

Engaging Adolescents with Down Syndrome in an Educational Video Game

César González-Ferreras¹, David Escudero-Mancebo¹, Mario Corrales-Astorgano¹,
Lourdes Aguilar-Cuevas², Valle Flores-Lucas³

¹*Departamento de Informática, Universidad de Valladolid, Valladolid, Spain*

²*Departamento de Filología Española, Universitat Autònoma de Barcelona, Cerdanyola del
Vallès, Barcelona, Spain*

³*Departamento de Psicología, Universidad de Valladolid, Valladolid, Spain*

Corresponding author:

César González-Ferreras. Email: cesargf@infor.uva.es.

The Version of Record of this manuscript has been published and is available in **International Journal of Human-Computer Interaction**, 11 Jan 2017, <http://www.tandfonline.com/10.1080/10447318.2017.1278895>

Engaging Adolescents with Down Syndrome in an Educational Video Game

This article describes the design, implementation and evaluation of an educational video game that helps individuals with Down syndrome to improve their speech skills, specifically those related to prosody. Special attention has been paid to the design of the user interface, taking into account the cognitive, learning and attentional limitations of people with Down syndrome. The learning content is conveyed by activities of production and perception of prosodic phenomena, aimed at increasing their communicative competence. These activities are introduced within the narrative of a video game, so that the players do not conceive the tool as a mere succession of learning activities, but so that they learn and improve their speech while playing. The evaluation strategy that has been followed involves real users and combines different evaluation activities. Results show a high level of acceptance by participants and also by professionals, speech therapists and special education teachers.

Keywords: Down syndrome, prosody learning, computer games

1 Introduction

Children and adolescents with Down syndrome (DS) are digital natives that use the information and communications technologies (ICT) with relative ease. They enjoy interacting with computers and mobile devices, browsing the web and playing with video games, just like any other people (Feng, Lazar, Kumin, & Ozok, 2010). There have been experiences that seek to exploit this fact to test ICT educational tools with this group of users (Ng, Bakri, & Rahman, 2014). Although in many cases they can use the same tools as the rest of the users without any kind of adaptation, people with DS present characteristics that determine the effectiveness

with which they can use the ICT educational tools (Pazos González, Raposo-Rivas, & Martínez-Figueira, 2015). Among the characteristics that limit the effectiveness of the ICT educational tools is the fact that many people with DS present attention deficit (Hernández Martínez, Pastor Duran, & Navarro Navarro, 2011), low tolerance to frustration, delayed language development (Chapman, Schwartz, & Bird, 1991), difficulty to process information that comes from different channels simultaneously (Lanfranchi, Cornoldi, Vianello, & Conners, 2004), lack of motivation (Wuang, Chiang, Su, & Wang, 2011), problems with the short term memory (Chapman & Hesketh, 2001), and difficulties in understanding the meaning of iconic symbols (Yussof & Zaman, 2011). All these limitations are an important challenge, taking into account the fact that our objective was to build a tool for speech training, thus requiring the development of a multimodal interaction system with both audio and visual input and output.

There is a well developed literature on the potential of games to improve motivation and engagement in education (McFarlane, Sparrowhawk, & Heald, 2002), but limited information on the advantages of using them to improve the speech production and prosodic skills of people with Down syndrome (Kent & Vorperian, 2013). Although some tools do exist (Saz et al., 2009), they are not effectively used by people with DS, since a high degree of motivation is required, which is not easily achievable by this type of users.

In this article, we describe the design and evaluation of an educational video game for speech training that has been designed taking into account the specific limitations of people with DS, in order to increase their chances of success in the formative tasks. A combination of design decisions that affect human-computer interaction (HCI) allowed us to achieve an increase in user motivation, and thus, an improvement of the effectiveness of the training sessions. The main idea is to apply typical elements of computer games, such as feedback and guidance (Kapp, 2012), in the development of an educational video game for speech training for people with DS called “The Magic Stone”.

“The Magic Stone” is a serious video game grounded in the genre of graphic adventure video games, in which the player assumes the role of the main character in an interactive story driven by exploration and problem-solving. Training activities are inserted into the video game and users must perform the activities in order to continue advancing in the adventure. Moreover,

some activities force players to interact with game characters using their voice, which helps them to feel included in the story. The use of a narrative where the learning activities are included is what makes the game immersive. The idea is that the users feel like they are in the world that the game creates, so they do the learning activities in an unnoticed way. Besides this, well established learning objectives concerning prosodic skills allow the progress of the users to be monitored.

The evaluation strategy of the video game was developed to answer three main research questions (RQ):

- RQ1: Do players with DS interact differently with the video game than people without intellectual disabilities? In spite of the fact that children and adolescents with DS are digital natives who are used to interacting with computers, we assess whether there are usability aspects affected by their special characteristics that condition the training performance.
- RQ2: Does the video game engage players with DS and keep them training? We evaluate whether the HCI elements included in the video game sufficiently motivate the users to keep on doing the training activities.
- RQ3: Do players with DS improve their communication skills using the video game? We evaluate whether the activities of the video game are effective from the point of view of speech training.

The results of the pilot test show that the special characteristics of users with DS cause a different use of the software compared to users without DS. At the same time, the conducted tests highlight the fact that the abundance and type of reinforcements used allow effective training to be created for pronunciation improvement. The results permit us to discuss how the HCI techniques employed have allowed us to address the users' attention deficit, memory difficulties and low tolerance to frustration. We also discuss how the concurrent interaction between teacher and the student while using the program is a key factor in obtaining a good training performance. The system was also tested by people without intellectual disabilities, with the aim of verifying if players with DS interact differently with the video game than people without intellectual disabilities.

The structure of the article is as follows. Section 2 reviews related works of the state of the art. Section 3 describes the features of the video game. Section 4 explains in detail the activities that must be done by the users. Section 5 presents the evaluation strategy and Section 6 shows the results. We discuss in Section 7 how the HCI elements introduced in the video game help to increase the users' motivation. Finally, the conclusions are presented in Section 8.

2 Background and Related Work

Several sources of knowledge are needed to design an educational tool addressed to people affected by DS. It is first of all necessary to know the characteristics of the Down syndrome, since their difficulties should be considered in the development of the tool (Section 2.1). Second, since the video game focuses on prosodic skills, an overview of prosody is also necessary, as well as the existence of tools that help in the learning of prosody (Section 2.2). Third, specific software and video games for training people with DS has to be reviewed in order to know which kind of solutions have been adopted in the state of the art (Section 2.3).

2.1 Down Syndrome and Language Difficulties

Down syndrome, according to the new diagnostic classification system DSM-5 (American Psychiatric Association, 2013), is a subtype of intellectual developmental disorder (IDD). IDD is characterized by significant limitations both in intellectual functioning (with an intelligence quotient -IQ- score equal to or below 70, with cognitive difficulties) and in adaptive behavior, showing difficulties in conceptual, social and practical adaptive skills. Besides, this condition appears before age 18 (Schalock et al., 2010).

Down syndrome is not the most common cause of IDD, but it is the most frequent of genetic origin. It is not a hereditary disorder and occurs equally in all ethnic groups, cultures and social classes. Although a clear cause is unknown, some risk factors have been established, among which the most widely agreed is the age of the mother.

People with DS present profiles of development and difficulties which are different in terms of different variables, such as IQ score, severity, etc. However, they generally present impor-

tant difficulties in their language skills, with better language comprehension than production (speech). Among their most common language difficulties are: articulation, in some cases their speech is almost unintelligible; phonological discrimination; more severe difficulties with the acquisition of morphological and syntactic aspects of the language, e.g. understanding and producing some complex morphological marks; and a very poor comprehension and production of syntax, so they do not understand or produce long and syntactically complex sentences, such as compound sentences. Although they have higher levels of lexical-semantic and pragmatic development, they also have limitations in these components: they have less vocabulary and problems to understand complex semantic relations. Even though they have good general conversational skills, they have difficulties with the use of mechanisms of discourse cohesion (anaphora, ellipsis, distinction between information already given and new information, etc.), difficulties to understand and produce indirect speech acts and figurative language, etc. (Flores Lucas & Belinchón Carmona, 2010; Kent & Vorperian, 2013; G. E. Martin, Klusek, Estigarribia, & Roberts, 2009). Individuals with DS also have limitations in the perception and production of prosody (Pettinato & Verhoeven, 2009; Stojanovik, 2011). Chapman (1997) concludes that interventions including language therapy targeting the practice of expressive language can increase communicative effectiveness.

Regarding their cognitive difficulties, perhaps one of the most severe and which compromises the following of the video game, is their memory difficulties, specifically, their difficulties to retain and process auditory and serial information (Chapman & Hesketh, 2001; Jarrold & Baddeley, 2001). They also have deficits in the integration of information, as well as in the processes of reasoning and deduction, thus making it necessary to provide more explicit information to alleviate their problems in making inferences. They also show difficulties with executive control functions (Daunhauer et al., 2014), thus, among others, they have deficits with the inhibitory response control, that is, they issue a response impulsively without having analyzed all the information previously or taking time to stop giving a wrong answer.

2.2 Training Speech Prosody with ICT

When people talk or read, not only sequences of phonetic segments are articulated, but also variations of the intensity, the speaking rate and the tone. In linguistics, all this extra information is referred to as prosody, which is relevant because it conveys a huge amount of information. Prosody is used in speech with a twofold purpose: 1) to indicate the segmentation of the discourse and its organization for the benefit of the interpretation of the messages and 2) to apply different intonation and stress patterns to convey the intended meanings, going from modality differences to nuances of emotions. Moreover, the information that is transmitted by prosody cannot be replaced by lexical or grammatical elements (Martínez Celdrán & Fernández Planas, 2007).

Prosody has been managed by computers for a long time, mainly due to the contribution of prosody to improving the naturalness of TTS systems (Escudero-Mancebo, González-Ferreras, & Cardeñoso-Payo, 2002; Taylor, 2009). There are algorithms that permit the acoustic signal of speech to be processed so as to code the relevant information encoded by prosody (Aguilar, Bonafonte, Campillo, & Escudero, 2009; Escudero, Cardeñoso, & Bonafonte, 2002; González-Ferreras, Escudero-Mancebo, Vivaracho-Pascual, & Cardeñoso-Payo, 2012). These types of algorithm are used by several tools, like the ones presented by Saz et al. (2009) for training speech and prosody. Video games are especially useful for studying and practising prosody with students affected by DS as they allow the incorporation of audio files to capture the differences between pronunciations (in order to work in the perception domain) and to record their own productions (in order to work in the production domain). Perception activities focus on the discrimination between meanings that cannot be attributed to lexical or grammatical elements (for instance, differences in the expression of emotions), while production activities focus on acquiring utterances pronounced by the student until a correspondence with the model is achieved.

There are tools that help in the learning of prosody. For instance, a software that augments text with visual prosodic cues to improve expressive reading (Patel, Kember, & Natale, 2014). The idea is to provide cues of the underlying prosody: pitch, duration and intensity. These authors carried out an experiment with children whose results suggest that beginning readers

could benefit from explicit visual prosodic cues and thus enhance oral reading expressiveness.

Other tools incorporate speech technologies to assist speech therapy with patients that have a speech disorder. For example, Shahin et al. (2015) describe a system for automated speech therapy in childhood apraxia of speech (CAS). The system is able to identify the three main types of error commonly associated with CAS: groping errors (delay in sound production), articulation errors (incorrect pronunciation of phones) and prosodic errors (inconsistent lexical stress). Rodríguez, Saz, and Lleida (2012) present another computer program for providing speech therapy to individuals with speech disorders. The tool is used to train such speech skills as voice production, intensity, blow, vocal onset, phonation time, tone and vocalic articulation. An experimental study of 27 subjects, the majority with cognitive disabilities, showed improvements in the voice capabilities of a remarkable number of users. Also interesting was the finding that the use of the program motivated the subjects, because they were attracted by the game-like applications. The tool consists of a set of game-like applications that are independent from each other. The main difference with our work is that “The Magic Stone” is designed as a graphic adventure game, and thus, all the activities are integrated into a story. This allows the different linguistic activities to be contextualized and increases the player’s sense of immersion in the video game.

2.3 ICT and Video Games for Training People with Down Syndrome

Children with DS are able to use computers in their daily lives to perform many tasks. A study about how children and young adults with DS use computers (Feng, Lazar, Kumin, & Ozok, 2008) found that 83% of the 561 participants started using computers at six years old. The majority of the children use the computer for learning (80%) and entertainment (95%). Educational software, video games and programs for the Internet were the most used applications. They spent an average of 3.5 hours a week using the computer at school and 4.94 hours a week using the computer at home.

The potential of games to improve motivation and engagement in education has been examined (McFarlane et al., 2002). The idea is to apply typical elements of computer games, such as feedback, guidance, time pressure, and rewards (Kapp, 2012). Moreover, these elements can be

included in non-gaming systems to improve user experience and engagement (Darejeh & Salim, 2016; Deterding, Sicart, Nacke, O'Hara, & Dixon, 2011). Games allow learners the chance to practice in environments close to real world situations. Rather than applying learning in a vacuum, games provide a context for learning, allowing players to apply what they have learned to solve real life problems (Prensky, 2007). Modern theories of effective learning suggest that learning is most effective when it is active, experiential, situated, problem-based and provides immediate feedback (Boyle, Connolly, & Hainey, 2011). Video games can include activities which have these features.

Graphic adventures are a video game genre that use graphics to convey the environment in which a story is developed during the game. Learners remember information more easily when they learn that information in the form of a story rather than from a bulleted list (Adams, 2013). Stories evoke emotions and provide a context for placing information. Involving a learner in a story can make the learning more powerful. A well-crafted story focuses on helping learners to solve problems, educates the learners, and is easily recalled when the actual situation arises or when a learner is in a similar situation.

In addition, there are some studies that show the efficiency of ICT and video games in the cognitive rehabilitation and teaching of people with intellectual disability: improvement in choice reaction time (Standen, Anderton, Karsandas, Battersby, & Brown, 2009), stimulating cognitive abilities of children (Brandão et al., 2010), independent decision making (Standen, Rees, & Brown, 2009) and improvement in mathematical skills (Brown, Ley, Evett, & Standen, 2011; Bruno et al., 2003; González, Noda, Bruno, Moreno, & Muñoz, 2015; Ortega-Tudela & Gómez-Ariza, 2006; Shafie et al., 2013; Wan Ahmad, Muddin, & Shafie, 2014). The program *Phonics Alive!* (Phonics Alive!, 1996) was used in (Gallaher, van Kraayenoord, Jobling, & Moni, 2002) to teach reading to a young adult with DS. Tangible interfaces are used in (Santana & Muro, 2011) to implement the reading method described in (Troncoso & Del Cerro, 1999), which is specific for children with DS. Torrente, Del Blanco, Moreno-Ger, and Fernández-Manjón (2012) developed two games whose main aim is to promote independent living for people with intellectual disability, such as interacting with others at work or the use of public transportation. A similar idea is presented in (Lopez-Basterretxea, Mendez-Zorrilla, Garcia-

Zapirain, Madariaga-Ortuzar, & Lazcano-Quintana, 2014), where players learn how to go shopping. Furthermore, other games are focused on therapy for children with speech disorders (Cagatay, Ege, Tokdemir, & Cagiltay, 2012). Reviews of educational software for people with DS can be found in (Black, 2006; Ng et al., 2014; Pazos González et al., 2015).

On the other hand, the evaluation of the user interface is important to guarantee the usability of the educational tool. However, there is little research in human-computer interaction literature related to people with DS. Lazar, Kumin, and Feng (2011) reported an ethnographic observation that examined computer skills of expert users with DS and Kumin, Lazar, Feng, Wentz, and Ekedebe (2012) carried out a usability evaluation of multi-touch tablet devices by adults with DS. Both studies report the great diversity in cognition, communication, skills, capabilities and computer experience of people with DS.

The direct involvement of people with DS during the design process is very challenging, due to their communication, cognitive and behavioral difficulties. One solution would be for designers to work with parents and carers as proxies for the user. Designers could also observe the interaction of the users with prototypes. There are some examples of this approach in the literature. A system that helps users with DS to learn addition and subtraction is described in (González et al., 2015). Different types of experts were involved in the design process (teachers, experts in usability, and experts in Down syndrome). Several prototypes were evaluated by the experts and redesigned with the proposed modifications. The final system was evaluated both by experts and by users with DS. Carmien et al. (2005) presented a system for supporting people with cognitive disabilities using public transportation. The interaction of people with cognitive disabilities was observed and interviews with specialists who help them were carried out. In the testing stages, there was a direct involvement of the users.

3 System Description

“The Magic Stone” is a video game whose main aim is to help people with Down syndrome to improve communication skills that have been affected due to their disability, especially those related with prosody. To do this, players will have to do some activities related with prosody and others not related with language, such as puzzle-solving, that have been introduced to add

variety to the game. Another aim of the video game is that it can be used by teachers or speech therapists in special education centers as part of their educational activities.

The video game has the structure of a graphic adventure game, including conversations with characters, getting and using items and navigating through different scenes. Players have to use the mouse to interact with the elements of the game. Players go through different scenes where they have to do some actions, like doing an activity or using an item. The main innovation is that players have to record sentences using their voice in some activities. The speech therapist or teacher decides if the player has done it correctly or not. In other activities, an audio is played in the context of a conversation and the player has to choose between some options to continue this conversation.

At the beginning of the game, the configuration screen is shown, where players can introduce a user name and configure their game profile. They can select an avatar, which represents the player in the game, the difficulty level and reader profile (reader/non reader). Then, a video is played to introduce the story of the game.

The plot of the adventure is a mysterious story in which the player takes the role of a hero who is responsible for saving the city from environmental destruction. To do this, the player has to retrieve the magic stone which has been stolen and which will restore order in the city. The search for the magic stone requires the user to carry out several activities in order to progress in the game. Four activity types are defined to practice speech, communication and prosodic skills. The activities will be described in Section 4.

The development of the video game is based on multimedia learning, presenting visual and verbal materials together simultaneously (Mayer, 2002). In this way, the users do not depend only on the textual channel to receive important information, as it can be completed by information with images, because this information modality has lower difficulties for people with intellectual disabilities (Chapman & Hesketh, 2001). In fact, it has been shown that using images to support and complement verbal information is a better educational strategy than using only verbal information (Buckley & Sacks, 2002). Other studies using multimedia learning systems have shown that it is an effective method for this population (Khan, 2010).

3.1 Learning Objectives

The learning objectives of the video game are to:

- Perceive and discriminate the different sentence types: declarative, interrogative and exclamatory.
- Identify the correct sentence type (declarative, interrogative or exclamatory) to use in a particular communicative exchange.
- Associate the different sentence types with the corresponding prosodic patterns.
- Produce the appropriate prosodic pattern according to the type of sentence.
- Produce sentences keeping the rhythm of the utterance, respecting the location and duration of pauses and the form of intonation.
- Control the volume and intensity.

3.2 System Architecture

Figure 1 shows the system architecture. Two users interact with the system: the player and the trainer. The player is normally a person with language deficits, specifically in prosodic comprehension and production. The trainer is typically a helper (teacher, speech therapist, family) that helps the player during game sessions. When trainer and player are working together on a game activity, the trainer will help the player in the correct use of voice and also to configure the system. The trainer evaluates the player's recordings in production and prosodic activities, making the player repeat the exercise when the result is not correct. To evaluate the player's recordings, the trainer uses the keyboard of the same computer that the player is using. Therefore, the trainer has to sit next to the player. The role of the trainer is essential to maximize the educational potential of the game. The trainer supports and guides players during the game, adapts the difficulty level, encourages players to continue when they have difficulties and helps them to solve such difficulties as understanding the story and the activities.

[Figure 1 about here.]

During the game session, information about user interaction is stored, as well as the audio recordings of the production activities. This information can be used by the speech therapist to analyze the evolution of the user in successive game sessions.

The application has a multimodal interface, as much for input as for output. Input is performed using voice and mouse. On voice training exercises, players will have to use their voice. Output is performed using visual and sound channels. The sound output channel is used for the voiceover that narrates the story, to play character voices and to play the voice of the virtual assistant. In any case, a recorded voice is used instead of TTS, as it allows emotions to be included that synthetic voices cannot yet express.

3.3 User Profile

The video game encourages the improvement of skills in a gradual way. Moreover, due to the variety of users that Down syndrome includes and their respective cognitive abilities, it is necessary for the video game to be able to adapt to the particular needs of each person. Therefore, the video game includes a number of levels and reinforcements adapted to the individuality of each player.

The video game allows the trainer to define the user profile of the player. First, the trainer chooses between “reader” and “non reader”, as the reading difficulty is an important aspect. The activities have more visual and audio instructions for non reader users. Second, the trainer selects the difficulty level (beginner, intermediate, advanced) which affects the complexity of the activities and the instructions that players receive.

To improve the feeling of immersion within the game, players can choose a personal avatar at the beginning of the game. It represents the players’ image in the game, allowing players to identify themselves with the character of the story. This is important because promoting the presence of the player in the game is necessary for the completion of the maximum number of activities.

3.4 Virtual Assistant

A virtual assistant is used to guide the player during the game. This assistant is a parrot, as shown in Figure 2. This assistant tells players to which place they must go or which object they must find or use. It also reminds players of the current goal.

[Figure 2 about here.]

3.5 Interface Design

The development of the scenes, items and characters has a uniform design, close to cartoons, but without making them too childish. Bright colors are used in accordance with the scenes represented in the game. A simple text font is used, with a larger size than usual to make it easier to read (Arial Rounded MT Bold, size 19pt).

The instructions of the activities are formulated in a way easy to understand for the users. We use simple sentences and high frequency words, so they can understand all the words used. An expert on developmental and language disorders and on intellectual disability revised the sentences to guarantee that they are in accordance with the cognitive level of people with DS. Furthermore, in all the activities, they are provided with visual clues to facilitate the task (like the traffic lights in the prosodic production tasks or the microphone in the other speech tasks). In addition, the provided feedback is both visual and auditory.

3.6 Narrative and Game Immersion

The scenes, characters and items that the players find in the game are representations of real world elements. The objective of this is that players can identify daily situations and be able to transfer the game lessons to real world situations. For instance, some scenes are the player's home, a library, a bus and a bookshop. The player will meet different characters, such as the bus driver, the librarian or the bookshop assistant. Some items that the player will find are: a wallet, a book, a magnifying glass and a map. The story also includes some imaginary elements to motivate the player.

3.7 Feedback

Each activity offers the users feedback according to the results obtained. However, due to the difficulties presented by the target users, it is important not to cause frustration that can produce an abandonment of the game. For this reason, errors are dealt with in a positive way. Thus, they are always allowed to progress regardless of the results. A negative feedback is shown when they go wrong, to show them that they committed an error. However, this feedback is complemented by a positive message that helps them to keep playing and not get demoralized, showing them that making mistakes is totally normal. Moreover, in order to not stress the users, we avoided the use of scores of any kind. For the same reason, the available time is not shown to the users.

3.8 Learner Model

An important feature of educational tools is the possibility of monitoring the learning of the users. For this reason, the video game stores the interaction history of each player. Then, the reporting module analyses the sequence of operations performed in a session and provides useful information to the instructors in order to see the evolution of the student's skills along the different sessions of the game. The following information is provided:

- Total time taken to finish the game.
- All the recorded audios of the players.
- For each activity:
 - Time spent to complete the activity.
 - Number of errors.
 - Number of additional listenings of the phrase.
 - Number of speech aids.
 - Number of incorrect clicks.
 - Number of phrase repetitions.

4 Activities

This section describes the activities included in the video game. These activities are included in the general context of the game and players need to do them in order to progress in the game. Four activity types are defined to practice speech, communication and prosodic skills. All the activities have been planned according to the principles of learning that have proven most effective for teaching and presenting information to people with DS (Buckley & Bird, 2000, 2001). Furthermore, we use visual clues to support the verbal information, because this type of combination has proven to be the best way for people with DS.

4.1 Comprehension Activities

These activities are focused on lexical-semantic comprehension and the improvement of prosodic perception in a specific context, such as asking a question or asking for something politely.

[Figure 3 about here.]

These activities introduce players to different conversations with game characters. They have to choose between a series of sentences to continue the conversation, of which only one is correct in the context (as shown in Figure 3). These sentences are played at the beginning of the activity to complement the textual information with audios, so users can get the information from two channels to prevent possible reading problems. Players have the option of replaying the sentences again. If players choose the correct sentence, the activity ends, showing a positive feedback. If players choose the wrong one, the activity plays an audio and they can choose again. If, after a given number of attempts (4 or 5 attempts, depending on the activity), players are unable to choose the correct one, the activity ends showing a negative feedback.

During the activity, the virtual assistant plays a series of audio clues to guide the player. These clues are played to explain the activity to players, when players choose an incorrect option, or when they do not do any action over a period of time (the timeout is set to 20 seconds).

The player cannot perform actions in the activity when a sound is played. This is to make audio instructions as effective as possible, to prevent possible distractions (because of their attentional deficits (Hernández Martínez et al., 2011)) and to avoid an impulsive answer without

hearing or understanding the instructions (because of their deficits in executive control functions (Daunhauer et al., 2014)).

The difficulty level chosen modifies the complexity of the sentences and their structure, but the mechanics of the activity is the same. It also modifies the number of available options.

4.2 Production Activities

These activities are focused on voice production, so players have to practice with their voice on such prosodic aspects as the tone of voice, the expression of emotions or the syllabic emphasis. Players engage in conversations with game characters, where they have to repeat different sentences related with the context of the conversation (as shown in Figure 4).

[Figure 4 about here.]

As in the comprehension activities, at the beginning of the activity, the virtual assistant introduces the activity, in order to relate it to the conversation in progress. Later, the game plays the sentence to be recorded. The player must record the sentence shown on the screen in order to continue the conversation with the character. To do this, the player must click on the microphone button and start recording. The trainer must decide if the player has repeated the sentence correctly or incorrectly. If the recording is correct, the game will show a positive feedback on the screen. If it is incorrect, the player can repeat the recording a certain number of times (2 or 3 times, depending on the activity). At the end, if the player cannot make the recording correctly, the game will show a negative feedback and the game continues.

These activities also have sound instructions during their development, in case of error and in case the player does not execute any action. As in the comprehension activities, the player cannot execute any action when a sound instruction is being played.

In these activities, the difficulty level also modifies the complexity of the sentence that the player has to record.

4.3 Prosodic Activities

The objective of these activities is to mimic the prosodic curve of a sentence. To do this, players have to move an object (a bus) on the screen using their tone of voice, as shown in Figure 5. In addition, the text of the sentence is shown below the curve, highlighting the ascent or descent of the curve, with the goal that the players can connect the curve with the sentence and learn to modify their tone of voice depending on the type of sentence (interrogative or affirmative). To indicate when the player has to make a break, the game uses a traffic light. Moreover, players are asked to reproduce the sentence in the normal way and humming. The objective is that players understand the intonation changes in both ways.

[Figure 5 about here.]

Before players start the activity, an example is reproduced: the game plays a recorded voice of the sentence and an object moves automatically, following the tone of speech. In these activities, the trainer also has to decide if the recording is correct or incorrect.

In order to translate the speech into object motion, we use a pitch tracker based on the fast lifting wavelet transform (FLWT) to extract the fundamental frequency of the user's voice in real time (Larson & Maddox, 2005). Production activities previous to the prosodic activities are used to calibrate the pitch tracker and to adapt it to the current player.

4.4 Visual Activities

The visual activities are more variable in their development. They are introduced to increase the playability of the video game.

- Hidden pieces: the player has to search, using a magnifying glass, for several pieces hidden between the pages of a book. These pieces are only seen when the player passes the magnifying glass over them.
- Puzzle: the player has to form a drawing of the moon using the pieces found in the previous activity. According to the difficulty level, the player will have more or less clues on how to build the shape. Also, a sound stimulus will be played when the player puts a piece incorrectly, in order to tell the player that an error has been made.

- Cookies: the objective of this activity is that the player understands the difference between an offer (question) and an affirmation (declarative), from the point of view of the character of the game. Two options (a declarative and a question) are presented and the player is asked to select the option that can be used to offer cookies to another person.
- Dusting a bookshelf: the player has to clean the dust from a bookshelf with a cloth handled by the mouse.
- Hidden map: this activity is similar to the hidden pieces activity. This time, the player has to find a map hidden on the bookshelf.

5 Evaluation Strategy

Figure 6 shows the strategy that has been followed for the design, implementation and evaluation of the video game. The goal of the strategy is to answer the three research questions (RQ1, RQ2, RQ3) presented in the introduction of the article. This fact determines the relative weight of the different usability aspects (represented by the size of the circles, according to (Quesenbery, 2003)) and the evaluation activities that have been performed. We used user centered design, in which the users are the main focal point and have a significant impact on the design decisions. In this approach, the use of both usability inspection methods and testing with real users is recommended, combining qualitative and quantitative methods (Lewis, 2014).

[Figure 6 about here.]

The development of the video game started with a script that specifies in detail the narrative of the graphic adventure and the activities. A set of sketches was also used to describe the different scenes, the characters and the interaction of the video game. Cognitive walkthroughs (Mahatody, Sagar, & Kolski, 2010; Nielsen, 1994; Polson, Lewis, Rieman, & Wharton, 1992) were applied by the experts in DS of the research team to ensure that the narrative and the activities were adapted to the way users with DS process information and to their cognitive and linguistic abilities. The goal was to make an attractive interface that engages the users. Another objective was to ensure an easy to learn interface that takes into account the special

characteristics of the users and that is tolerant to the players' possible mistakes. As a result of these cognitive walkthroughs, a series of design principles, presented in the discussion section, were collected. The authoring tool (Adobe Flash) permitted us to adapt the interface to the experts' suggestions. In this way, the different prototypes were reviewed by the experts under the perspective of the users with DS and their special characteristics.

A pilot test was performed with real users. The goals were to check the degree of engagement of the players, detect the difficulties of use, observe the reaction of the users to the video game feedback and check whether the system is effective for practicing pronunciation. Users did not have any training before the pilot test. During the pilot test, one of the researchers took the role of the trainer. One of the staff members of the education center played the role of observer during the test. In this test, objective and subjective evaluation methods were combined. On the one hand, the game itself recorded data about the interaction between the player and the game. On the other hand, at the end of the game session, evaluators gave a questionnaire on general aspects of usability to the player; this was complemented by observations that evaluators collected during the test. In addition, at the end of the questionnaire, the opinions of speech therapists or teachers who are dedicated to special education were collected. As the video game records the voice of the players, we used these recordings in a perception test, to check whether the players' pronunciation improves as they use the video game (see Section 6.3).

The pilot test was performed by 14 participants with Down syndrome: 10 boys and 4 girls (chronological ages: 13, 16, 16, 17, 18, 20, 21, 22, 23, 24, 25, 25, 26 and 39). All of them have a moderate intellectual disability. All tests were done face to face at the place where the participants performed their educational activities for people with intellectual disabilities. In particular, these tests were conducted at the Niu School (Barcelona), Aura Foundation (Barcelona), and the School of Special Education "El Pino de Obregón" (Valladolid).

In addition, tests were performed on 10 children aged between 6 and 9 years and on 10 adults. These tests were done with the purpose of comparing the results obtained with the results of the players with DS. We can thus see if there are significant differences in the use of the video game depending on the different player profiles, in order to shed light on the research question RQ1.

The questionnaire was prepared to obtain information from the users relating to various aspects of the game, such as the story, the assistant or the difficulty. We collected information from children and users with DS. We ignored the information that adults could provide for being not relevant to drawing conclusions about the game. The questionnaire contains a series of yes/no questions and some with two possible answers. Moreover, in some questions, users were asked to explain why they gave that answer or to give more details of things that they would change in order to improve the game. Children filled in the questionnaire on their own. In the case of the users with DS, the questionnaire was used to guide a short interview in which the evaluation team, with the help of the teachers, retrieved the information while trying to avoid the use of directed questions. The questionnaire is included in Appendix A.

After the pilot test, an evaluation report was prepared which is summarized in Section 6. A number of small changes were done in the interface concerning the aspect and position of the recording button, the simplification of the prosodic activity, and some improvements on the audio instructions and help messages.

Finally, the School of Special Education “El Pino de Obregón” made more field tests with real users (both players and trainers). After two months’ work (two hours per week in the language classes), they completed three sessions with each of the children with DS from the school that participated in the pilot test. We carried out a semi-structured interview with the teachers of the school in order to know their opinions of the video game and to get some feedback. We prepared a set of open-ended questions (see Appendix B) and allowed the teachers to respond in an open way. This interview triangulates the information obtained in the pilot test to shed light on the research questions RQ2 and RQ3, concerning the degree of engagement of the players, and the effectiveness of the use of the software for training pronunciation.

6 Results

In this section we present the results obtained in the different tests. The results are grouped according to the research question to which they relate.

6.1 Analysis of Interaction Difficulties of Players with DS

The main source of information to address this topic is the log file, which records the actions performed by the users during the pilot test. The number of players that participated in the pilot test and the total playing time are shown in Table 1. There are significant differences in time between the three user groups ($p < 0.001$ Kruskal-Wallis test). The game does not offer a playable challenge to adults and children, so the game is carried out without interruption. Adults play the game faster than children ($p = 0.005$ Mann-Whitney test). Moreover, the game time for people with DS is significantly higher than for adults ($p < 0.001$ Mann-Whitney test) and children ($p < 0.001$ Mann-Whitney test). There is also a high standard deviation in people with DS, mainly due to their heterogeneity. They have specific characteristics which could produce varied response times. Therefore, people with DS have more difficulty solving the challenges offered by the game than adults and children.

[Table 1 about here.]

Table 2 presents the data obtained in the comprehension activities. This table contains average data on resolution time in seconds, errors (wrong choice of sentence in conversation), clicks on the speaker button (to hear a phrase again), sound aids (those reproduced when the user takes more time than expected to perform some action), and clicks that are made on the interface when mouse clicks are disabled (e.g. when an audio is playing). The Kruskal-Wallis test showed differences between the three user groups in time ($p < 0.001$), errors ($p = 0.003$), and clicks ($p < 0.001$). Resolution times are higher for people with DS than for adults (166.0 vs. 105.9 with $p < 0.001$ Mann-Whitney test) and for children (166.0 vs. 118.8 with $p = 0.007$ Mann-Whitney test). Besides, people with DS made more errors than adults (1.0 vs. 0.1 with $p = 0.007$ Mann-Whitney test) and children (1.0 vs. 0.1 with $p = 0.007$ Mann-Whitney test). Finally, the number of incorrect clicks made by children is worth noting. It is higher than the number of incorrect clicks made by adults (1.6 vs. 0.0 with $p < 0.001$ Mann-Whitney test) and by people with DS (1.6 vs. 0.1 with $p < 0.001$ Mann-Whitney test). This is because the audio instructions are numerous and repetitive, something that is not harmful to people with DS, but which slows the game down and makes children get impatient because they want to move forward but they have to wait until the audio ends.

[Table 2 about here.]

Table 3 shows the data obtained in production activities (the variables represented are the same as those shown in Table 2). There are significant differences in time ($p < 0.001$), number of errors ($p < 0.001$), phrase repetitions ($p < 0.001$), and clicks ($p = 0.027$) between the three user groups using the Kruskal-Wallis test. Again, observed times are higher for people with DS (335.9 vs. 173.1 compared with adults, with $p < 0.001$; 335.9 vs. 189.5 compared with children, with $p < 0.001$; Mann-Whitney test), but with greater differences with respect to the comprehension activities. The number of errors is much higher for people with DS, due to the difficulties in speech production of this population (4.4 vs. 0.1 compared with adults, with $p < 0.001$; 4.4 vs. 0.3 compared with children, with $p < 0.001$; Mann-Whitney test). Also significant is the number of clicks on the speaker button (phrase repetitions) that people with DS made (3.7 vs. 0.0 compared with adults, with $p < 0.001$; 3.7 vs. 0.6 compared with children, with $p = 0.012$; Mann-Whitney test). Some users with DS had problems with the reading and needed to hear the sentence again in order to do the recording better. In these activities, again, the children did a lot of incorrect clicks, more than other user types (6.2 vs. 1.0 compared with adults, with $p = 0.013$; 6.2 vs. 2.2 compared with people with DS, with $p = 0.026$; Mann-Whitney test).

[Table 3 about here.]

Table 4 presents the data of prosodic activities. This activity was difficult for all users because the melodic curve proposed was too demanding to follow in real time. It has two parts, one with lexical content and another without it, based on a humming activity. The part based on the humming of the sentence causes confusion in all players, because it is hard to remember the hummed pattern once the recording starts.

[Table 4 about here.]

The results of the visual activities are shown in Tables 5 and 6. First, the average time spent in the hidden pieces activity is shown. The Kruskal-Wallis test shows that there are significant differences between the three user groups ($p < 0.001$). Again, players with DS employ more

time to complete the activity (100.9 vs. 27.4 compared with adults, with $p < 0.001$; 100.9 vs. 50.1 compared with children, with $p = 0.002$; Mann-Whitney test). This result is symptomatic of the deficits in executive control functions, and more specifically in the inhibitory control of response, which provides the ability to not answer a question or stop giving the answer if it is not adequate (Daunhauer et al., 2014).

Second, average time spent and average number of errors is shown for the puzzle activity. There are significant differences between the three user groups (time: $p < 0.001$; errors: $p = 0.010$; Kruskal-Wallis test). On the one hand, players with DS take longer to perform the activity (54.0 vs. 21.5 compared with adults, with $p < 0.001$; 54.0 vs. 27.9 compared with children, with $p < 0.001$; Mann-Whitney test). On the other hand, adult and child players committed very few errors in assembling the puzzle, but players with DS made many more mistakes until they assembled the puzzle correctly (1.8 vs. 0.0 compared with adults, with $p = 0.010$; 1.8 vs. 0.2 compared with children, with $p = 0.045$; Mann-Whitney test). People with DS have difficulties processing information, because, in this population, the two types of local and global processing seem to be dissociated (i.e. the ability to attach pieces of information to a whole, or the inverse, to segment comprehensive information into parts). In addition, these people often have a bias toward global information processing; they do not process and integrate details appropriately, especially if they are not connected to a whole, as in this case (D'Souza, Booth, Connolly, Happé, & Karmiloff-Smith, 2016).

Third, the average time spent and average number of errors is shown for the cookies activity. In this activity, there is not much difference between the times of all types of users ($p = 0.051$ Kruskal-Wallis test) or between errors ($p = 0.098$ Kruskal-Wallis test). Finally, the average time spent in the dusting a bookshelf and the hidden map activities is shown. There are significant differences in both activities between the three user groups ($p = 0.001$ for dusting a bookshelf; $p < 0.001$ for hidden map; Kruskal-Wallis test). As can be seen, players with DS take longer to perform both activities, especially searching the map (dusting a bookshelf: 33.3 vs. 15.9 compared with adults, with $p = 0.001$; 33.3 vs. 21.2 compared with children, with $p = 0.016$. Hidden map: 55.9 vs. 11.2 compared with adults, with $p < 0.001$; 55.9 vs. 16.2 compared with children, with $p < 0.001$. Mann-Whitney test).

[Table 5 about here.]

[Table 6 about here.]

[Table 7 about here.]

[Table 8 about here.]

Apart from the actions recorded in the log file, the second source of information for analyzing the differences between the types of players is the pilot test questionnaire, whose results are presented in Tables 7 and 8. The most relevant differences between the answers of the children and the players with DS are related to the understanding of the story (100% vs. 64.3%), the perceived ease of use (100% vs. 78.6%), and the perception that the video game is long (33.3% vs. 71.4%).

Users' observation performed by the research team during the pilot test confirms that, in spite of the fact that players with DS can easily interact with the system, the differences reported in the previous paragraphs are clearly observed during the playing sessions: players with DS are slower and make more mistakes.

6.2 Degree of Engagement of the Players

Users' observation carried out during the pilot test and the opinions collected in the teachers' interview performed after the field test allow us to describe the behavior of the users with DS during the gaming sessions as attentive, taking the video game seriously; they like the video game and feel motivated to play; and they focus their attention, even students with attentional problems. Teachers reported:

“It seemed curious to us how quiet they are, even the ones who are more active. Some of them are very focused”.

“They feel the responsibility of doing it well, they are super serious and super formal”.

“It caught our attention how they remain staring at the screen”.

“Paula is a girl that has attention deficit. [During the game session] she has been very attentive, focused and giving the answer immediately. She has done everything right. But above all, the attitude of being calm, following everything correctly, super responsible”.

“What I liked most is the responsibility of the children. The video game motivates them. For instance, Paula is a hyperactive girl and with the video game she is focused ... with full attention, she was delighted, answering correctly”.

“The video game is useful for motivating the students in their daily learning, for focusing their attention”.

Tables 7 and 8 show the results of the questionnaire for each group of players interviewed after the pilot test. Players expressed the fact that they liked to test the game and would like to continue the story later. There was also fair unanimity in relation to the assistant of the game and the opportunity to talk to the characters of the game. Finally, the vast majority were satisfied with the character graphics and game scenes. Teachers reported that their students insisted on replaying “The magic stone” even weeks after the playing sessions. There was also much unanimity with respect to the entertainment and ease aspects of the game. Most of the players believed that the game is funny and easy (first and second row of Table 8). The fun aspect is important to improve motivation. The opinions about the ease of the game suggest that the difficulty level is adequate. The overall assessment of the game is thus positive, as described by the teachers:

“When they were finishing, we asked them for an assessment: ‘what do you think?, do you like it?, don’t you like it?’. All of them said that they liked it a lot. ‘Is there anything that you don’t like?, what would you remove?’ They said no, that they liked everything”.

Players with DS believe that the game is long, while the children think that it is short (third row of Table 8). We already mentioned in the previous subsection that players with DS took more time to finish according to the actions recorded in the log file (see Table 1). Although they had to work longer, all of them kept working until the end of the game session, which is also a

consequence of the high degree of engagement of the players with DS. The engagement is also reflected by the fact that there is a high level of immersion in the game, as commented by the teachers:

“Some children are so imaginative that they get so much into the role that they are incorporated inside the video game, they get into the video game. They are so imaginative. Some of them have come to dream of the video game. A boy has come to say: ‘I am a hero’ ”.

6.3 Analysis of the Improvement in Pronunciation

In order to assess the efficiency of the activities of production, we analyzed the quality of the pronunciation of the players with DS in the sentence “Hola tío Pau, ¿sabes dónde vive la señora Luna?” (Hello uncle Pau, do you know where Mrs. Moon lives?). This sentence has been selected because users are required to repeat it several times during the video game so that we can analyze whether the last production improves with respect to the first one.

Seven evaluators judged in an ABX perception test (Pisoni & Lazarus, 1974) which one of the utterances in a pair seems more natural or appropriate. We selected people without any specific background on speech therapies in this test, as we were interested in the perception of normal people concerning the possible improvement of the players. For that reason, we decided not to make any more specific question about intonation and prosody because we didn't want to condition the opinion of the evaluators. This is not an intelligibility test (the spoken sentence was always the same) but a perception test of the overall differences between a pair of utterances of the same sentence.

The pair is composed of the first and last production of the sentence during the working session. The first of them was uttered after reading and listening to it. The second one was produced after receiving a visual stimulus. The list of pairs (one per player) of utterances are presented in a web interface. Utterance A and utterance B are assigned randomly to the first and last production of the sentence. The evaluators can listen to utterance A and utterance B as many times as they consider necessary.

Two questions are presented to the testers:

- Q1: Do you perceive any difference between both utterances?
- Q2: Do you have any preference?

The possible answers to Q1 are in a 5-point Likert scale: 1 means “I do not perceive any difference” and 5 means “I perceive the differences very clearly”. We obtained 98 answers (14 subjects and 7 raters) with a percentage of consistency of 71.4% (tolerance = 3). Table 9 shows the statistics per speaker. Low values evidence small differences in all the cases. All the mean values are over 3, which indicates that the evaluators perceive differences between the utterances.

[Table 9 about here.]

The possible answers to Q2 are: “I prefer utterance A”, “I prefer utterance B”, “No preference: both are good”, “No preference: none is good”, “No preference: I do not perceive any difference”. We obtained 98 answers (14 subjects and 7 raters) with an inter-transcriber consistency Kappa Fleiss index of 0.224 which is “Fair agreement” in the Landis and Koch (1977) scale. Most of the judgments consider that the last utterance is the best one (64.2% with p -value=0.014 binomial test). Additionally, as can be seen in Table 9, there is a clear dependence on the player. None of the evaluations observed any improvement when evaluating the players spk2 and spk10. After checking the control forms of the test, we observed that these two players were not able to pay attention to the game, which justifies the bad results. Indeed, the percentage of preference of the last utterance is 74.3%, after removing these two speakers. The speakers that improve are 8 out of 14, as can be seen in Table 9.

Apart from this perception test, we have also obtained the opinions of the teachers from the teachers’ interview performed after the field test. Teachers think that the video game helps the students to improve their pronunciation. Even in the same session, they noticed pronunciation improvements. Players also improve along the different sessions. Mainly, they noticed improvements in the use of pauses and in the control of intonation of interrogative sentences. Teachers reported:

“They have improved. Within the same session it may be that they improve, but over the days they do clearly better”.

“Paula is a hyperactive girl ... She was making very short breaks and she has corrected it a lot. She reads well, but she read super fast and with low intonation, and with the repetition she has improved a lot. Of the girls you see that change, amazing”.

“The video game allows us to work on interrogative sentences and on pauses”.

When teachers were asked about the reasons for the engagement and the improvement of the players, they assessed that the context and the story of the video game are appropriate, which is essential for developing the activities with proficiency. Although players miss some details about the story, the teachers assessed that the players understand the main objective of the video game (searching for a magic stone) and use of a significative context is very helpful for interpreting the pragmatics of prosody. They commented:

“They do comprehension activities quite well. We are surprised that they have been working so positively. They understand the context, even when they have to choose between three possible answers. It is not difficult for them, just the opposite”.

“They do understand the story, they know that they have to search for a magic stone”.

Teachers recommend the use of the video game, because of the following reasons. First, because it has interesting activities, which they think are fun. Second, because students can play independently. Finally, because it allows them to work with the students in a different way. In regular lessons, teachers are used to being in front of the students explaining the exercises. However, using the video game, they do not need to explain anything, just sit next to the players, help them if required and evaluate the activities. Comments of the teachers:

“I would recommend the video game to a colleague, to a teacher of children with Down syndrome. I would recommend the video game because it provides a dynamic of repeating in a playful way, functional and quite independent. I’m beside, not in front of them”.

“I would use the video game for repetitive learning, for pauses, and for making interrogative sentences in a mechanical way. For training it is great”.

“I would use it in the classroom”.

Teachers think that the instructions and the feedback of the video game are suitable. Besides, they believe that the virtual assistant is useful. Teachers also like the graphics of the video game. They reported:

“The feedback provided by the video game, the applauses, is wonderful”.

7 Discussion

The experimental results show that performing repetition exercises using the video game can lead to improvements in the production and use of prosody by speakers with DS. Despite the fact that players with DS use the video game with greater difficulty than other users (RQ1), they have a high motivation to use it (RQ2), which makes them strive to do the activities, with the consequent improvement (RQ3). We think that this is an excellent result because it encourages a repetition of the strategy in other scenarios to achieve engagement of users with DS in serious games. We give the credit for this to the application of a number of HCI design elements that have taken into account the special characteristics of the users such as: deficits in language development, problems with short term memory, attention deficit, lack of motivation, bad response to frustration, problems to process information coming from different channels, and difficulties in understanding the meaning of iconic symbols. We focus on these characteristics in the following subsections, referring to the HCI design elements that could have influenced the achievement of such results.

7.1 Motivating the Users with a Graphic Adventure

People with DS usually have low motivation for difficult tasks like educational ones, because of their cognitive and learning problems. The inclusion of learning activities that must be accomplished in order to progress in the game has been one of the keys to motivating the users.

The initial story (an animated cartoon of about 3 minutes) engages the players because they feel represented as the main character of the action of the video game. For the users, doing the activities in a proper way became a significant goal. As a result, almost all the players keep doing the activities until the end of the game without showing signs of fatigue or boredom.

The scenes of the video game represent real life situations (for practicing prosody in context), but the narrative is based on a fantastic story of a magic stone to be rescued. One of the teachers argued this fact as an important engaging factor, as many of the children and adolescents with DS have a remarkably imaginative character.

7.2 Attention Deficit and Multimodal Interface

The teachers were surprised by the result of the game with students with deficit of attention. Even the most problematic students in terms of attention carried out the activities until the end. We attribute this level of engagement to the high motivation mentioned in the previous subsection, and to the combination of attractive graphics and sound stimulation.

As for the aesthetics, bright colors are used, but we avoid childish cartoons. One of the challenges of developing the video game is to find an adequate content and design presentation. People with DS are more competent in social and emotional skills than in cognitive capabilities. As they grow older, their preferences and tastes become more mature, but the cognitive capabilities lag behind (Feng et al., 2010). On the one hand, computer programs that appeal to young people are usually too advanced for adolescents with DS. On the other hand, programs to learn skills at the right level tend to be boring for them, because they do not match their age and interests.

The sound stimulation consists of the voice of the voiceover on the one hand, and on the turns of the assistant (the parrot) on the other. Both voices use a friendly tone to help and guide the user during the adventure. The attitude of the voiceover and the assistant is patient, repeating the messages as many times as necessary, with the aim that the user will focus on the activity.

7.3 Multimodal Feedback and Intolerance to Frustration

Performing tasks that can cause experiences of failure may demotivate people with DS, because of their low tolerance to frustration. Therefore, the feedback offered to the user by the game is always positive, both when the user performs the activity correctly and incorrectly. When the player fails, a sad face appears accompanied by a message that encourages the player to try again. When the player performs the activity successfully, the sound of applause and the corresponding icon rewards the player. We have observed that, in spite of the fact that this positive message is always the same, the users respond positively to it, mostly with a smile. Moreover, after a number of incorrect attempts, the video game continues with the following activity.

We avoided the use of rankings, scores and other similar rewarding strategies, because we didn't want to stress the users. Instead, the role of the teacher was shown to be very relevant to assisting the student in improving and getting positive feedback. The possibility of giving automatic feedback was also abandoned because the risk of having an incorrect feedback in real time is high, leading to potential undesirable situations. We advise against the use of the game in autonomous mode (without the assistance of a teacher or a parent) because we predict that some players would focus on progressing in the graphic adventure without putting the required effort into doing the activities (each activity has a maximum number of tries after which the game continues).

7.4 Audio Messages to Help the Abstraction of Concepts

The icons and scenes of the video game are visually intuitive and simple, with the goal of avoiding ambiguities. Furthermore, the multimodal output helps the players as the voiceover pronounces the name of the represented objects when they appear (magnifying glass, book, map, etc.) and also informs about the current scene (library, bus stop, etc.).

All the activities provide oral instructions to the user. At the beginning of the activity, the voiceover explains what has to be done in the activity. If the user gets stuck on the activity, more instructions are provided. These instructions seem to work properly, guiding the user through the different activities. In addition, the sound auxiliary aids, which are reproduced when the

player does nothing for a certain time, are an essential resource, as most players with DS need more time to analyze the activity and do it.

In some cases, we could not avoid confusions, as we observed that some players started talking before pushing the microphone button, or pushed the speaker button to record. This highlights the difficulties of users with DS to understand clearly the purpose of each interface element.

7.5 Deficits on Language Development

The learning goals of the video game are related with the training of a particular aspect of language such as prosody. It has been shown that the repetition activities, in combination with the help of the teachers, result in the improvement of the oral production of the players. Additionally, experimental results show that users with DS make more mistakes and need more time to finish these activities than adults and children. Therefore, comprehension activities introduce some level of difficulty for users with DS, which helps to improve both lexical-semantic comprehension and prosodic perception in specific contexts.

The use of short and simple sentences that appear in the messages of the program (both written and oral messages) has been shown to be useful in minimizing the impact of these language development deficits. The implementation of three levels of difficulty permits the teacher to adapt the game to the specific characteristics of each user.

7.6 Difficulties for Processing Information and Poor Short Term Memory

Visual activities are introduced to increase the playability of the game. The results obtained showed the difficulties of people with DS to do these activities, because of their cognitive difficulties. For instance, in the hidden pieces activity, the pieces that the player has to look for are too small and some players have trouble finding them. In the activity of dusting a bookshelf, some users leave areas uncleaned, because those areas are too small to be seen with the naked eye. In the hidden map activity, some players have trouble locating the map for the same reasons as in the hidden pieces activity.

Additionally, results show significant differences with respect to the time taken to do the

activities by the different user groups. Users with DS are slower than the rest of the users, and again, audio stimulus has been shown to be useful to remind them of the activity's goal when the players take more time than expected to perform the task.

7.7 Problems to Integrate Information Coming from Different Sources

This article presents a good example in which people with DS benefit from the use of multi-modal information coming from two different channels, such as audio and image. J. C. Martin (1998) classifies the possible channel combination modes into transfer, equivalence, specialization, complementarity and redundancy. Redundancy has been shown to be useful to disambiguate the icons and scenes of the video game and to offer feedback to the user. The audio channel is specialized in giving messages that reinforce the argument of the graphic adventure and remind them of the goal of the activities, given the poor short term memory of the users.

Problems for integrating information from different channels have been observed mainly in the prosodic activities. In these activities, players have to move an object (a bus) on the screen using their tone of voice. The objective is to mimic the prosodic curve of the sentence. Players with DS find this activity difficult to understand. Users able to read generally focus on the text, ignoring the bus animation and the prosodic contours. Even for non reader users, teachers manifest serious doubts about whether the players understand that the bus movement is a result of their voice. The difficulty may also be connected with the fact that the elements of these activities are too abstract for users with DS.

7.8 Limitations

We have included three levels of cognitive and linguistic difficulty in the video game (see section 3.3), but this categorization may not be accurate enough to adjust the game dynamics to the specific profiles of all potential users. It should be noted that there is great heterogeneity in the profiles of cognitive and linguistic development among people with Down syndrome (Abbeduto, McDuffie, Thurman, & Kover, 2016; Dierssen, Ortiz-Abalia, Arqué, Martínez de Lagrán, & Fillat, 2006). During the testing, the teacher or speech therapist helped players with the obstacles they found. We think that adapting the video game to the whole possible spectrum

of people with Down syndrome is a difficult challenge and may not even be possible. We assume that even though we have made a remarkable effort in adapting the interface, the role of the assistant teacher or speech therapist remains key.

The video game is a tool of educational support in the areas of prosody and pragmatics, while at the same time it serves to build a corpus of knowledge about prosodic and pragmatic limitations of the individuals affected by Down syndrome. The perceptual test performed has demonstrated that playing with the video game improves the quality of the vocal production of the users (see section 6.3). In this research, we have focused on the definition of the interface that the players use to perform the activities, but the effectiveness of the training was not measured in terms of each particular prosodic ability, but in terms of the overall perceived quality. In further steps of the project, specific tests of prosodic abilities will be conducted to obtain qualitative and quantitative measures of the learning of prosodic skills. The results of these tests will serve to refine the activities, but the structure and the elements of the interface will remain, since it has been shown that they serve to engage and motivate adolescents affected by DS.

8 Conclusions and Future Work

This article has described the design and implementation of a video game whose main objective is to improve communication skills, especially those related to prosody, of individuals with intellectual disabilities, in particular, those affected by Down syndrome. The use of a video game proved to be a promising research line, since it improves the motivation and engagement in the learning process. There exist tools in the field of speech therapy and audition that serve the professionals as assistance systems, but the innovative aspect of our work is that it incorporates activities of production and perception of prosodic phenomena through the combined strategies of reading aloud, imitative speech and resolution of speech acts. Players have to do activities that involve comprehension, production and prosodic tasks, together with visual activities. Nevertheless, due to the difficulties that people with DS encounter in human-computer interaction, we have paid special attention to the user interface, since it determines the interaction of the players with the video game. This task has turned out to be difficult, since there are not too

many video games oriented to this type of users.

In this sense, the following eight key elements have been shown to be efficient in engaging and motivating adolescents with Down syndrome in the activities of the video game. The authors of this paper encourage game designers and designers of interactive applications to use them as guidelines for the development of tools for people with cognitive impairments because they are founded in theoretical and descriptive studies about limitations of individuals with intellectual disabilities.

1. Adapt the video game to the users' needs with a system of levels and reinforcements according to their individuality. The use of several levels of difficulty allows the complexity of the activities to be adjusted. Positive reinforcements motivate users when they succeed in the activities. In order to avoid frustration, allow the players to continue the game even if they do not succeed in the activities.
2. Adapt the design of the scenes, items and characters to the specific difficulties of people with DS. Avoid using an excessive number of graphic elements that could distract attention from the relevant elements of the activity.
3. Use short and simple sentences and high frequency words. The instructions and sentences used in the video game must be in accordance with the cognitive level of people with DS.
4. Use audio messages to give instructions to the players and to assist them when their response time is high. The information provided by the audio channel helps users when it is redundant with the visual information.
5. Include visual clues in the scenes to help players continue the game and prevent abandonment. Highlight the visual elements needed to advance in the interactive story.
6. Consider the inclusion of a virtual assistant to guide the player during the game. The presence of this assistant ensures that the player knows at every moment what to do.
7. Use realistic scenes to favor game immersion and to provide a context for learning. Real life situations allow individuals affected by DS to better interact with other people.

8. Use fantastic elements to maintain the play-oriented nature of the proposed activities. These elements reinforce the players' interest in the video game.

The most relevant aspect of the results, due to the special characteristics of the people with DS, is that the majority of users kept playing throughout the video game with hardly any distractions. The players understand the main points of the narrative and are motivated by the fact that they have to solve the problem faced by the residents of a city. In general, they interact properly with the video game and succeed in doing the different activities (comprehension, production, prosodic and visual) and react well to positive reinforcement, even when they do not complete the activity perfectly. The users like to use the microphone to communicate with the characters of the game, and the different levels of difficulty in the video game allow them to progress through the various activities without getting blocked.

Another aim of the video game is that it can be used by teachers or speech therapists in special education centers as part of their educational activities. Related to this, the evaluation carried out shows a high level of acceptance and that the final product can be really useful for different areas of disability education.

Overall, we have succeeded in developing a tool that motivates players with Down syndrome because it imposes challenges and, as a consequence, they are attentive and improve their speech in a pleasant environment. Thus, we conclude that the use of a video game to improve the oral and communication skills of people with Down syndrome is a promising research line.

Further research includes the adaptation of the video game to other collectives with intellectual disabilities that also present communicative limitations, especially in prosodic and pragmatic abilities, such as people with Williams syndrome (Martínez-Castilla, Sotillo, & Campos, 2011; Martínez-Castilla, Stojanovik, Setter, & Sotillo, 2012), fragile-X syndrome (Ferrer, Bashir, Meryash, Johnston, & Wolff, 1991) or other neurodevelopmental disorders such as autism spectrum disorders (Filipe, Frota, Castro, & Vicente, 2014; Peppé, Cleland, Gibbon, O'Hare, & Martínez-Castilla, 2011). Testing the video game with users of new profiles will allow us to determine whether the HCI elements that have succeeded with adolescents with DS are also effective in other populations with intellectual disabilities.

The current version of the video game focuses on prosodic and pragmatic limitations of

adolescents with Down syndrome. It is future work to redesign the activities for the users to train other linguistic levels, such as phonetics (Rochet-Capellan & Dohen, 2015), syntax (Fortunato-Tavares et al., 2015) or vocabulary.

Currently the game is only available in Spanish, but we intend to extend the video game to users of other languages (among others, Catalan, English and French). To do this, besides the obvious aspects of translation, other modifications may also be required to adapt the video game to the specific cultural context.

Acknowledgments

This work was supported by Recercaixa, ACUP, Obra Social “la Caixa” (project “¡Juguemos a comunicar mejor! La mejora de la competencia prosódica como vía de integración educativa e inclusión social del alumnado con necesidades específicas de soporte educativo” - PZ611683-2013ACUP00202), by Fundacion BBVA (project “Pradia: la aventura gráfica de la pragmática y la prosodia” - CF613399), by Ministerio de Economía y Competitividad y Fondos FEDER (project “Videojuegos sociales para la asistencia y mejora de la pronunciación de la lengua española” - TIN2014-59852-R) and by Junta de Castilla y León (project “Evaluación automática de la pronunciación del español como lengua extranjera para hablantes japoneses” - VA145U14). The authors would like to thank all the participants who took part in the evaluation. We also would like to thank Yurena Gutiérrez, Patricia Sinobas, Valentín Cardeñoso, Ferran Adell and Juan María Garrido for their collaboration in the project.

A Script of the Interview with the Player

- Do you like video games?
- Have you ever played video games before?
- In general, did you like trying it?
- Would you like to play again in the future? Would you like to know how the story continues?

- Did you understand the story?
- Did you like the story?
- Did you like the parrot as the adventure's companion?
- Did you like talking to the characters of the video game?
- Would you like to create your own avatar (face, eyes, mouth, clothes, etc)?
- Did you learn something with the video game? What?
- Did you like the city images and the people?
- The video game seemed funny or boring.
- The video game seemed easy or hard.
- The video game seemed short or long.

B Script of the Interview with the Teachers

- How have you used the video game?
- Do you think that the video game helped them to improve their pronunciation?
- Do you think that the context and the story of the video game are useful to perform the activities?
- Are the instructions suitable? Is the feedback suitable?
- Is it useful that the teacher can decide if the activity is right or wrong?
- What is the reaction of the students during the game sessions?
- Do you think that the virtual assistant is useful?
- What do you think of the graphics?
- Would you recommend the video game?

References

- Abbeduto, L., McDuffie, A., Thurman, A. J., & Kover, S. T. (2016). Language development in individuals with intellectual and developmental disabilities: From phenotypes to treatments. In R. M. Hodapp & D. J. Fidler (Eds.), *Fifty years of research in intellectual and developmental disabilities* (Vol. 50, pp. 71–118). Cambridge, MA: Academic Press.
- Adams, E. (2013). *Fundamentals of game design*. Berkeley, CA: Pearson Education.
- Aguilar, L., Bonafonte, A., Campillo, F., & Escudero, D. (2009). Determining intonational boundaries from the acoustic signal. In *10th Annual Conference of the International Speech Communication Association (Interspeech)* (pp. 2447–2450). Brighton, UK: ISCA.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*. Arlington, VA: American Psychiatric Publishing.
- Black, B. (2006). Educational software for children with Down syndrome - an update. *Down Syndrome News and Update*, 6(2), 66–68.
- Boyle, E., Connolly, T. M., & Hainey, T. (2011). The role of psychology in understanding the impact of computer games. *Entertainment Computing*, 2(2), 69–74.
- Brandão, A., Brandão, L., Nascimento, G., Moreira, B., Vasconcelos, C. N., & Clua, E. (2010). Jecripe: stimulating cognitive abilities of children with Down syndrome in pre-scholar age using a game approach. In *7th International Conference on Advances in Computer Entertainment Technology* (pp. 15–18). Taipei, Taiwan: ACM.
- Brown, D. J., Ley, J., Evett, L., & Standen, P. (2011). Can participating in games based learning improve mathematic skills in students with intellectual disabilities? In *IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)* (pp. 1–9). Braga, Portugal: IEEE.
- Bruno, A., González, C., Moreno, L., Noda, M., Aguilar, R., & Muñoz, V. (2003). Teaching Mathematics to Children with Down's Syndrome. In *11th International Conference on Artificial Intelligence in Education* (pp. 1–8). Sydney, Australia: AIED.
- Buckley, S., & Bird, G. (2000). *Education for individuals with Down syndrome - an overview*. Portsmouth, UK: Down Syndrome Educational Trust.

- Buckley, S., & Bird, G. (2001). *Memory development for individuals with Down syndrome - an overview*. Portsmouth, UK: Down Syndrome Educational Trust.
- Buckley, S., & Sacks, B. (2002). *An overview of the development of teenagers with Down syndrome (11-16 years)*. Portsmouth, UK: Down Syndrome Educational Trust.
- Cagatay, M., Ege, P., Tokdemir, G., & Cagiltay, N. E. (2012). A serious game for speech disorder children therapy. In *7th International Symposium on Health Informatics and Bioinformatics (HIBIT)* (pp. 18–23). Nevsehir, Turkey: IEEE.
- Carmien, S., Dawe, M., Fischer, G., Gorman, A., Kintsch, A., & Sullivan, J. F. (2005). Socio-technical Environments Supporting People with Cognitive Disabilities Using Public Transportation. *ACM Transactions on Computer-Human Interaction*, *12*(2), 233–262.
- Chapman, R. S. (1997). Language development in children and adolescents with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews*, *3*(4), 307–312.
- Chapman, R. S., & Hesketh, L. (2001). Language, cognition, and short-term memory in individuals with Down syndrome. *Down Syndrome Research and Practice*, *7*(1), 1–7.
- Chapman, R. S., Schwartz, S. E., & Bird, E. K. R. (1991). Language Skills of Children and Adolescents With Down Syndrome I. Comprehension. *Journal of Speech, Language, and Hearing Research*, *34*(5), 1106–1120.
- Darejeh, A., & Salim, S. S. (2016). Gamification solutions to enhance software user engagement - a systematic review. *International Journal of Human-Computer Interaction*, *32*(8), 613–642.
- Daunhauer, L., Fidler, D., Hahn, L., Will, E., Lee, N., & Hepburn, S. (2014). Profiles of everyday executive functioning in young children with Down syndrome. *American journal on intellectual and developmental disabilities*, *119*(4), 303–318.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification. Using Game-design Elements in Non-gaming Contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (pp. 2425–2428). Vancouver, Canada: ACM.
- Dierssen, M., Ortiz-Abalia, J., Arqué, G., Martínez de Lagrán, M., & Fillat, C. (2006). Pitfalls and hopes in Down syndrome therapeutic approaches: in the search for evidence-based

- treatments. *Behavior Genetics*, 36(3), 454–468.
- D’Souza, D., Booth, R., Connolly, M., Happé, F., & Karmiloff-Smith, A. (2016). Rethinking the concepts of ‘local or global processors’: evidence from Williams syndrome, Down syndrome, and Autism Spectrum Disorders. *Developmental Science*, 19(3), 452–468.
- Escudero, D., Cardeñoso, V., & Bonafonte, A. (2002). Corpus based extraction of quantitative prosodic parameters of stress groups in Spanish. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Vol. 1, pp. I-481–I-484). Orlando, FL: IEEE.
- Escudero-Mancebo, D., González-Ferreras, C., & Cardeñoso-Payo, V. (2002). Quantitative evaluation of relevant prosodic factors for text-to-speech synthesis in Spanish. In *Seventh International Conference on Spoken Language Processing (ICSLP)* (pp. 1165–1168). Denver, CO: ISCA.
- Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2008). Computer Usage by Young Individuals with Down Syndrome: An Exploratory Study. In *10th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 35–42). Halifax, Canada: ACM.
- Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2010). Computer Usage by Children with Down Syndrome: Challenges and Future Research. *ACM Transactions on Accessible Computing*, 2(3), 1–44.
- Ferrier, L. J., Bashir, A. S., Meryash, D. L., Johnston, J., & Wolff, P. (1991). Conversational skills of individuals with fragile-X syndrome: A comparison with autism and Down syndrome. *Developmental Medicine & Child Neurology*, 33(9), 776–788.
- Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, 44(8), 1972–1981.
- Flores Lucas, V., & Belinchón Carmona, M. (2010). Dificultades en la comprensión de la ironía en personal con TEA, e implicaciones para la hipótesis de un déficit de habilidades de Teoría de la mente. In M. Belinchón Carmona (Ed.), *Investigaciones sobre autismo en español: problemas y perspectivas* (pp. 139–154). Madrid, Spain: Centro de Psicología Aplicada.

- Fortunato-Tavares, T., Andrade, C. R. F., Befi-Lopes, D., Limongi, S. O., Fernandes, F. D. M., & Schwartz, R. G. (2015). Syntactic comprehension and working memory in children with specific language impairment, autism or Down syndrome. *Clinical Linguistics & Phonetics*, 29(7), 499–522.
- Gallaher, K., van Kraayenoord, C., Jobling, A., & Moni, K. (2002). Reading with Abby: A case study of individual tutoring with a young adult with Down syndrome. *Down Syndrome Research and Practice*, 8(2), 59–66.
- González, C., Noda, A., Bruno, A., Moreno, L., & Muñoz, V. (2015). Learning subtraction and addition through digital boards: a Down syndrome case. *Universal Access in the Information Society*, 14(1), 29–44.
- González-Ferreras, C., Escudero-Mancebo, D., Vivaracho-Pascual, C., & Cardeñoso-Payo, V. (2012). Improving automatic classification of prosodic events by pairwise coupling. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(7), 2045–2058.
- Hernández Martínez, M., Pastor Duran, X., & Navarro Navarro, J. (2011). Attention deficit disorder with or without hyperactivity or impulsivity in children with Down's syndrome. *International Medical Review on Down Syndrome*, 15(2), 18–22.
- Jarrold, C., & Baddeley, A. (2001). Short-term memory in Down syndrome: Applying the working memory model. *Down Syndrome Research and Practice*, 7(1), 17–23.
- Kapp, K. M. (2012). *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*. San Francisco, CA: Pfeiffer.
- Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: A review. *Journal of Speech, Language, and Hearing Research*, 56(1), 178–210.
- Khan, T. M. (2010). The effects of multimedia learning on children with different special education needs. *Procedia - Social and Behavioral Sciences*, 2(2), 4341–4345.
- Kumin, L., Lazar, J., Feng, J. H., Wentz, B., & Ekedebe, N. (2012). A Usability Evaluation of Workplace-related Tasks on a Multi-touch Tablet Computer by Adults with Down Syndrome. *Journal of Usability Studies*, 7(4), 118–142.
- Lanfranchi, S., Cornoldi, C., Vianello, R., & Conners, F. (2004). Verbal and visuospatial working memory deficits in children with Down syndrome. *American Journal on Mental*

- Retardation*, 109(6), 456–466.
- Larson, E., & Maddox, R. (2005). Real-time time-domain pitch tracking using wavelets. In *Proceedings of the University of Illinois at Urbana Champaign Research Experience for Undergraduates Program* (pp. 1–12).
- Lazar, J., Kumin, L., & Feng, J. H. (2011). Understanding the Computer Skills of Adult Expert Users with Down Syndrome: An Exploratory Study. In *13th International ACM SIGACCESS Conference on Computers and Accessibility* (pp. 51–58). Dundee, UK: ACM.
- Lewis, J. R. (2014). Usability: Lessons learned ... and yet to be learned. *International Journal of Human-Computer Interaction*, 30(9), 663–684.
- Lopez-Basterretxea, A., Mendez-Zorrilla, A., Garcia-Zapirain, B., Madariaga-Ortuzar, A., & Lazcano-Quintana, I. (2014). Serious games to promote independent living for intellectually disabled people: Starting with shopping. In *Computer Games: AI, Animation, Mobile, Multimedia, Educational and Serious Games (CGAMES)* (pp. 1–4). Louisville, KY: IEEE.
- Mahatody, T., Sagar, M., & Kolski, C. (2010). State of the art on the cognitive walkthrough method, its variants and evolutions. *International Journal of Human-Computer Interaction*, 26(8), 741–785.
- Martin, G. E., Klusek, J., Estigarribia, B., & Roberts, J. E. (2009). Language Characteristics of Individuals with Down Syndrome. *Topics in Language Disorders*, 29(2), 112–132.
- Martin, J. C. (1998). TYCOON: Theoretical framework and software tools for multimodal interfaces. In J. Lee (Ed.), *Intelligence and Multimodality in Multimedia Interfaces* (pp. 1–25). Menlo Park, CA: AAAI Press.
- Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: articulaciones y sonidos de español*. Barcelona, Spain: Editorial Ariel.
- Martínez-Castilla, P., Sotillo, M., & Campos, R. (2011). Prosodic abilities of Spanish-speaking adolescents and adults with Williams syndrome. *Language and Cognitive Processes*, 26(8), 1055–1082.
- Martínez-Castilla, P., Stojanovik, V., Setter, J., & Sotillo, M. (2012). Prosodic abilities in Spanish and English children with Williams syndrome: A cross-linguistic study. *Applied*

- Psycholinguistics*, 33(1), 1–22.
- Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, 41, 85–139.
- McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the educational use of games*. Cambridge, UK: TEEM (Teachers evaluating educational multimedia).
- Ng, K. H., Bakri, A., & Rahman, A. A. (2014). A Review on Courseware for Down Syndrome Children. *Journal of Information Systems Research and Innovation*, 8, 56–65.
- Nielsen, J. (1994). Usability inspection methods. In *Conference Companion on Human Factors in Computing Systems* (pp. 413–414). Boston, MA: ACM.
- Ortega-Tudela, J. M., & Gómez-Ariza, C. J. (2006). Computer-assisted teaching and mathematical learning in Down Syndrome children. *Journal of computer assisted learning*, 22(4), 298–307.
- Patel, R., Kember, H., & Natale, S. (2014). Feasibility of augmenting text with visual prosodic cues to enhance oral reading. *Speech Communication*, 65, 109–118.
- Pazos González, M., Raposo-Rivas, M., & Martínez-Figueira, M. E. (2015). Las TIC en la educación de las personas con Síndrome de Down: un estudio bibliométrico. *Virtualidad, Educación y Ciencia*, 6(11), 20–39.
- Peppé, S., Cleland, J., Gibbon, F., O'Hare, A., & Martínez-Castilla, P. (2011). Expressive prosody in children with autism spectrum conditions. *Journal of Neurolinguistics*, 24(1), 41–53.
- Pettinato, M., & Verhoeven, J. (2009). Production and perception of word stress in children and adolescents with Down syndrome. *Down Syndrome Research & Practice*, 13, 48–61.
- Phonics Alive! (1996). *Advanced software*. Retrieved from <http://www.phonicsalive.com.au>.
- Pisoni, D. B., & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America*, 55(2), 328–333.
- Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. *International Journal of Man-Machine*

- Studies*, 36(5), 741–773.
- Prensky, M. (2007). *Digital Game-Based Learning*. Saint Paul, MN: Paragon House Publishers.
- Quesenbery, W. (2003). The five dimensions of usability. In M. Albers & B. Mazur (Eds.), *Content and Complexity: Information Design in Technical Communication* (pp. 81–102). Mahwah, NJ: Lawrence Erlbaum Associates.
- Rochet-Capellan, A., & Dohen, M. (2015). Acoustic characterisation of vowel production by young adults with Down syndrome. In *18th International Congress of Phonetic Sciences (ICPhS)* (pp. 1–5). Glasgow, UK: IPA.
- Rodríguez, W. R., Saz, O., & Lleida, E. (2012). A prelingual tool for the education of altered voices. *Speech Communication*, 54(5), 583–600.
- Santana, P. C., & Muro, B. P. (2011). Tangible interfaces to support the teaching of reading and writing to children with Down syndrome. *IEEE Learning Technology Newsletter*, 13(2), 9–12.
- Saz, O., Yin, S., Lleida, E., Rose, R., Vaquero, C., & Rodríguez, W. R. (2009). Tools and Technologies for Computer-Aided Speech and Language Therapy. *Speech Communication*, 51(10), 948–967.
- Schalock, R. L., Borthwick-Duffy, S. A., Bradley, V. J., Buntix, W. H. E., Coulter, D. L., Craig, E. M., ... Yeager, M. H. (2010). *Intellectual disability. Definition, Classification, and Systems of Supports (11th Edition)*. Washington, DC: American Association on Intellectual and Developmental Disabilities.
- Shafie, A., Wan Ahmad, W. F., Mohd, N., Barnachea, J. J., Taha, M. F., & Yusuff, R. L. (2013). “SynMax”: A Mathematics Application Tool for Down Syndrome Children. In H. B. Zaman, P. Robinson, P. Olivier, T. K. Shih, & S. Velastin (Eds.), *Advances in Visual Informatics: Third International Visual Informatics Conference* (pp. 615–626). Selangor, Malaysia: Springer.
- Shahin, M., Ahmed, B., Parnandi, A., Karappa, V., McKechnie, J., Ballard, K. J., & Gutierrez-Osuna, R. (2015). Tabby Talks: An automated tool for the assessment of childhood apraxia of speech. *Speech Communication*, 70, 49–64.

- Standen, P., Anderton, N., Karsandas, R., Battersby, S., & Brown, D. (2009). An evaluation of the use of a computer game in improving the choice reaction time of adults with intellectual disabilities. *Journal of Assistive Technologies*, 3(4), 4–11.
- Standen, P., Rees, F., & Brown, D. (2009). Effect of playing computer games on decision making in people with intellectual disabilities. *Journal of Assistive Technologies*, 3(2), 4–12.
- Stojanovik, V. (2011). Prosodic deficits in children with Down syndrome. *Journal of Neurolinguistics*, 24(2), 145–155.
- Taylor, P. (2009). *Text-to-speech synthesis*. Cambridge, UK: Cambridge University Press.
- Torrente, J., Del Blanco, Á., Moreno-Ger, P., & Fernández-Manjón, B. (2012). Designing serious games for adult students with cognitive disabilities. In *Neural Information Processing: 19th International Conference ICONIP* (pp. 603–610). Doha, Qatar: Springer.
- Troncoso, M. V., & Del Cerro, M. M. (1999). *Síndrome de Down: lectura y escritura*. Barcelona, Spain: Masson.
- Wan Ahmad, W. F., Muddin, H. N. B. I., & Shafie, A. (2014). Number skills mobile application for Down syndrome children. In *2014 International Conference on Computer and Information Sciences (ICCOINS)* (pp. 1–6). Kuala Lumpur, Malaysia: IEEE.
- Wuang, Y. P., Chiang, C. S., Su, C. Y., & Wang, C. C. (2011). Effectiveness of virtual reality using Wii gaming technology in children with Down syndrome. *Research in developmental disabilities*, 32(1), 312–321.
- Yussof, R. L., & Zaman, H. B. (2011). Scaffolding in early reading activities for Down syndrome. In H. B. Zaman et al. (Eds.), *Visual Informatics: Sustaining Research and Innovations* (pp. 180–192). Selangor, Malaysia: Springer.

ABOUT THE AUTHORS

César González-Ferreras is an Assistant Professor at the Department of Computer Science at the University of Valladolid, Spain. His research interests include human-computer interaction, spoken language processing and prosody recognition.

David Escudero-Mancebo is an Associate Professor at the Department of Computer Science

at the University of Valladolid, Spain. He is co-author of several publications in the field of computational prosody, both concerning modeling of prosody for text-to-speech systems and prosodic labeling of corpora. His research interests also include human-computer interaction.

Mario Corrales-Astorgano is working on a Ph.D. in Computer Science at the University of Valladolid, Spain. His research interests mainly focus on human-computer interaction, especially in the design and evaluation of user interfaces for people with disabilities.

Lourdes Aguilar-Cuevas is an Associate Professor at the Department of Spanish Philology at the Autonomous University of Barcelona, Spain. She is a specialist in experimental phonetics and phonology of intonation in Spanish and Catalan. She has extensive experience in the field of prosody and automatic modeling of prosody.

Valle Flores-Lucas is an Associate Professor at the Department of Psychology at the University of Valladolid, Spain. Her research interests focus on the study of the problems of language development and communication, mainly in autism and Down syndrome, and their relationship with the theory of mind.

Table 1: Number of players and total time taken to finish the game in seconds (mean and standard deviation for each type of player).

	#Users	Time	
		Mean	SD
Children	10	942.2	130.3
Adults	10	824.8	72.7
Down syndrome	14	1342.1	301.4

Table 2: Comprehension activities: time spent in seconds, number of errors, number of additional listenings of the phrase, number of speech aids and number of incorrect clicks (mean and standard deviation for each type of user in the 5 comprehension activities).

	Time		Errors		Phrase repetitions		Speech aids		Incorrect clicks	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Children	118.8	17.2	0.1	0.3	1.2	1.3	0.0	0.0	1.6	1.5
Adults	105.9	7.4	0.1	0.3	0.3	0.5	0.0	0.0	0.0	0.0
Down syndrome	166.0	43.3	1.0	0.9	1.6	1.7	0.6	1.0	0.1	0.5

Table 3: Production activities: time spent in seconds, number of errors, number of additional listenings of the phrase, number of speech aids and number of incorrect clicks (mean and standard deviation for each type of user in the 7 production activities).

	Time		Errors		Phrase repetitions		Speech aids		Incorrect clicks	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Children	189.5	31.3	0.3	0.7	0.6	1.1	0.1	0.3	6.2	6.0
Adults	173.1	6.5	0.1	0.3	0.0	0.0	0.0	0.0	1.0	0.9
Down syndrome	335.9	102.6	4.4	2.3	3.7	3.7	1.6	2.8	2.2	3.5

Table 4: Prosodic activities: time spent in seconds, number of phrase repetitions and number of humming repetitions (mean and standard deviation for each type of user).

	Time		Phrase repetitions		Humming repetitions	
	Mean	SD	Mean	SD	Mean	SD
Children	99.5	7.1	0.1	0.3	0.0	0.0
Adults	98.6	10.0	0.1	0.3	0.1	0.3
Down syndrome	106.5	10.4	0.4	0.7	0.4	0.5

Table 5: Visual activities: hidden pieces (time spent in seconds), puzzle (time spent in seconds and number of errors) and cookies (time spent in seconds and number of errors).

	Hidden pieces		Puzzle				Cookies			
	Time		Time		Errors		Time		Errors	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Children	50.1	15.3	27.9	5.7	0.2	0.6	27.6	6.7	0.4	0.5
Adults	27.4	8.6	21.5	2.3	0.0	0.0	22.1	5.0	0.0	0.0
Down syndrome	100.9	48.3	54.0	20.5	1.8	2.8	30.2	9.6	0.3	0.5

Table 6: Visual activities: dusting a bookshelf (time spent in seconds) and hidden map (time spent in seconds).

	Dusting a bookshelf		Hidden map	
	Time		Time	
	Mean	SD	Mean	SD
Children	21.2	10.9	16.2	7.1
Adults	15.9	9.1	11.2	3.3
Down syndrome	33.3	17.7	55.9	47.5

Table 7: Questionnaire results (yes/no questions).

	Children		Down syndrome	
	Yes	No	Yes	No
The user likes video games	100.0%	0.0%	92.9%	7.1%
The user is used to playing video games	80.0%	20.0%	78.6%	21.4%
The user enjoyed trying the game	90.0%	10.0%	92.9%	7.1%
The user would like to play again in the future / the user would like to know how the story continues	90.0%	10.0%	85.7%	14.3%
The user understands the story	100.0%	0.0%	64.3%	35.7%
The user likes the story	90.0%	10.0%	78.6%	21.4%
The user likes the parrot as the adventure's companion	90.0%	10.0%	85.7%	14.3%
The user likes talking to the characters of the video game	88.9%	11.1%	92.9%	7.1%
The user would like to create his/her own avatar (face, eyes, mouth, clothes, etc)	88.9%	11.1%	64.3%	35.7%
The user manifests to have learned something with the video game	70.0%	30.0%	78.6%	21.4%
The user likes the city images and the people	100.0%	0.0%	85.7%	14.3%

Table 8: Questionnaire results (two options questions).

	Option1	Option2	Children		Down syndrome	
			Option1	Option2	Option1	Option2
The video game seemed ...	Funny	Boring	88.9%	11.1%	85.7%	14.3%
The video game seemed ...	Easy	Hard	100.0%	0.0%	78.6%	21.4%
The video game seemed ...	Short	Long	66.7%	33.3%	28.6%	71.4%

Table 9: ABX test

Speaker	Q1		Q2				
	Mean	SD	Last better	Last worse	Both are good	None is good	No difference
spk1	4.4	0.8	6	0	0	1	0
spk2	4.3	1.1	0	5	0	2	0
spk3	4.3	1.1	7	0	0	0	0
spk4	3.6	1.3	3	4	0	0	0
spk5	3.7	1.0	3	1	3	0	0
spk6	3.7	1.1	5	1	0	0	1
spk7	4.3	1.0	4	3	0	0	0
spk8	3.0	1.3	4	2	1	0	0
spk9	3.0	1.2	3	2	2	0	0
spk10	3.1	1.3	0	6	1	0	0
spk11	4.4	0.8	7	0	0	0	0
spk12	3.1	1.2	1	3	3	0	0
spk13	3.3	1.4	4	2	0	0	1
spk14	3.0	1.2	5	0	0	0	2

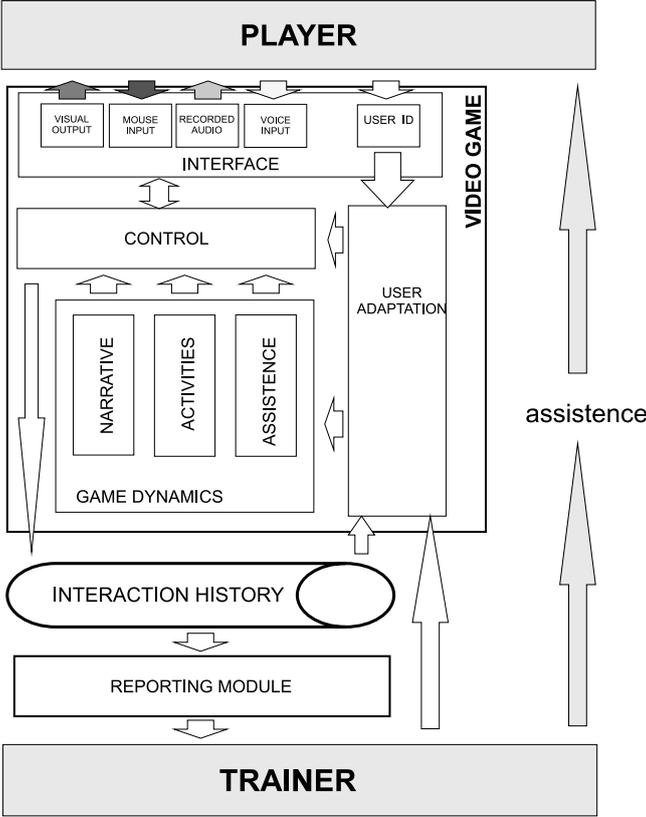


Figure 1: System architecture.



Figure 2: Main game screen, in which the player’s avatar (the girl) and the virtual assistant (the parrot) are visible.



Figure 3: Comprehension activity. The player has to choose the right option to continue the conversation.



Figure 4: Production activity. The player has to say the sentence to continue the conversation.



Figure 5: Prosodic activity. The user has to reproduce the intonation contour and to make the breaks at the traffic lights.

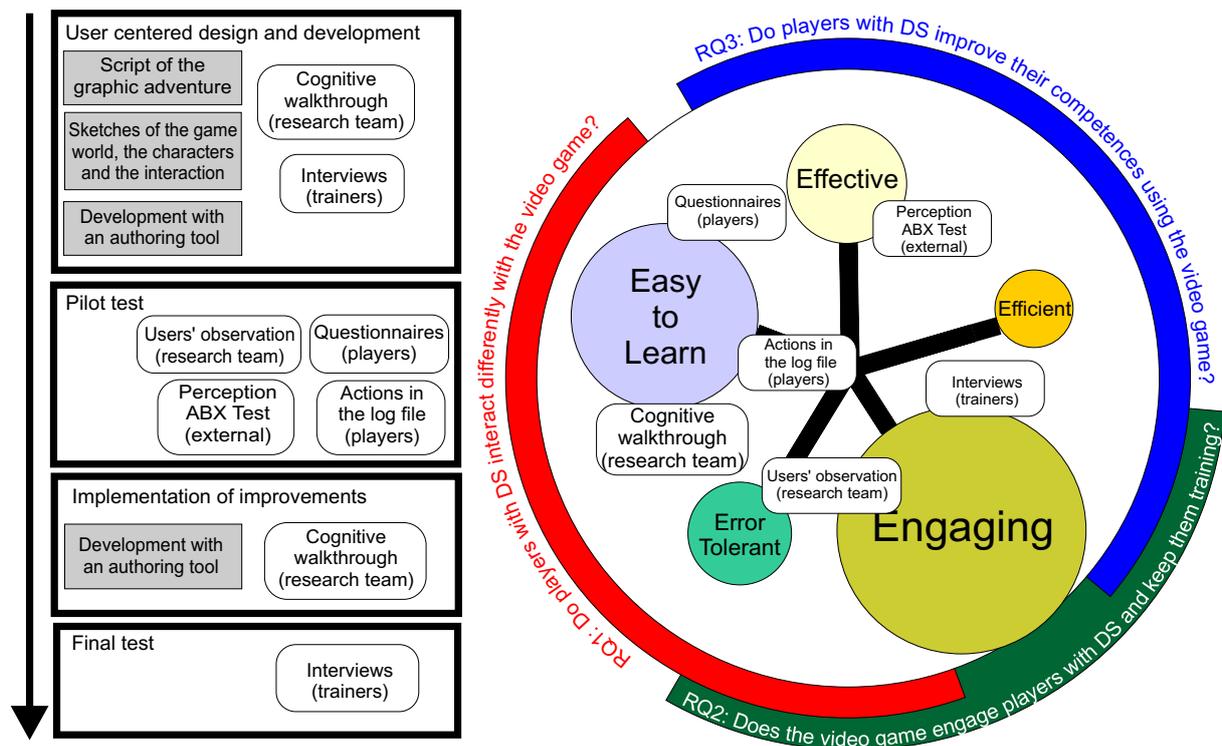


Figure 6: Evaluation strategy. On the left, temporal sequence of the design, implementation and evaluation activities. On the right the five dimensions that describe the different aspects of usability according to (Quesenbery, 2003), and their relation with the three research questions and the evaluation activities. The circles represent the usability aspects (the size reflects the importance of the aspect), the rectangles represent the development and evaluation phases, and the rounded rectangles are the evaluation activities with the participants in brackets.