Virtual peers as partners in storytelling and literacy learning

K. Ryokai, C. Vaucelle & J. Cassell

MIT Media Laboratory

Abstract Literacy learning — learning how to read and write — begins long before children enter school. One of the key skills to reading and writing is the ability to represent thoughts symbolically and share them in language with an audience who may not necessarily share the same temporal and spatial context. Children learn and practice these important language skills everyday, telling stories with the peers and adults around them. In particular, storytelling in the context of peer collaboration provides a key environment for children to learn language skills important for literacy. In light of this, an embodied conversational agent, Sam, who tells stories collaboratively with children was designed. Sam looks like a peer for pre-school children, but tells stories in a developmentally advanced way, modelling narrative skills important for literacy. Results demonstrated that children who played with the virtual peer told stories that more closely resembled the virtual peer’s linguistically advanced stories: using more quoted speech and temporal and spatial expressions. In addition, children listened to Sam’s stories carefully, assisting her and suggesting improvements. The potential benefits of having technology play a social role in young children’s literacy learning is discussed.

Keywords: Collaboration; Empirical; Literacy; Pre-school; Storytelling; Virtual peer

Introduction

As new technologies are increasingly present in classrooms, efforts are being made to prepare children for computer literacy. Yet, the traditional literacy skills — the ability to read and write — remain critical for children’s academic success and may also be aided by advances in technology and research. Young children’s acquisition of skills leading to literacy begins with everyday interactions in informal settings with both adults and peers, and is not isolated to formal, academic environments. Whether it is for an adult or a child peer, constructing language for someone encourages children to practice many skills essential to later literacy.

This research addressed the specific discourse genre of storytelling as a bridge to literacy. Storytelling occurs in the context of peer play and while a fun activity for children, it also involves the kind of linguistic activities that can bridge children’s
competence and knowledge of oral language with that of written language. A novel approach to supporting children’s literacy learning is presented and discussed. Technology plays a social role, as a listener of children’s stories, thus providing opportunities for children to foster linguistic expressions in an oral mode that are useful for their later literacy skills.

Storytelling and literacy

This research on literacy learning and storytelling is based on the theory of emergent literacy (Teal & Sulzby, 1986). In the emergent literacy view, aspects of language — both oral and written — develop concurrently rather than sequentially (Goodman, 1986). According to this view, literacy learning does not happen only in formal classroom settings, but also in informal settings, in both oral and written modes, and in collaboration and interaction with others.

Whitehurst & Lonigan (1998) distinguish between the ‘inside-out’ and ‘outside-in’ skills of literacy. Inside-out skills are concerned with children’s phonological and syntactic awareness, and grapheme-phoneme correspondence, thus facilitating children’s ability to decode information within a sentence. Outside-in skills are concerned with children’s ability to take the meaning of a sentence from the context in which the sentence is placed (e.g. understanding who ‘she’ refers in a phrase, ‘then she ate the poisoned apple’). Children must bring their knowledge about the world and apply that to the text. These outside-in skills of literacy —children’s knowledge about language and how it works in a given context — are what drives this research about the kinds of language activities important for the transition between pre-school to school.

Young children’s language is initially limited to concrete here-and-now talk. Early words rely on physically present objects and scaffolding from a familiar conversational partner with whom the child can assume shared knowledge (Ninio & Bruner, 1978; Nelson, 1996). Thus, the acquisition of outside in skills, which requires gaining independence from physical and temporal context, marks a significant transition in a child’s literacy development.

Snow (1983) introduced the term ‘decontextualised language’ to refer to language that is not bound to spatial or historical context. Storytelling, then, provides an ideal forum for children to practice decontextualised language since it avoids any laborious writing tasks. Rather than concrete ‘here-and-now’ talk, storytelling encourages the use of ‘then-and-there’ language (Scarlet & Wolf, 1979). In order to tell a comprehensible story, children must be able to hold the audience’s perspective in mind and reconstruct the original context (Cameron & Wang, 1999). Children learn these skills through interaction with both adults and peers.

Learning with adults

Vygotsky defined the zone of proximal development as ‘the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers’ (Vygotsky, 1978; p. 86). According to this theory, a child performs at a higher developmental level of abstraction and performance with a knowledgeable and skilled partner than she would achieve independently.
Adults serve as the competent partner in emergent literacy activities to support children’s literacy learning. With parents, teachers, children engage in many different kinds of conversations together exchanging information, disciplining, socialising, showing feelings. Within those various types of conversations, children are given opportunities for syntactic planning, careful lexical selection, making explicit cross-utterance relationships and integrating successive utterances into a particular structure (Nelson, 1996). For lexical selection, the use of rare words during parent-child book reading has been shown to correlate with children’s vocabulary acquisition (Snow, 1993). Dickinson et al. (1993) found that pre-school teachers’ use of rare words during mealtime and in free-play settings were positively correlated with story understanding and definitional abilities (e.g. a cat is a kind of animal) in addition to vocabulary growth.

Learning with peers

In contrast to adults, a learning companion can be characterised as someone of a similar developmental age who understands the world in similar ways (Damon & Phelps, 1989). For this reason, a peer partner does not bring the sophisticated strategies and knowledge that an adult partner would. Learning with peers is more heuristic than rule-oriented. Children working with peers may settle for an ungrammatical use of language or may not come to a solution or conclusion simply because they forget to do so (Neuman & Roskos, 1991).

Despite these drawbacks, peers offer a unique learning opportunity for children that adults may not. While parents and teachers may not always be available to listen to children’s everyday stories, peers are available and also scaffold their equal-status partners. In both naturalistic kindergarten environments (Paley, 1984) and experimental conditions (Pellegrini, 1987; Morrow & Rand, 1991; Neuman & Roskos, 1993) children engage in instructional conversation with their peers — designating, negotiating, and coaching each other’s literacy activities. In a mixed-age (5 to 8 years old), K-2 classroom, both younger and older children engage in modelling, assisting, directing, tutoring, negotiating, affirming, and contradicting each other in literacy activities (Stone & Christie, 1996). These behaviours indicate not only that children can provide scaffolding to their equal-status partner, but also that peer interactions offer a forum for the participants to explain and clarify what they mean. According to Daiute et al. (1993), through these explanations, children linguistically externalise their thoughts and ideas for their peers, which is key to producing audience-based language. Thus, the social nature of the interaction around literacy learning is just as important as the absolute expertise of any partner.

Related work

New technologies have been developed and proven to improve both reading fluency and other linguistic skills important for literacy. Mostow et al. (1994) used state-of-the-art speech recognition technology to develop a reading tutor that gave appropriate feedback for children reading storybooks out loud. The reading tutor was found to increase oral reading fluency in children significantly. In contrast to Mostow’s intelligent tutor approach, the Cognition and Technology Group at Vanderbilt used a situated learning approach in developing their Young Children’s Literacy series (CTGV, 1996). A series of animated video stories challenged
children to write a story to save the animals they saw in the video. Interaction with both the teacher and peers was key to literacy learning, as the teacher modelled the story writing activity for the children, and children worked together as a group. The series has resulted in significant improvements in children’s word and sentence fluency and story complexity.

Previous work at MIT has established the paradigm of the Story Listening System (Cassell, 2001). One of these systems, the StoryMat (Ryokai & Cassell, 1999) recorded children’s oral stories and the movements of stuffed animals made on a technologically enhanced play mat. When another child played with a mat, the stories were played back as animations – echoes of the previous playmate. Results demonstrated that interacting with peer stories on StoryMat led children to tell more imaginative and structurally advanced stories. Another SLS, TellTale (Ananny, 2001) invited children to record segments of a story into the body parts of a plastic toy caterpillar. After a short period of play, including deciding how to arrange and segment story sequences, children exhibited more sophisticated use of discourse connectives (e.g. ‘and’, ‘then’, ‘because’) and story event language. These previous systems raise the question of the role a virtual partner’s feedback might have on children’s stories. Could children’s literacy skills be fostered by incorporating a kind of virtual companion who could be a listener of children’s stories?

Chan & Baskin (1990) created ‘learning companion systems’ that employed a set of agents — one as an intelligent tutor and the other as an artificial student that was designed to be at about the same level as the student (both agents were non-embodied). The idea was that a student would learn from an intelligent tutor (in regards to programming LISP), but then was asked to teach the artificial student (learning companion) what he learned. By having the two tasks — learning by being tutored and tutoring, these learning companion systems offer a learning protocol that is similar to ‘reciprocal teaching’ (Palincsar & Brown, 1984) in which children take both the teacher’s and learner’s role. While their preliminary results did not show significant improvements on problem solving tests, their interviews revealed that the students enjoyed teaching an agent over a real student because they felt it was like a game.

Technology to provide opportunities for children to learn by teaching others was explored further by Brophy and colleagues in the Teachable Agent project (Brophy et al., 1999). In their work, children learned ecology by teaching it to a naïve cartoon character. Brophy and colleagues found that children who studied in order to teach the agent did better on the post-test than control children who studied just for the subject test, as the students who prepared to teach spent time trying to understand ‘the why’ of the studies.

There seems to be an advantage in making technology play a more social role in supporting children’s learning. In literacy learning, such social interactions are important as they provide opportunities for children to gain knowledge about language and communication, and also to test their knowledge of language.

**Sam**

Sam was created to give technology a social role in supporting young children’s literacy learning (Cassell, 2001). The Sam system has two parts: the character Sam, an embodied conversational agent who is designed to look like a child around age 6, and a toy castle with a figurine. Sam’s androgynous appearance (and accordingly, the
name, *Sam*) was chosen intentionally so that both girls and boys could relate to *Sam*. For the sake of simplicity, in this paper, *Sam* is considered to be female.

*Sam* is projected on a screen behind the castle, and can both listen to a child’s stories and tell her own. The figurine can exist in either the physical world or on the screen, so that *Sam* and the child can pass it back and forth between their worlds (Cassell *et al.*, 2000). When a child arrives in front of the toy castle, *Sam* looks at the child and says, ‘Hi, I’m *Sam*!’ After the child greets *Sam*, *Sam* tells a story as she moves the figurine around the castle, occasionally looking up to draw the child in to the story. When *Sam* finishes her story, she says, ‘I’ll put the toy in the magic tower so you can tell a story’, and places the figurine inside the tower. When the child opens the door, she finds the figurine *Sam* had been playing with and tells her story. While the child does so, *Sam* watches the child (following where the child is moving the figurine with head and eye movements), nodding, smiling, and prompting, ‘And then what happens?’ When the child is done, she places the figurine back in the castle where *Sam* can access it.

*Sam* then starts her story in the same part of the magic castle where the child finished hers. *Sam* tells stories using more advanced forms of linguistic expressions (quoted speech, and enough temporal and spatial information for the audience to be able to reconstruct the story). In Vygotsky’s terms, children learn through their participation in activities that are slightly beyond their competence, with the assistance of adults or more skilled children. Thus, by interacting with a peer who tells stories in a developmentally more advanced form than the child, the child may enter his/her ‘zone of proximal development’ (Vygotsky, 1978). The hypothesis is that by interacting with a slightly more advanced peer, children model *Sam*’s linguistic behaviour and therefore, perform their storytelling task in a more mature form themselves. However, in addition, *Sam*’s young appearance and playful environment (with the toy castle) may invite children to critique *Sam*’s behaviour, giving them an opportunity to externalise their thoughts and communicate their points using language. The intention is for *Sam* to provide just the right amount of challenge. *Sam*’s storytelling is more advanced than the child’s, but not too advanced, as she is a partner who is just a head taller than the child.

**Technical implementation**

The *Sam* system detects a child’s presence through a microphone and a motion detector sensor in front of the castle. When the child is playing with the toys and narrating, the system uses audio threshold detection to determine when to give feedback (backchannels such as ‘uh-huh’ nods, and explicit prompts such as ‘and then what happens?’). Radio frequency tag readers are embedded inside of every room in the castle. The tag attached to the figurine tells the system in which room in the castle the figurine is at the moment. A switch in the door tells the system whether the figurine is inside the magic tower and when the magic tower door is opened, so that the child will never see the physical and virtual instantiations of the toy.
simultaneously (when the door is opened and Sam has the figurine, it disappears instantly and Sam expresses surprise). In order to make Sam’s character believable, Sam’s stories and other utterances were recorded from a real child, as the quality of children’s synthesised voices is still poor. The software is written in Java and C++ and can run on a single PC with a graphics acceleration card. The animation is displayed on a back-projection screen behind the castle.

Method

Participants
Thirty-one children (all five-year-old girls) volunteered for the study. Three children indicated that they did not wish to participate during the study. Thus, the final sample consists of 28 five-year-old girls.

Procedure
In a 2 × 2 design, eight children played alone with a castle without Sam, eight children played alone with a castle with Sam, six children played with a copresent playmate with a castle but without Sam, and six children played with co-present playmate with a castle and with Sam.

The study was done in a ‘Wizard of Oz’ setting where Sam’s response was controlled by a researcher behind the screen. All children played for about 15 minutes: 5 minutes introduction with an experimenter, and a 10 minute play session on their own. All the children’s 10-minute play sessions were fully transcribed.

Results

Sam as a capable partner
The following is an example of a child interacting with Sam as a capable partner:

Ann (age 5) is telling stories with Sam by herself

SAM: OK. Let me start. Today I’m going to ride horses in the meadow. My parents said I could ride the big horse named Star. Oh, no. Star has been stolen. I better go tell the sheriff. ‘Oh, sheriff, my favourite horse Star has been stolen and I don’t know where she is.’ ‘Oh, no. No need to worry. A kind old lady from the other side of the forest has found her, and she is just coming back home.’ ‘Yippee! Thanks. Come on, Star.

ANN: Once upon a time there was a little girl and she went downstairs. She looked in the magic mirror. She went downstairs and looked in the mirror and turned on the lights, and then went back up the stairs. and she looked at her magic, and she looked in the magic mirror, then went back downstairs, and there was her mom and dad. The end. Your turn to tell the story.

SAM: Cool! OK, let me start.

[…] I’ll put the toy in the magic tower so you can tell a story.
ANN: Once upon a time, there was a little girl who wanted her mommy and poppy, but she didn’t have one so she could do anything she wanted. She hopped downstairs and then she saw the, she went upstairs and told the magic mirror that she wanted a mama and papa. The magic mirror told her that she couldn’t get one. So she went back downstairs and she saw a monster. She went back upstairs to the magic mirror and said, ‘Magic mirror, why there’s a monster?’ and she went back downstairs and there wasn’t a monster anymore. The end!

Ann took turns with Sam, listened to Sam’s stories, and in that process, seemed to tell her stories with increased complexity. In her first turn, Ann’s sentences were fairly simple. Her speech may be categorised as an eventcast (i.e. the form of ‘then she went there, and then she went there . . .’) rather than a story with a causal connection between clauses (Labov, 1972).

Sam’s stories were designed to involve complicating actions (e.g. losing a horse) and resolution of stories (e.g. finding the horse). They also modelled decontextualised language, such as quoted speech (e.g. ‘Oh, sheriff . . .’), temporal expressions (e.g. today I’m going to . . .), and spatial expressions (e.g. a kind old lady from the other side of the forest), and relative clauses (e.g. the big horse named Star) that help the audience reconstruct the event. In the example above, hearing Sam’s stories seemed to encourage Ann to use such decontextualised language (e.g. ‘a little girl who wanted her mommy and poppy’) and quoted speech (e.g. ‘she said, ‘Magic mirror . . .’).

A team of two researchers coded the occurrence of spatial expressions, temporal expressions, and quoted speech in the children’s stories. Following Peterson et al. (1999), a spatial expression was coded as definite information about where the event took place (e.g. ‘then the boy went to the kitchen’) and temporal expression as explicit information about when the event took place (e.g. ‘he went downstairs when he heard the noise’). The quoted speech was coded for both direct speech with a framing clause (e.g. then she said, ‘Oh no!’) and indirect speech such as ‘he said that he wasn’t hungry’ (Hickmann, 1993). The occurrences were tallied, and then normalised with respect to the child’s total storytelling time.

Sam’s presence as a storytelling partner dramatically increased the frequency with which children used quoted speech and temporal and spatial expressions. Figure 3 shows the mean frequency (tally of occurrences of expressions by each child/total time that child spent speaking) of spatial expression across the four conditions. Thus, for the dyads, the bar represents the mean frequency for each of the children in dyads. A full-factorial ANOVA revealed a main effect due to the presence or absence of Sam, $F_{3,24} = 68.04, p < 0.01$. There was no main effect for number of children (the one child vs. the dyad condition), nor were there any interactions. Children used significantly more spatial expressions when playing with Sam than they did alone, or with another child. Findings were equally significant for quoted speech ($F_{3,24} = 10.58, p < 0.01$) and temporal expressions ($F_{3,24} = 30.52, p < 0.01$). The children in the ‘dyad with Sam’ condition had equally high frequencies of quoted speech and temporal and spatial expressions as in the ‘one child with Sam’ condition.
condition. This suggests that Sam succeeds in evoking decontextualised language even in the presence of a real flesh-and-blood playmate.

Were the children’s uses of literate expressions attributable to the fact that Sam modelled these behaviours? This was examined by looking at whether the literate expressions increased over the course of the interaction with Sam. Remember that as the children took turns with Sam, every one of their stories was preceded and followed by a story by Sam. Figure 4 illustrates the mean number of spatial expressions per story produced by the children in the ‘one child with Sam’ condition. The figure illustrates the increased amount of spatial expressions as the children tell their stories with Sam. The Pearson product-moment correlation test revealed a significant positive correlation between the chronology of stories and occurrence of spatial expression, \( r = 0.35, \ p < 0.05 \), and of quoted speech \( r = 0.27, \ p < 0.06 \). No significant correlation was found for temporal expressions \( r = 0.065 \). However, if one looks only at the first three stories, the use of temporal expressions increases significantly successively. This suggests that children may have become tired after the third interaction and were no longer able to push their linguistic behaviour to its limits.

The result suggests that Sam did succeed at eliciting more literate language from children over time. However, the duration of the study is not sufficient to conclude that the children actually learned these behaviours from Sam. Perhaps, the children already had the ability to use decontextualised language, but did not necessarily know when or why to use it. In that sense, may have helped the children perform to the best of their ability. By telling stories in a developmentally advanced way, Sam may have modelled the use of literate expressions and provided an opportunity for the children to practice them. A future study will investigate children’s interactions with Sam over a longer term in order to determine how repeated interactions with Sam affect children’s development.

**Conversation vs. storytelling**

Pairs of children who played with the castle without Sam treated each other more as conversational partners rather than taking turns being the storyteller and the story listener. In the example below, the two children engage in pretend play and move seamlessly between talking to one another as characters in a shared story, and as children in shared play:

**Wendy and Sarah (both age 5) are playing without Sam**

Wendy: You broke this after I had fixed it.
Sarah: Not me.
Wendy: It’s probably the ghost.
Sarah: There’s no such thing as monsters. Did that door just open, or was it just my imagination?
Wendy: It was just your imagination.

While the two children are engaged in a conversation, instead of storytelling, their speech is more dependent upon contextual cues. For example, the child did not introduce or explain what ‘this’ was in the utterance ‘You broke this . . .’ because the referred item was immediately shared with her partner and in their conversation. The children who played with Sam also shared the physical context with Sam (e.g.
sharing the castle). However, Sam explicitly invited the children to tell stories and modelled decontextualised storytelling behaviour. Further, perhaps because Sam’s method of narration did not rely on contextual cues, the children’s narration also became less context-dependent. In a way, the children and Sam shared the same invisible audience. Therefore, Sam’s presence as a partner who took turns with children and told stories using diverse linguistic expressions appears to have been important in making the stories more sophisticated, fostering children’s use of linguistic expressions in storytelling.

**Sam as a storytelling partner**

The children regarded Sam as a storytelling partner. This was evident both from how the children took turns with Sam and from various comments directed at Sam. Many children acknowledged Sam’s turn by giving ‘Your turn!’ acknowledgement. When things were not clear, as in the following example, children seemed to ‘ask’ Sam questions as if to check if Sam was OK:

Simone (age 5) is playing alone with Sam.

**SAM:** Cool! OK, my turn. Today I’m going to ride horses in the meadow. […]

**SIMONE:** Thanks. Come on, Star. [pause]

**SAM:** I’ll put the toy in the magic tower so you can tell a story.

**SIMONE:** What should I tell, Sam? Do you have an idea? [gaze Sam] Hmmm.

**SAM:** Tell me what happens next.

**SIMONE:** Oh, the girl was happy. She came back from, her husband was there, she was very happy. Everyone. I mean everyone knew she was a good girl. She always had fun playing with her sisters.

**SAM:** Cool.

Simone talks to Sam in a way that indicates she considers her to be a real storytelling partner. In addition, Simone’s nonverbal behaviours — looking at Sam when she is asking her questions, looking at the toys while narrating, and then looking back at Sam at surprise points in the narrative — also demonstrates her willingness to treat Sam as a peer. Thus the children not only regarded Sam as a fellow narrator, but also treated Sam as if she was a real child. Although there are only preliminary results on eye gaze patterns used by children in the study, observation leads to the belief that children looked back-and-forth from Sam to the castle in similar ways as they did when they were playing with another child. And, in fact, even with a copresent playmate, children seemed to take Sam into account. The following is an example from two children playing with Sam:

Amy and Beth (both age 5) are playing together with Sam. Beth has already told her story. Now Amy is telling her story.

**AMY:** So, then, the mother and father put her bed.

**BETH:** Because she lied?

**AMY:** Because she lied, and because she wasn’t supposed to do that.

**BETH:** OK. My turn.

**AMY:** *Sammy*. I want *Sammy* to do it. I’ll put it back.

[Amy puts the toy in the magic tower for Sam to take her turn]

Even with a co-present playmate, the children seemed to take Sam into account. In everyday storytelling, children become collaborators and facilitators of peer narrations (Preece, 1992). Thinking about Sam’s turn and acknowledging Sam’s role as a fellow collaborator is similar to what children go through with peers in everyday collaborative storytelling. Literacy learning is more profound in situations where
children assist each other or collaboratively engage in activities than it is in parallel or solitary behaviours (Stone & Christie, 1996). In these experiments, Sam seemed to play the role of an engaging peer, and may thus have been able to elicit linguistic behaviours predictive of future literacy.

Teachable Sam

Unlike the Teachable Agent project (Brophy et al., 1999) in which the agent was specifically designed for students to teach, Sam was not explicitly designed to elicit help from children. However, Sam was designed to appear as a child to the children who play with her. In effect, children not only seemed to regard Sam as a storytelling partner to model after, but also as a peer they needed to coach. While interacting with Sam, children spontaneously helped her by making various comments about her stories and behaviours. The following is an example of a child ‘coaching’ Sam:

Jane (age 5) is playing alone with Sam.

JANE: [talking to Sam] Try to make a longer story next time. It’s like this. The little boy was outside. He flipped all around and he went inside, he did a flip […]

In response to Sam’s story, Jane told a relatively long story. After listening to Sam’s short story, Jane criticised and then went on to model what she was looking for.

The following is another example of a child correcting Sam:

Ann (age 5) is playing alone with Sam. Sam tells a story which Ann has heard before. Ann interrupts Sam and says that Sam has already told that story before.

ANN: Sam, you already told that story. You can still tell it though. Go ahead.

[pause]

SAM: I’ll put the toy in the magic tower so you can tell a story.

ANN: OK. Let’s see. [pause]

SAM: Why don’t you tell me a story?

ANN: Just a minute, Sam.

Ann corrected Sam’s storytelling, but did so politely, allowing Sam to finish her story. In everyday storytelling, children become not only collaborators and facilitators, but also active critics and correctors of peer stories (Preece, 1992). Accordingly, Jane and Ann, in the above examples, became critics and correctors of Sam’s storytelling. Sam seemed to act as a co-storyteller, but also a peer the children felt responsible to critic and coach. By coaching, not only do peers provide substantive input to one another’s learning (Cazden, 1988; Rogoff, 1990; Neuman & Roskos, 1991) but also practice verbalising their thoughts (Youngblade & Dunn, 1995). Therefore, children’s interactions with Sam, both as co-storyteller and as critic, may contribute to them becoming critical thinkers who can evaluate and challenge others’ linguistic behaviours while reflecting on their own knowledge.

Sam as a facilitator of peer interactions

Sam does not only play with a single child but also facilitates interactions among dyads of children. Using Stone and Christie’s labels of collaborative behaviours, the number of collaborative behaviours the children exhibited towards their partner during their play session were measured and categorised into two types: story and non-story. Collaborative behaviours labelled as ‘story’ were comments about the ongoing story (e.g. ‘pretend that she was eaten but she escapes’). Collaborative behaviours labelled as ‘non-story’ were about any topic except the on-going story (e.g. ‘I have a toy like this’). Figure 5 summarises the results.
The children in both the 'dyad with Sam' and 'dyad without Sam' groups exhibited collaborative behaviours. However, the nature of the collaboration differed between the two groups. The children in the 'dyad with Sam' group engaged in more ‘story’ collaborations ($m = 7.5$) than ‘non-story’ collaborations ($m = 2$), $t(5) = 3.18$, $p < 0.05$; that is, they more often helped each other by commenting on one another’s stories. The children in the ‘dyad without Sam’ group, on the other hand, engaged in more ‘non-story’ collaborations ($m = 5.67$) than ‘story’ collaborations ($m = 1$), $t(5) = –12.79$, $p < 0.01$; that is, they more often commented to each other on things not related to the target task of storytelling. With Sam, then, the children talked more about storytelling than without Sam. Sam engaged children more fully in collaboration related to storytelling.

Limitations

Sam successfully engaged the children in collaborative storytelling (i.e. taking turns being a listener and a storyteller. However, the primary limitation of the current Sam system arises from Sam’s inability to understand the content of the children’s speech. Since Sam uses silences to determine when it is her turn to speak, when she does not detect a silence, or incorrectly detects a silence, her feedback can appear inappropriate. Likewise, children may understand Sam’s silences as a cue for them to take the turn, and Sam is unable to recover from this and give over the turn gracefully.

In this first study of Sam’s interaction with children, a limited number of research questions were explored. First of all, the study was limited to 5-year-old girls. Would this kind of storytelling play and interaction with Sam and her toy castle engage both girls and boys equally? What age range does this type of storytelling effectively engage? Secondly, the children in the study played with Sam only once and for only 15 minutes. In order to investigate whether children really learn new linguistic skills from interaction with Sam, children’s interactions with Sam need to be prolonged in the short and long-term. The ability to understand the child’s speech and respond more appropriately will allow an extension of children’s interactions with Sam.

Future work

The results of this study with children revealed a number of possible modifications and improvements. Firstly, Sam’s response behaviour can be improved through an investigation of keyword spotting speech recognition technology. In addition to speech input, Sam’s toy castle is being enhanced with more sensors to follow movements children make while they are narrating. For example, movement of furniture in the castle while children tell their story may be cues for Sam to give feedback to their actions.

A recent study has shown that children’s ability to take multiple perspectives in storytelling is positively correlated with their mathematical skills, so Sam’s new stories will also include narrative perspective taking. It seems that Sam could model such a perspective taking by introducing and maintaining different characters in her
stories. To encourage such perspective taking, multiple figurines have now been incorporated so that Sam and children can tell stories with multiple perspectives using the figurines.

To carry out a longitudinal study, Sam’s interactions with children also need to evolve over time. For example, Sam cannot simply greet ‘Hi, I’m Sam!’ every time a child plays with her. How could Sam establish a long-term relationship? A study has shown that friends, compared to nonfriends, resolved more conflicts and performed better at emergent literacy activities during pretend play (Pellegrini et al., 1998). Can Sam be a friend to a child? This will be investigated through the kind of interactions and relationships Sam can have with children over a longer term.

Finally, which aspects of Sam contribute to its success as a literacy learning companion: a linguistic model for the children, or a peer who promotes constructive criticism and perspective taking? For example, does Sam’s child-like appearance make children more comfortable to critique Sam’s stories and behaviours? In order for Sam to produce the positive effect of multiage collaboration where children learn by both modelling and coaching their peer (Christie & Stone, 1999), a more explicit model of a peer who could both teach and be criticised is needed.

Discussion

Sam became a partner for children to model their own stories after, as well as a peer in need of didactic coaching.

Sam’s ability to model and draw children’s attention to linguistic behaviours crucial for literacy. By taking turns with Sam and by listening to Sam’s stories, the children’s stories became more sophisticated and explicit through the use of quoted speech and spatial and temporal expressions. In effect, children practised ways of clearly presenting narrative ideas for an audience, which is one of keys to literacy learning.

Unlike traditional computer-assisted learning, where computers are enlisted to support learning between a teacher and pupils or to support collaborative learning between pupils, this work explored the role of computers as participants in collaborative learning. This work contributes to the field of computer-assisted learning as it illustrates how computers could play a more social role in supporting young children’s literacy learning by both scaffolding and reproducing social learning environments.

Acknowledgements

We thank all the children who participated in the study, the administration of the Thompson School (especially the wonderful Vicky Rose), Linda Lin, Elisabeth Sylvan, Dona Tversky, Ian Gouldstone, and the members of the Gesture and Narrative Language group at the MIT Media Laboratory

References


Virtual peers as partners


