigan and Loffredo, 2010) will be taken into account.

5.4. Social aspects

According to the literature (Nikas et al, 2014), collaborative tasks are a determinant factor which influences not only individuals’ behavior towards the adoption of a collaboration technology, but also affects the effective use of such technologies from individuals.

Regarding social aspects, the scenario proposed in this experiment considers the involvement of two learners for practicing the conversation. Nevertheless, for now we have focused on the individual needs. We believe that new recommendation opportunities will appear when following the proposed scenario, George asks his friend to practice with him the conversational situations. Thus, factors that have not been considered so far should be analyzed. Specifically, issues related to interaction in a real context, where other issues are involved, such as intensity of collaboration, type of collaborative task, individual acceptance of the technology usage to support collaboration activities as well as specific personality traits such as extraversion, conscientiousness and neuroticism), which might be related to the level of participation, tolerance to stress or frustration, critics acceptance, etc.

As a result, future work should consider enriching the scenario with collaborative support, which opens new modelling issues from the affective and collaborative viewpoints.
6. Conclusions

In this paper we have reported the design, implementation and evaluation of the an infrastructure (so called AICARP) to explore the potential of delivering context-aware affective feedback that goes beyond computer-based recommendation approaches (i.e., those that consists in presenting some text or telling aloud the learner what to do) by taking advantage of the possibilities of ambient intelligence. As a result, the corresponding personalized support can be provided without interrupting the learning activity by delivering the recommended action to the learner at the same time she is carrying out the learning activity (e.g., while the learner is talking, the system can recommend her to slow down by switching on a light or playing a sound). This requires enriching the system with capabilities to detect changes in the learners’ affective state (for instance, from physiological sensors), as well as to interact with the learner through her preferred sensory channel (i.e., sight, hearing, touch, and even smell).

TORMES methodology complemented with the development of an Arduino based infrastructure (i.e., the AICARP) that can sense changes in the learners’ affective state and deliver recommendations in a more interactive way, has served to design and evaluate an interactive affective context-aware educational recommendation for the elicited scenario. In particular, this research opens a new avenue in related literature which focuses on managing the recommendation opportunities that an ambient intelligent scenario can provide to tackle affective issues during the language learning process when preparing for the oral examination of a second language learning course.

Up to our knowledge, based on a systematic review of 82 educational recommender systems (Drachsler et al, 2015), AICARP is the first recommender system that takes advantage of Arduino to gather contextual information and deliver interactive recommendations aimed to enrich the personalized support provided to learners in technology enhanced learning settings. Both the functional and human computer interaction aspects have been evaluated with ten users. In particular, a researcher to support the recommendations elicitation process, an educator with ambient intelligence expertise, a psycho-educational expert who defined the personalized affective support to be provided, a telecommunications engineer who implemented the system and six learners (i.e., the participants of the pilot study) that were learning English as a second language.

Results have shown that the recommendation elicited with TORMES is detecting a problematic affective situation and reacting accordingly in a rather efficient way since participants agreed that when received, it helped them retake the speech with a better disposition. However, as expected, they found the system quite intrusive regarding the affective information gathering since participants were attached uncomfortable sensors during the experiment to collect the physiological signals. This should not be considered a strong limitation since physiological variables can already be registered in a less intrusive way (as discussed in Section 5.1.4).

We are currently working on the automation of the feedback delivery in those situations that have been identified of interest with the Wizard of Oz user centred design method. In this way, we can run new studies where we can address some of the open issues identified in Section 5 regarding how personal, cultural, social and contextual factors may affect the interaction approach. Then, we will be focusing on another crucial open issue, which is to measure the actual impact of this approach on the learning, target by studying if the learning outcomes improve when interactive recommendations are delivered. This requires defining a user-centric evaluation framework for educational recommender systems in social ubiquitous networking.
environments which can be built from generic existing frameworks, such as those proposed by Pu et al. (2012) and Knijnenburg et al. (2012).

Acknowledgements

Authors would like to thank and acknowledge the implementations carried out by Raúl Uría Rivas to gather and visualize the physiological signals as well as to deliver the interactive recommendation. Part of this work focused on the signals acquisition was implemented in his Final Career Project at Universidad Rey Juan Carlos entitled “Diseño de una plataforma inalámbrica basada en hardware libre para el registro y procesamiento de señales fisiológicas” supervised by Susana Borromeo López during the school year 2013–2014 and the rest within the aDeNu Research Group at UNED. In addition, authors would like to thank Sergio Salmerón Majadas, also from the aDeNu research Group, for the support provided for the statistical analysis carried out on the collected data.

Disclosure statement

Authors declare no conflict of interest.

Funding

This work was supported by the Spanish Ministry of Economy and Competence [grant number TIN2011-29221-C03-01 (MAMIPEC: Multimodal approaches for Affective Modeling in Inclusive Personalized Educational scenarios in intelligent Contexts)] and [grant number TIN2014-59641-C2-2-P (BIG-AFF: Fusing multimodal Big Data to provide low-intrusive AFFective and cognitive support in learning contexts)].

References


Artificial Intelligence (IBERAMIA 2012). Lecture Notes in Computer Science, 7637, 672-681.


Annex

The problem and solution scenarios compiled in the second TORMES activity, and which focus on the requirements specifications, are detailed in this Annex. First, a general description of the scenario is presented. Next, the problem scenario is detailed. After that, the problem scenario is turned into the solution scenario by modifying the reported situations so that corresponding affective support is provided to the learner.

General description of the scenario.
The learner is placed in the following scenario:

“George is a 23 years old student enrolled at the Spanish Open University (UNED) to study Computer Science. He is also learning English as a second language and preparing for the CERT\(^3\) certification B2 with the UNED Distance Language University Centre\(^4\). He has been learning English for 13 years (since he started Primary School) but he still fears the oral expression part of the exam, that is, the part where he is given a situation and has to play it in English with another student for 10 minutes. All along the years he has been given advices on how to deal with that situation (which he also knows is very common among all students in the same situation). He has read some tips online\(^5\) that suggest him to breathe deeply before starting, prepare a generic introductory text to get confidence, and so on.

The oral exam is taking place in four weeks and will be done through a videoconferencing system. The system picks two learners randomly and gives them a situation to play. Each learner is at her own place and they do not know each other nor have a way to communicate between them. Each of them is given

---

\(^3\) [http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp](http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp)

\(^4\) [http://portal.uned.es/portal/page?_pageid=93,154330&_dad=portal&_schema=PORTAL](http://portal.uned.es/portal/page?_pageid=93,154330&_dad=portal&_schema=PORTAL) (in Spanish)

\(^5\) [http://oupeltglobalblog.com/2011/05/03/7-tips-for-helping-learners-minimize-anxiety-in-speaking/](http://oupeltglobalblog.com/2011/05/03/7-tips-for-helping-learners-minimize-anxiety-in-speaking/)
a specific role to play in the situation and has 2 minutes to individually think about how to address it. They can individually write some notes on their own computer that they can later see while playing the situation. When the 2 minutes are done, the videoconferencing system starts and each student sees each others’ face through the screen and can talk to play the situation. The conversation has to last 10 minutes and it is recorded to be later evaluated by the teacher.

In order to prepare for this, George has decided to practice on his own computer for half an hour every day recording with a webcam his talk (to simulate the usage of the videoconferencing system while rehearsing). He has compiled 50 typical situations that can be asked in the B2 certification level and is going to follow the same procedure. This procedure consists of picking one situation randomly, preparing it for 1 minute (e.g., typing some notes in the computer) and then (when an alarm rings) the system starts recording while he talks for 5 minutes (till the alarm rings again). When he feels that he has done some progress, he will ask a friend to practice with him a conversation. But for the moment, he prefers to do a monologue and on his own.

Today he picks the situation ‘Claiming in the airport for a luggage loss’. The situation is not difficult, and he has already practiced with it in class”.

**Problem scenario.**
The problem scenario reported the following situations:

“George has only 1 minute to think about what to say and that blocks him. His heart rate increases and starts moving his legs. He then touches his cheek with the finger, a typical gesture he does when he is nervous.

When the 1 minute available for preparation is over, the alarm rings and recording starts. At first George is quiet, but then he remembers his general introductory speech and speaks it aloud. His voice sounds confident.

But then he gets nervous again. His heart rate increases and starts moving his legs. His fingers are also sweating and his body temperature is higher. After 2 minutes, he remembers to try to breathe deeply and he gets a bit relaxed (the look of his face tells so). This helps him to start talking about the situation.

However, he talks very quickly using very basic vocabulary and expressions instead of others with higher level of English competence that he also knows.

Then he gets blocked again and decides to stop the recording after 4 minutes. Once more, he has not been able to produce a full speech with introduction, development and closing for the situation. He ends up very frustrated”.

**Solution scenario.**
The above reporting of problems was then rewritten to provide solutions for it, as follows (changes are bolded):

“George has only 1 minute to think about what to say and that blocks him. His heart rate increases and starts moving his legs. He then touches his cheek with the finger, a typical gesture he does when he is nervous. **The system detects these combined actions and considers that George is getting blocked and thus, needs some support. In particular, a relaxing music from George playlist stored in his mobile is played in low volume (the system already knows that the**
learner likes this kind of feedback). He then forgets about the time limitation and focuses on the task to be done. His heart rate and legs movements reduce.

When the 1 minute available for preparation is over, the alarm rings and the recording starts. At first George is quiet, but then he remembers his general introductory speech and speaks it aloud. His voice sounds confident.

But then he gets nervous again. His heart rate increases and starts moving his legs. His fingers are also sweating and his body temperature is higher. The system detects again these combined actions and considers that he has to make George to relax a bit by making him breathe deeply. For this, the system sends him some signals that George can perceive but do not interrupt his task. In particular, next to the table is a light that turns on and buzzes (reproducing a quiet melody at the same time that produces some vibrations) when the system wants George to breathe deeply (and George already knows that these signals are meant to recommend this). As soon as George notices these signals (either by the sight, the sound or by the movement of the table due to the buzzer vibrations) he decides to follow the advice given by the system (i.e., the enhanced recommendation supported by these signals) and starts to breathe deeply. In a few seconds, George gets more relaxed (the look of his face tells so). This helps him to start talking about the situation.

However, he talks very quickly using very basic vocabulary and expressions instead of others with higher level of English competence that he also knows. Using a semantic speech analyzer, the system is able to identify this situation and suggest the learner to speak slower so he has more time to think before speaking aloud and thus, he is able to select more appropriate expressions and vocabulary. For this, another light with buzz is used to help communicate the recommendation to the learner. In this case, both the light and the buzz swing at the speed the learner should be talking, so he can mirror that rhythm as the rhythm he should be breathe.

Thanks to the support received during the rehearsal, the learner does not get blocked and keeps on talking in the recording for the 5 minutes, till the second alarm rings. He has been able to produce a full speech with introduction, development and closing for the situation. He ends up very happy and with high self-esteem. Tomorrow, George will ask a friend to practice the situation together in a conversation.
<table>
<thead>
<tr>
<th><strong>Recommended action (what)</strong></th>
<th><strong>Breathe deeply</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recommendation rules (when and who)</strong></td>
<td>Learner’s hearth rate (pulse) increase over baseline  &lt;br&gt; Learner’s movements rate (in face and body) increases  &lt;br&gt; Learner stops talking for several seconds  &lt;br&gt; Learners’ speaking rate increases and impacts in the speech intelligibility</td>
</tr>
<tr>
<td><strong>Justification of the recommendation (why)</strong></td>
<td>In order to prepare the oral part of the language exam, the learner needs to relax and this can be done by breathing deeply</td>
</tr>
<tr>
<td><strong>Recommendation format (how and where)</strong></td>
<td>Activate a light (green LED) and a buzzer  &lt;br&gt; (both complementary sensing channels are used to cope with universal design)</td>
</tr>
<tr>
<td><strong>Recommendation attributes (which)</strong></td>
<td>Category: affective support  &lt;br&gt; Task: training second language speaking</td>
</tr>
</tbody>
</table>

Table 1. Recommendation modelling in terms of TORMES recommendation model
<table>
<thead>
<tr>
<th>ID</th>
<th>Questions for the participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Did you feel relaxed during the experience?</td>
</tr>
<tr>
<td>Q2</td>
<td>Do you consider that the recommendations provided during the experience had an impact on your performance?</td>
</tr>
<tr>
<td>Q3</td>
<td>Do you think that the recommendations had been provided at the right time?</td>
</tr>
<tr>
<td>Q4</td>
<td>Do you consider that the recommendation format is appropriate? Would you prefer any other alternative format?</td>
</tr>
<tr>
<td>Q5</td>
<td>Do you consider that the recommendation was effective in modifying or improving your performance?</td>
</tr>
</tbody>
</table>

Table 2. Open questions asked during the interview to evaluate the delivery of the recommendations
<table>
<thead>
<tr>
<th>Participant</th>
<th>Task 1</th>
<th>Task 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>baseline participant 1</strong></td>
<td>5.20 Learner is blocked</td>
<td>12.11 High rate of legs and hand movements and low speech speed</td>
</tr>
<tr>
<td></td>
<td>6.27 Increased rate of legs and hand movements, high speech speed and increased heart rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.00 High rate of legs and hand movements, low speech speed and increased heart rate</td>
<td></td>
</tr>
<tr>
<td><strong>baseline participant 2</strong></td>
<td>4.29 High rate of mouth and eyebrow movements, low speed speech and increased heart rate</td>
<td>n/a (Criteria to deliver the recommendation were not met in this task)</td>
</tr>
<tr>
<td><strong>baseline participant 3</strong></td>
<td>4.00 High rate of legs movements, low speech speed and increased heart rate</td>
<td>10.17 High rate of mouth and eyebrow movements, low speed speech and increased heart rate</td>
</tr>
<tr>
<td></td>
<td>4.52 High rate of legs and mouth movements, low speech speed and increased heart rate</td>
<td>11.24 High rate of legs and hand movements and increased heart rate</td>
</tr>
<tr>
<td><strong>baseline participant 4</strong></td>
<td>3.44 High rate of legs and hand movements and increased heart rate</td>
<td>11.40 High rate of body movements, low speech speed and increased heart rate</td>
</tr>
<tr>
<td></td>
<td>4.31 Increased heart rate and low speech speed</td>
<td>14.52 High rate of legs movements, low speech speed and increased heart rate</td>
</tr>
<tr>
<td></td>
<td>5.11 Increased heart rate and high speech speed</td>
<td>16.26 Increased heartbeats rate</td>
</tr>
<tr>
<td></td>
<td>5.42 Increased heart rate and high speech speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.45 Increased heart rate and low speech speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.36 Increased heart rate and low speech speed</td>
<td></td>
</tr>
<tr>
<td><strong>baseline participant 5</strong></td>
<td>3.54 Increased heart rate and low speech speed</td>
<td>11.37 Increased heart rate</td>
</tr>
<tr>
<td></td>
<td>4.34 Increased heart rate and low speech speed</td>
<td>13.17 Increased heart rate and low speech speed</td>
</tr>
<tr>
<td></td>
<td>6.51 Low speech speed</td>
<td>14.49 Increased heart rate and low speech speed</td>
</tr>
<tr>
<td></td>
<td>7.34 Low speech speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.56 Increased heart rate and low speech speed</td>
<td></td>
</tr>
<tr>
<td><strong>baseline participant 6</strong></td>
<td>4.33 Increased heart rate</td>
<td>5.51 Increased heart rate and low speech speed</td>
</tr>
<tr>
<td></td>
<td>5.06 Increased heart rate</td>
<td>7.50 Increased heart rate and low speech speed</td>
</tr>
<tr>
<td></td>
<td>5.51 Increased heart rate and low speech speed</td>
<td>7.50 Increased heart rate and low speech speed</td>
</tr>
<tr>
<td></td>
<td>7.50 Increased heart rate and low speech speed</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Participants’ interaction data (HR at baseline and at recommendations delivery) as well as criteria used to deliver the recommendation

*Legend: HR: Heart Rate; RTD: Recommendation Time Delivery*
<table>
<thead>
<tr>
<th></th>
<th>Heartbeat/minute</th>
<th>Skin temperature</th>
<th>Skin resistance</th>
<th>Skin conductance</th>
<th>Observed status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participant 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>93</td>
<td>37</td>
<td>292375</td>
<td>5.64</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>99</td>
<td>38</td>
<td>17000</td>
<td>7.5</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>94</td>
<td>37</td>
<td>20005</td>
<td>6</td>
<td>relaxed</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>95</td>
<td>37</td>
<td>220375</td>
<td>5.24</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>99</td>
<td>38</td>
<td>30000</td>
<td>6.03</td>
<td>nervous</td>
</tr>
<tr>
<td>End T2</td>
<td>98</td>
<td>37</td>
<td>32000</td>
<td>5.75</td>
<td>nervous</td>
</tr>
<tr>
<td><strong>Participant 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>86</td>
<td>36</td>
<td>17000</td>
<td>6.72</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>110</td>
<td>37</td>
<td>15000</td>
<td>7.3</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>87</td>
<td>36.5</td>
<td>24000</td>
<td>7</td>
<td>relaxed</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>99</td>
<td>36</td>
<td>23000</td>
<td>7.56</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>114</td>
<td>37</td>
<td>13000</td>
<td>8.18</td>
<td>nervous</td>
</tr>
<tr>
<td>End T2</td>
<td>99</td>
<td>36</td>
<td>27000</td>
<td>5</td>
<td>relaxed</td>
</tr>
<tr>
<td><strong>Participant 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>108</td>
<td>37</td>
<td>33500</td>
<td>8</td>
<td>nervous</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>121</td>
<td>37</td>
<td>44000</td>
<td>9</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>111</td>
<td>37</td>
<td>43000</td>
<td>8</td>
<td>nervous</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>109</td>
<td>38</td>
<td>43000</td>
<td>7.79</td>
<td>nervous</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>111</td>
<td>38</td>
<td>27000</td>
<td>7.8</td>
<td>nervous</td>
</tr>
<tr>
<td>End T2</td>
<td>103</td>
<td>37</td>
<td>27000</td>
<td>8.27</td>
<td>nervous</td>
</tr>
<tr>
<td><strong>Participant 4</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>67</td>
<td>32.73</td>
<td>102330</td>
<td>9.87</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>101</td>
<td>33</td>
<td>93804</td>
<td>10.77</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>86</td>
<td>32.71</td>
<td>92335</td>
<td>10.73</td>
<td>relaxed</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>77</td>
<td>32.67</td>
<td>90586</td>
<td>11.10</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>101</td>
<td>32.51</td>
<td>90133</td>
<td>11.14</td>
<td>nervous</td>
</tr>
<tr>
<td>End T2</td>
<td>91</td>
<td>32.54</td>
<td>89900</td>
<td>11.14</td>
<td>nervous</td>
</tr>
<tr>
<td><strong>Participant 5</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>100</td>
<td>32.6</td>
<td>104389</td>
<td>9.70</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>120</td>
<td>32.71</td>
<td>95793</td>
<td>10.24</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>111</td>
<td>32.24</td>
<td>103471</td>
<td>9.77</td>
<td>nervous</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>100</td>
<td>32.67</td>
<td>107930</td>
<td>9.27</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>98</td>
<td>32.79</td>
<td>104609</td>
<td>9.95</td>
<td>relaxed</td>
</tr>
<tr>
<td>End T2</td>
<td>94</td>
<td>32.67</td>
<td>105690</td>
<td>9.75</td>
<td>relaxed</td>
</tr>
<tr>
<td><strong>Participant 6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preparation T1</td>
<td>85</td>
<td>33.59</td>
<td>211577</td>
<td>4.81</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T1</td>
<td>105</td>
<td>33.73</td>
<td>206479</td>
<td>4.91</td>
<td>nervous</td>
</tr>
<tr>
<td>End T1</td>
<td>96</td>
<td>33.8</td>
<td>197520</td>
<td>5.06</td>
<td>relaxed</td>
</tr>
<tr>
<td>Preparation T2</td>
<td>90</td>
<td>33.95</td>
<td>178314</td>
<td>5.66</td>
<td>relaxed</td>
</tr>
<tr>
<td>Speech in T2</td>
<td>102</td>
<td>34.07</td>
<td>186612</td>
<td>5.45</td>
<td>nervous</td>
</tr>
<tr>
<td>End T2</td>
<td>82</td>
<td>34</td>
<td>193779</td>
<td>5.26</td>
<td>relaxed</td>
</tr>
</tbody>
</table>
Table 4. Average values of physiological signals in each of the experiment phases
<table>
<thead>
<tr>
<th>Participant</th>
<th>SUS score</th>
<th>usable (&gt;68)</th>
<th>SUS level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>75</td>
<td>yes</td>
<td>C</td>
</tr>
<tr>
<td>Participant 2</td>
<td>67.5</td>
<td>no</td>
<td>D</td>
</tr>
<tr>
<td>Participant 3</td>
<td>52.5</td>
<td>no</td>
<td>F</td>
</tr>
<tr>
<td>Participant 4</td>
<td>62.5</td>
<td>no</td>
<td>D</td>
</tr>
<tr>
<td>Participant 5</td>
<td>65</td>
<td>no</td>
<td>D</td>
</tr>
<tr>
<td>Participant 6</td>
<td>60</td>
<td>no</td>
<td>D</td>
</tr>
</tbody>
</table>

Table 5. SUS scores per participants’ responses
<table>
<thead>
<tr>
<th>Participant</th>
<th>Q1 Relax</th>
<th>Q2 Impact</th>
<th>Q3 Suitability</th>
<th>Q4 Alternative format</th>
<th>Q5 Recommendation effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No</td>
<td>Low</td>
<td>Low</td>
<td>Sometimes</td>
<td>Low</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>Low</td>
<td>High</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>Low</td>
<td>High</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>High</td>
<td>High</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>Low</td>
<td>High</td>
<td>Yes</td>
<td>High</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>Low</td>
<td>High</td>
<td>Yes</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 6. Recommendation evaluation
Fig. 1: Ambient Intelligence Context-aware Affective Recommender Platform (AICARP)