Designing a virtual teacher for non-verbal children with autism: Pedagogical affordances and the influence of teacher voice.

Paul Herring, Kieron Sheehy, Roger Jones and Karen Kear.

The Open University, U.K.

1. Introduction

Children with autism demonstrate deficits in language and communication, social interaction, thought and behaviour. Various pedagogical approaches are commonly used to support communication development, one of which is the Picture Exchange Communication System (PECS) (Preston & Carter, 2009). PECS is a simple, physical symbol-based communication method used by teachers to develop communication and social interaction in children with autism (see figure 1, below). Such children are often diagnosed pre-school, but are rarely introduced to PECS until after they begin school at around four to five years of age. This delays language and social development with accompanying behavioural problems associated with the condition.

Figure 1: a PECS symbol exchange. Herring (2009)

Our research is developing a computer-based PECS approach involving a virtual teacher. This blending of traditional approaches with those facilitated by computers must put the children's needs first, rather than being primarily technological test-beds (Sheehy, 2010). If successful this might provide earlier and wider access to language support systems. In this paper we consider one particular element of this approach, the voice of the virtual teacher and its influence on a non-verbal child's learning interactions.

2. Computer Assisted Picture Exchange (CAPE)

The Computer Assisted Picture Exchange (CAPE) uses information and communication technologies to deliver PECS pedagogy in a virtual or mixed virtual/real environment. It uses a software toolkit originating at the Center of Spoken Language and Understanding (CSLU) at the University of Oregon. This provides 'virtual teacher' avatars with synthesised verbal utterances with associated lip movements or human utterances using pre-recorded tracks. Functionality has been extended in the CAPE system to include short animations of a 'Super Monkey' character. This is intended to interest and motivate the children interacting with the system (see figure 2, below).

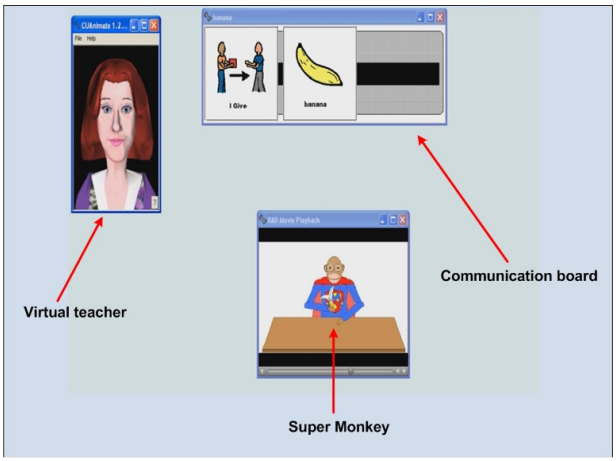

Figure 2: The CAPE interface

CAPE is also equipped with RFID-enabled physical symbols similar to those used in PECS. RFID enabled symbols used in conjunction with an RFID reader (see figure 3) allows CAPE to simulate the physical exchange of symbol tokens used in PECS pedagogy (see figure 3, below). In response to prompts from the virtual teacher, the child chooses and places symbol tokens on an interface board and CAPE responds to this selection.

Figure 3: RFID reader and RFID enabled symbols

A key design issue may be that of avatar voice selection as children with autism find some voices 'too synthetic' (Williams et al. 2002). This implies a tension between the children's preference for non-human teacher avatar, and a possible dislike of a non-human voice.

To investigate this issue a pilot study compared how children with autism reacted to tasks set by a virtual teacher using either a human voice or a synthesised voice. The interactions studied were centred on three tasks set for each participant (Table 1.)

Table 1. Key tasks in pilot study

Task number	Description
1	The virtual teacher asks the participant to select a symbol to indicate which of three foods and one drink, displayed on screen, Super Monkey should eat or drink. A symbol is placed on the RFID reader by the participant to indicate their choice, initiating a PECS communication exchange between the participant and the virtual teacher. Selection of an appropriate symbol is acknowledged by the virtual teacher, giving positive reinforcement. A representation of the chosen symbol is shown (visual cue), followed by an animation of Super Monkey, for example, eating the chosen food item. This provides reinforcement and an associated visual reward.
2	After each symbol interaction the participant is asked to remove the symbol from the RFID reader and put it back into their folder. This task does not have an explicit visual cue (other than the virtual teacher's talking head). This allows any differences in participant responses between tasks having visual and aural-only cues to be noted.
3	The virtual teacher asks the participant whether he or she would like to select another symbol. A symbol choice of Yes and No is shown on screen, (visual cue). The participant chooses Yes to continue or No to finish.

3. Method

Eight participants between seven and ten years of age took part in the study. Each participant had a formal diagnosis of autism and was able to use a maximum of one word utterances.

The research team consisted of a researcher and two experienced Special Educational Needs (SEN) teachers of the participating children. For continuity and control against participant anxiety a teacher was assigned to each cohort. During each CAPE session the cohort teacher provided one-to-one support for the participant and the other teacher observed the session.

A related two condition, counterbalanced design was used to reduce order effects. The participants were split into two cohorts, of four children. Each group took part in two, ten minute teaching sessions spaced over four days. One group's first experience of CAPE was mediated by the virtual teacher speaking with a female synthetic voice. The other group's first experience was mediated by the virtual teacher having a natural, female, recorded voice. In the second session the voice type was switched for each group. In this way any preferences that might be order dependent (for example, being associated with increasing familiarity with the system) were more likely to be counter-balanced.

Sessions were video recorded and observed by the researcher and a member of the school teaching staff. Observational data was coded and quantitatively analysed with regard to:

1. How many symbols were selected appropriately
2. The time spent looking at the computer screen and looking at the physical symbols (see figure 3 above).
3. The time spent looking at the teacher in the room.
4. The average speed of response to the virtual teacher's requests.
5. The average speed of response to verbal requests from the virtual teacher with and without a visual cue.

In addition a qualitative analysis considered how voice type influenced participant levels of engagement with CAPE.

4. Results and discussion

Seven participants found CAPE engaging. Participant attention and engagement was elicited by a verbal greeting of "Hello" from the virtual teacher followed by a short greeting and animation by Super Monkey. This elicited smiles in five participants and laughing and hand clapping by one child. Another child resisted engagement in his first session (averting his eyes when the virtual teacher spoke). In the second session however, he did engage with the system, with support from the human teacher.

Participants found different aspects of the system engaging. Five of the participants appeared to find Super Monkey more interesting, whilst two preferred the virtual teacher. One participant in particular found the virtual teacher so fascinating that he spent much time in both sessions gazing at, and touching her face when she stopped talking, possibly in an attempt to prompt continuing activity.

The process of symbol selection was initially demonstrated by the cohort teacher. In the first sessions teacher prompts were required to encourage participant symbol selection. However, by the second session, six children were able to complete tasks without this.

Preliminary statistical analysis found significant differences between responses to visual\auditory and auditory-only cues, suggesting that the children benefited most from visually supported, rather than auditory-only, stimuli.

After each session the support teacher and the observing teacher were interviewed. Both were surprised to observe that synthetic voice audio cues appeared to be the most effective. The teachers also noted that six students were beginning to use the 'Yes' symbol during the two brief sessions and three of the children also correctly selected 'No' when they wished to finish. The learning of the concept, and correct use of 'Yes' and 'No', is a significant difficulty for children with autism.

5. Conclusion

The RFID token system provided an effective interface, allowing participants to engage with the activities. CAPE elicited high levels of engagement, suggesting that this approach could be developed to teach symbol communication.

Natural voice seemed to produce lower levels of correct responses than synthetic voice, but could be effectively supported if verbal requests were combined with visual cues.

Future CAPE systems will aim to enhance PECS pedagogical approaches. CAPE will do this by combining synthetic voice with visual cueing that will be gradually phased out as the participants' PECS skills improve.

6. References

Bignell, s. (2008) Social communication skills of people diagnosed with autism in 3D multi- user virtual worlds. *ReLIVE08*, 20-21 November 2008, Milton Keynes, England.

Preston, D. & Carter, M. (2009) A Review of the Efficacy of the Picture Exchange Communication System Intervention. *Journal of Autism and Developmental Disorders* 39, 10: 1471-1486

Sheehy, K. (2010). Inclusive education and virtual worlds: The Teacher Embodiment And LEarning Affordance Framework (TEALEAF). In *Virtual Worlds: Controversies at the Frontier of Education.* Nova Science Publishers

Williams, C. Wright, B. Callaghan, G. Coughlan, B. (2002). 'Do children with autism learn to read more readily by computer assisted instruction or traditional book methods? A pilot study'. *Autism* 6(1): 71-91.