Assisting Speech Therapy for Autism Spectrum Disorders with an Augmented Reality Application

Camilla Almeida da Silva1, António Ramires Fernandes2 and Ana Paula Grohmann3

1Faculdade de Computação e Informática, Universidade Presbiteriana Mackenzie, São Paulo, Brazil
2Department of Informatics, University of Minho, Braga, Portugal
3Associação para a Inclusão e Apoio ao Autista, Braga, Portugal

Keywords: Autism Spectrum Disorder, Augmented Reality, Augmented and Alternative Communication, Human-computer Interaction.

Abstract: Graphics based systems of Augmented and Alternative Communication are widely used to promote communication of people with Autism Spectrum Disorders. However, studies indicate the inability of some of these people in understanding the used symbols. This study discusses an integration of Augmented Reality in communication interventions, by relating elements of Augmented and Alternative Communication and Applied Behaviour Analysis strategies. An Augmented Reality based interactive system to support interventions is discussed, and the report of its usage in interventions with children suffering from Autism Spectrum Disorders is presented.

1 INTRODUCTION

Children with Autism Spectrum Disorders (ASD) are affected with various impairments in communication, social interaction and imagination, three major components of self-development. Regarding communication issues, these range from the total absence of language, 20% to 30% (Klin, 2006), to the lack of effectiveness in the communication process, reinforcing the need to promote initiatives to improve communication skills.

Applied Behaviour Analysis (ABA) interventions for autism in early childhood are an effective practice to improve socially relevant behaviours and teach new skills, through several established teaching tools and positive reinforcement strategies. Augmented and Alternative Communication (AAC) based approaches, in particular those based on graphics, are the most used in interventions in children with ASD, being highly relevant in communication promotion, reduction of behavioural problems, and environment awareness (Committee on Educational Interventions for Children with Autism, 2001).

The graphic based system usage is supported by the strong visual processing ability seen in many children with ASD. Nevertheless, imagination impairments caused by rigidity and lack of imagination can prevent the usage of some symbols (Herrera et al., 2012). In (Herrera et al., 2012) it is argued that some children with ASD, when confronted with a symbol, only capture a set of lines, shapes and colours, and the usage of communication symbols would lead to memorization and context association. The authors show that changes in the background or width of the lines are sufficient to prevent recognition on a previously learned symbol. Hence, they conclude that these symbols are meaningless to those children. In (Committee on Educational Interventions for Children with Autism, 2001) it is stated that the lack of symbolic capacity is one of the main handicaps in communication, reflecting the difficulty in learning conventional or shared meanings of symbols. To achieve communication, symbols must be meaningful to the child (Avila, 2011).

Computer animations, 3D graphics and sounds, may provide deeper engagement for children in their activities as suggested in several studies on the benefits of Information Technologies in interventions with children with ASD (Moore et al., 2000; Moore and Calvert, 2000; Goldsmith and LeBlanc, 2004). The use of tangible interfaces has also been suggested by (Raffle et al., 2009) to promote alternative ways of collaboration and communication, engaging the children for longer periods of time than with regular interventions.

Augmented Reality (AR), merging virtual ob-
jects into real environments, can be explored in some intervention areas for ASD children, such as self-awareness, augmented communication, emotion awareness and identification, social commitment and concept development (Herrera et al., 2006). These visually oriented approaches are highly suitable due to the remarkable ability of most autistic people to excel at visual spatial skills (Grandin, 2009).

Considering the relevance of AAC approaches, and the impact of symbols to promote communication, we performed a study on the usage of AR to assist interventions and expand language with children with ASD. The interactive AR system developed in this study uses both AAC and ABA principles to support interventions focused on facilitating graphic symbol comprehension. The system reinforces the communication symbols superimposing virtual objects and animations over the real environment screen view.

The system’s human-computer interaction was previously discussed with speech-language therapists to ensure its suitability for both therapists and children. The system’s features include the creation of activities with AR graphics and sound to enrich the symbol’s meaning, with both visual and audio reinforcements. The system is adaptable to each individual child, allowing to have multiple activity configurations, and supports a number of different activity templates to increase the options for the therapist.

In order to evaluate the system, the application was tested in speech-language sessions with four children with ASD. The tests fulfilled the framework proposed by Moore (Moore et al., 2000), where computer assisted learning systems for ASD children should focus on at least one on the major impairments; the projects should be based on current established practices; and evaluation should be performed in cooperation with actual therapists.

2 AUTISM SPECTRUM DISORDER

ASD is a group of developmental brain disorders affecting individuals from all races and cultures. It presents a wide range of possible symptoms and varying degrees of severity, the common feature being an early developmental disruption in socialization processes (Klin, 2006).

The Centers for Disease Control and Prevention (Centers for Disease Control and Prevention, 2013) estimate that about 1% of individuals have ASD in Europe, Asia, and North America. ASD is more common among boys than among girls, with a ratio of 5 to 1.

2.1 Interventions

The main goal of an intervention is to promote functional and spontaneous communication, social interaction and symbolic play to increase functional independence and quality of life. These interventions address communication, social skills, unsuitable behaviours, among other common issues caused by ASD (Myers et al., 2007). The intervention programs require a multidisciplinary base, involving behavioural therapy, educational programs, speech-language and communication therapy (Gadia et al., 2004).

Applied Behavioural Analysis methods, such as intense behavioural interventions, directed towards long term behaviour modification, have as goals to increase and maintain adequate behaviours, reduce maladaptive behaviours, and teach new skills (Myers et al., 2007). These methods have a short duration and are effective in both adults and children (Committee on Educational Interventions for Children with Autism, 2001). Discrete Trial Training (DTT) is also an ABA method to enhance skills such as attention, compliance, imitation, and discrimination learning, among others (Smith, 2001). It consists of a structured learning method whereby a complex learning sequence is divided in small steps, with reinforcements and assistance by the therapist where required.

The American Speech-Language-Hearing Association defines AAC as an area of clinical practice focused on impairment and disability patterns of individuals with severe expressive communication disorders (Sevcik and Romski, 2000). AAC methods and tools are also highly adaptable to the autistic child needs and personality.

Graphics based systems take advantage of the strong visual processing skills found in many individuals with ASD, and have been proven effective, through functional communication training, in increasing communication reception in small children, and to replace disruptive behaviour such as aggression, self-aggression and crying (Committee on Educational Interventions for Children with Autism, 2001). The environment, space, and time concepts are other examples where AAC can be used.

The Picture Exchange Communication System (PECS), developed by Bondy and Frost in 1995 (Bondy and Frost, 1994), is an AAC method for functional communication based on picture exchange. PECS follows ABA principles being structured in a sequence of steps. Initially the child is taught to initiate requests through figures. Later the child will start
building sentences, answering questions and, on latter stages, to make comments. Several studies (Charlop-Christy et al., 2002; Kravits et al., 2002; Ganz and Simpson, 2004) have shown the benefits of PECS, resulting in vocabulary acquisition and social behaviour improvements.

3 RELATED WORK

VR and AR techniques have been used extensively to assist interventions for children with ASD.

Project INMER-II (Belen, 2004) was designed based on an a serious game in a VR environment related to shopping activities. The results show increased levels of functional and symbolic comprehension, and imagination. Furthermore, one child was able to successfully generalize the experience to a real environment.

Virtual avatars with emotional facial expressions were used in project Virtual Messenger (Fabri and Moore, 2005). The goal was to help users to recognize facial expressions. In the first stage, the user would be asked to select the emotion corresponding to the avatar presented on screen, or to select a facial expression that corresponds to an emotion. The second level requires the user to select an emotion that is appropriate for a scenario presented on screen. When in the third level, the user is presented with an avatar’s facial expression, and is required to select a scenario that could cause the avatar’s expression. The results show that people with ASD were able to increase their level of understanding of emotions based on facial expressions.

Project AS Interactive was developed to support people with ASD to improve social skills through VR environments simulating a cafeteria and a bus (Parsons et al., 2006). The study performed with teenagers with ASD led to the conclusion that these simulated environments could be successfully used to help people with ASD to engage in social interactions.

In (Herrera et al., 2006) a system based on AR with infrared markers was described. The system would superimpose the card symbols on the user’s image in real time. However, this prototype was never tested on children with ASD. Later, in (Herrera et al., 2012) a AR system using Microsoft Kinect was proposed, having as main goals teaching self-consciousness, body scheme and postures, communication and imitation, using serious games to promote interaction with the children.

Tentori and Hayes proposed the Mobile Social Compass framework for the development of mobile AR systems. The Social Compass Curriculum is a behavioural and educational curriculum for social skills training, based on stories and visual paper clues to guide the child in both active and passive social interactions. The framework focuses in the concept of Interaction Immediacy, providing a set of visual clues to assist the child in anticipating situations (Tentori and Hayes, 2010).

MOSOCO is a mobile application implemented using the above described framework, using AR and visual guidance for children with ASD, to assist children with ASD in social interactions in real life. The system guides the children in the basic social skills defined in The Social Compass curriculum, encouraging them through interactive resources. The system’s evaluation was performed with three children and the results show that the system increased social interactions, both in quantity and quality, mitigating behavioural and social disruptions, and allowed the social integration of these children in neurotypical groups.

Another mobile AR application, Blue’s Clues, was proposed in (Escobedo and Tentori, 2011), to help children with ASD moving from one place to another. Blue’s Clues concept was developed for a school environment, aiming at providing the necessary mobility instructions, visual and audio clues, to guide students with ASD.

4 THE AR BASED PROTOTYPE

In this section we propose an AR based system that allows the creation of interactive activities supported on AAC and ABA principles to assist interventions with children with ASD, aided by a therapist.

The goal was not to impose a new methodology to the therapist, but to offer a support tool that could enrich the current daily activities in the speech-language sessions. Hence, all the development was initially discussed with speech-language therapists, to ensure that the prototype would fit painlessly in the current activities, and that its interface was clear and intuitive for both the therapist and the children.

The prototype allows the design of activities for each individual child attending to the difference in degree of the ASD, the individual personality, and the previous activities developed by the therapist. Within each activity, we aimed at providing as much flexibility as possible. Several templates were provided for this initial evaluation to guide the therapist in designing the individual activities.

The usage of AR is directed towards superimposing 3D models and animations on cards with communication symbols that the child holds or places on the
table. This creates a richer experience, and takes advantage of the visual skills found in many children with ASD.

The selection of AR as the underlying technology for this project aims at:
1. Promote a natural interaction with the computer through the use of tangible interfaces;
2. Facilitate the generalization of the acquired knowledge;
3. Allow the enrichment of traditional interventions based on communication symbols without imposing a new methodology on the therapy;
4. Eliminate the distractions caused by traditional interfaces such as mouse and keyboard.

In (Billinghurst et al., 2005) is suggested that for the use of tangible interfaces to be intuitive and natural, it is recommended that the selection of physical objects and interaction metaphors should be familiar to users, allowing them to interact with the system relying on previous skills and experiences. In our project we use communication symbols which are already familiar to the children. The child interfaces with the system by showing cards with communication symbols, which is a natural behaviour for the child under intervention. Furthermore, the acquired knowledge is implicitly associated with the communication symbols, hence, the child should be able to recall its experiences afterwards.

The prototype was built with open software tools and APIs such as irrKlang for audio, ARToolkit for the AR component and MySQL. These options aim at providing a low cost solution, and multiplatform future support.

4.1 The Back Office

The activity construction process takes place in a simple back office support application where the therapist can select a template and design interactive activities to explore with the children. Figure 1 shows a screenshot of the activity construction stage.

Initially the therapist associates words with symbol cards, 3D graphics, sounds and animations, and builds phrases with word sequences. The words can be further defined as belonging to user-defined categories such as actions, fruits, or animals.

The data set of words, and their associated media, is then used to construct the activities based on the available templates. Each template provides full parameterization including reinforcement definition, and error management.

4.2 Activity Templates

The available templates are: Free, Category, Discriminate, and Phrase. These are based on stages I (Free), III (Category and Discriminate), and IV (Phrase) of PECS. The possibility to use different templates also seeks to avoid having the autistic child engaging in repetitive tasks for long periods of time.

The basic procedure consists in showing a card in the field of view of the web cam. The recognition of the pattern in the card will trigger an action.

The Free template, the simplest one, allows both, child and therapist, to interact freely with the prototype. When designing an activity with this template the therapist can associate the action to be triggered when the card is shown or occluded. This action, which can be a graphic, an animation, or a sound, can act as a positive reinforcement depending on the intervention context, or just be used to engage the child into exploring the system. Additional actions can be specified when the card is occluded. As an example assume an activity based on cards with animals. When the child shows a card the system may display on screen a 3D model of the animal on top of the card. If the child covers the card with the hands then the animal sound can be played.

In the Discriminate template the child is required to identify an object based on a clue provided on screen, and present the card that matches the clue. Positive and negative reinforcements can be associated with each individual clue. The system supports the configuration for special cards that the therapist can include in the activity. An 'hint' card will trigger a user-defined hint when shown to the system. A 'next' card tells the system to move on to the next object. Finally the 'end' card will end the activity. These cards, if provided by the therapist, allow for a freer interaction from the child, since the child can control the activity when possessing all three special cards, or some subset. The special cards can also be used by the therapist to control the activity.

The Category template can be configured in two different ways. Initially the system will present a hint on screen, namely a graphic or sound. The child is then requested to select and show a card either with the respective category or with an item in the same category. Positive reinforcements, applied when the answer is correct, are configurable, as well as the actions to be triggered for showing the wrong card.

In the Phrase template the child is requested to show a sequence of cards, which represent the subject, verb and noun in a phrase, see Figure 2 for an example using symbols from the ARASAAC collection (available at
When each card is placed on the table a reinforcement can be triggered by the system. This reinforcement will hint the child if the right card is being used, and also if the card is in the correct position within the phrase. The system will detect when a phrase is complete and trigger the user-defined positive reinforcement.

Figure 2: An example of a phrase.

5 TESTING THE PROTOTYPE

The test was performed in an association that supports children with ASD, the Associação para a Inclusão e Apoio ao Autista, located in Braga, Portugal (http://www.aia.org.pt/).

The prototype was tested by a speech-language therapist, the third author, that performs interventions regularly with children, and four children diagnosed with ASD. The goal was to test if the software was clear to the therapist, and to evaluate its usage as a supporting tool in the interventions performed.

The therapist had 13 sessions with four boys diagnosed with ASD. The children were between the ages of 6 and 10, had some level of orality, and attended school.

Some children were previously assessed as participative in the speech-language sessions, while others have an extremely passive behaviour and barely engaged in spontaneous interaction and communication with the therapist. One child had several verbal and motor stereotypes and expressed himself mainly through echolalia.

A qualitative study was performed with data gathered through observation by the team, including the therapist, in order to evaluate the suitability of the software usage by the children, namely what behaviours were triggered by the usage of AR, and if the children had benefited from the interaction with the software. To support the study we recorded the sessions in video, and filled observation grids with the interaction stages as well as observations that emerged in the sessions with the children (Figure 3).

5.1 Strategies

Based on the previous history with the children, as well as the existing intervention plan for each child, the therapist set the following scripts for these sessions.

For three children the goal was set to: (1) identify animals, nutrients, and daily objects; (2) Identify and designate gender; (3) Identify and designate actions; (4) Discriminate and identify sounds; (5) Answer yes/no questions, together with the respective head nod or shake; and (6) Build sentences with subject, verb and noun.

The other child had as goals to: (1) Enhance declarative skills; (2) Improve language skills; and (3)
Create and imagine using language.

The therapist created six activities based on the four available templates to achieve these goals.

5.2 Evaluation

In general the children were highly motivated by the usage of AR. Superimposed 3D models and computer animations on top of cards displaying familiar symbols caused a higher degree of engagement in the activities, and had the children requesting the therapist to 'play' with the software when arriving.

The child with stereotypies, when in traditional interventions, used to keep asking to go to the window to watch for cars passing by, and kept repeating out of context phrases. After the AR based session had ended, when they were ready to start a different activity (not related to this project), the child asked the therapist for the dog marker and autonomously presented it to the system. A joyfully expression emerged when the dog’s bark was heard.

A child defined by the therapist as passive, regarding communication and behaviour, revealed to be highly motivated, repeating everything the computer 'said'.

The activities were easily grasped by the children, possibly due to the use of tangible interfaces based on the cards with symbols, something which was highly familiar to these children. The prototype did not require any new form of communication, showing cards was something that was already present in the sessions prior to this test. Hence, the interaction with the computer presented itself as familiar to the child.

After 13 sessions the therapist reported her impressions. On a positive note, the therapist stated that the children kept the interest on the activities for a longer period of time, and inclusively that children had the initiative to ask and use the software on their own. The children acquired new vocabulary and consistently managed to perform correctly the proposed activities.

However, the therapist also noted that a particular child would sometimes disregard everything else concentrating only on his image on the screen. The joint attention was not as high and eye contact occurred less often, suggesting that a balance must be achieved between the several different types of interventions and activities.

The therapist concluded that an AR based approach has the potential to present benefits in speech-language therapy although the social area may be neglected with some children.

5.3 Detected Issues

Some issues came up during the test which are worth mentioning in this report. First we had some situations where a card would be incorrectly recognized, triggering an inappropriate action. Other times an action was triggered without having any card on display. These issues came up due to the variations in lighting during the day.

One child actually enjoyed the fact that the system made 'mistakes' and was eager to correct these mistakes. The therapist also took these system misclassifications and explored them within the intervention leading to an even more engaging session. These unintended recognition failures must be fixed, as children with more rigid cognitive processes may be unable to deal with these situations. Nevertheless, this obviously leads to the question: should the system fail occasionally on purpose? Should this feature be introduced in the configuration?

On another session when the system triggered a sound without a card being shown, the child immediately sought for the associated card and showed it to the system. Again, while these unattended failures must be avoided, they may become part of an activity with some children, introducing an element of surprise and breaking the pattern of the activity.

The team also noted that the Phrase template needed to be redesigned to include variations and more reinforcements. The system should be able to provide hints regarding the type of the missing elements and their relative position. Furthermore, the system should allow reinforcements when, although all the correct cards are in the table, the order is not correct. Another option to explore is to start with incomplete phrases and ask the child to identify the type of the missing element, as well as the specific missing card.

The last issue is related to the child’s interaction when motor stereotypies are present. In our study a child needed assistance from the therapist to be able to show the cards. This issue requires support for a different form of interaction.

6 CONCLUSIONS

The present work has proposed an approach, relating elements from AAC and ABA, to apply AR in communication and language interventions with children with ASD to mitigate the issues caused by the impairments in communication skills and imagination.

The resulting prototype has been qualitatively evaluated in an association that supports children with
ASD and has therapists conducting interventions in many areas. The sessions were conducted by a therapist and the behaviours of the children and their interaction with the system was recorded on video and using observation grids.

A common issue in qualitatively evaluations relates to the generalization of results to the remaining population. We have no intention of suggesting that the data presented in here can be generalized. This is particularly relevant in the case of ASD due to the wide range of disorders and their varying degrees. This study presents only a contribution to the development of technological systems to support interventions.

Our study shows that the usage of information technologies, in particular AR based systems, in speech-language interventions, complementing and supporting the traditional approaches is an option that must be further explored. The results suggest that the system usage could represent an added value since we could observe greater motivation, word acquisitions, and commitment in performing correctly, as observed by the therapist.

Due to the use of AR and tangible interfaces, familiar physical objects, and an interaction procedure similar to the one in the previous AAC sessions, where cards are exchanged between child and therapist, the system was easily grasped by the children. As expected, the children were able to use their past experience to interact effortlessly with the system and hopefully they will be able to generalize the acquired knowledge.

We believe that the positive results obtained in this study are, to some extent, due to the engaging environment provided by the application, and its tangible interface. These factors stimulate cognitive processing and lead to visual learning.

These results illustrate the potential of AR systems to be used in linguistic skill development due to the significant increase in interaction and communication initiated by the children.

While technologies such as the ones in our prototype can provide very positive results, we must pay particular attention to avoid having the child completely immersed on some application details and neglecting social interaction. The therapist plays a fundamental role in these augmented ‘interventions’.

Game oriented activity templates could provide an extra level of engagement for the children. Adding templates should be explored to increase the range of possibilities for the therapist to conduct the interventions. The phrase template, as mentioned before, would also benefit from a subdivision or at least more configuration options.

The card recognition subsystem also needs to be fine-tuned to increase robustness, hence suppressing unintentional card misclassification. These system errors, due to the way children reacted to them, raise an interesting question that we believe is worth pursuing: Would a system that intentional provides a wrong answer be a way of further engaging at least some children? What is a ‘good’ wrong answer?

Children with motor stereotypies also require a different interaction procedure. This condition makes it harder for the system to correctly classify the cards. This is an important issue since a significant number of children with ASD suffer from motor stereotypes.

In this study we focused on speech-language interventions. We would like to explore other types of interventions, such as behavioural and sensory therapies.

Finally, the limited number of participating children does not reflect the broad range of the spectrum, nor do they allow for any generalization in a particular subset of these disorders. While the generalization for the whole spectrum of disorders is an impossible task, more participating children would allow for a more relevant analysis, regarding the significance of the contribution of a system such as the one presented in here when used in interventions.

**ACKNOWLEDGEMENTS**

Work partially funded by National Funds through FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project PEst-OE/EEI/UI0752/2011 and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-Brazilian National Council for Scientific and Technological Development) within program Ciência sem Fronteiras.

**REFERENCES**


