StoryFaces: Children Exploring Emotional Expressions in Storytelling with Video

Kimiko Ryokai

School of Information

Berkeley Center for New Media

University of California Berkeley

Berkeley, CA 94720 USA

kimiko@ischool.berkeley.edu

Robert Kowalski

University of Munich

Amalienstr. 17

80333 Munich

kowalski@cip.ifi.lmu.de

Hayes Raffle

Nokia Research Center Palo Alto

955 Page Mill Road #200

Palo Alto, CA 94304

hayes.raffle@nokia.com

Abstract

We introduce *StoryFaces*, a new composition and storytelling tool for children to explore the role of emotional expressions in children's narrative. StoryFaces invites children to record emotional expressions and then automatically composes these recordings in storybook illustrations. After children watch their faces bring a story to life, they can "go backstage" to play with the story by rearranging the videos and altering the story text. This paper presents our exploratory prototype, a design rationale that focuses on supporting children's emotional growth through storytelling play and reflection, and reports on a formative evaluation with two children ages 4-6. Results from the evaluation suggest that children ages 4-6 are engaged in the activity, are excited to create a variety of emotional expressions, find the narratives funny yet clear, and work to re-craft and reinterpret story meanings through iterative editing and play with both video and textual content. Our goal is to provoke new ideas about how pretend play with digital tools can empower young children in a narrative process.
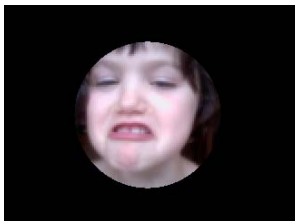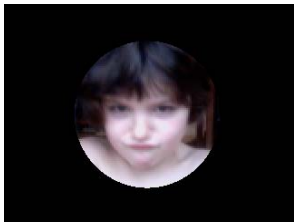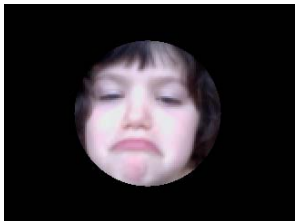
Figure 1: "Paula," age six records different emotional expressions for the story "Expressions with an Alien" in the Video Booth.

Keywords
Storytelling, children, emotional expressions, video recording, communication tools.

ACM Classification Keywords
H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—artificial, augmented, and virtual realities.

General Terms
Design

Introduction
One of the ways in which young children make sense of the world around them is by telling stories [2]. Storytelling not only plays an important role in children's linguistic and literacy skills [10], but it also offers a unique opportunity for their emotional development. When young children tell stories, they pretend that they are going through different scenarios and associated emotions from the perspective of the story characters, e.g., "Then the bunny got lost in the woods, and she got scared…" In storytelling, the children know it is only pretend. Therefore, storytelling is a safe place where children can explore various emotional states without the risk of experiencing the actual emotions [8].

Our goal is to create a tool to support children's natural ability to engage in storytelling and explore emotional expressions in their storytelling play. While many non-technological tools can also have this goal, multimedia tools can easily let children record and replay performances, and software can allow them to arrange and rearrange those performances. Authoring and editing tools have long been lauded for their potential to let children reflect upon their own performances [1] [4][6][11] and our work builds on this trend, investigating how children's emotional play can help them to learn about the roles of emotional expression in narrative structures.

This paper presents a novel prototype to provoke new ideas about the ages and ways in which children can create and edit their own video narratives. We will discuss our prototype design and implementation, a formative evaluation, and directions for future work.

StoryFaces
StoryFaces is a tool for children to capture their facial expressions in a virtual video booth (Figure 1) and create animated stories with their recorded videos and audio. StoryFaces has three parts: 1) The *Video Booth* (Figures 1, 2) where the children record their facial and gestural expressions, 2) the *Story World* (Figure 3) where the children see the animated story with their recorded videos, and 3) the *Backstage* (Figure 4) where they can edit the story with the videos and text.

The Video Booth
In StoryFaces, the children first enter the virtual video booth where they record various emotional expressions. In the video booth, they can record any expressions they wish, but there is an animated character (named "Ema" by the children in our study) who elicits various expressions from the children. For example, Ema says, "Let's make a happy face! Yeah! Happy, happy!" or "Let's pretend that you are sad. Can you make a sad face?"(Figures 1, 2). This is designed to structure the story to follow, but Ema also supports those children who may not know what they want to record when they first arrive at the booth. After each

recording, the recorded video of the children appears at the bottom of the video booth as thumbnails.

The Story World
Once the children have recorded the elicited 3 to 5 facial expressions, they enter the animated story world (Figure 3). In the Story World, an illustrated story incorporates their recorded facial expressions and animates automatically. Each character's face is brought to life with the children's recorded facial expressions and voices. Ema narrates the text of the story as it proceeds so that the young children who do not yet read can understand the story.

The Backstage
Once the storytelling finishes, the children may watch the story again or "go backstage" to edit the story by rearranging the order of the recorded videos and changing the story's text. Backstage reveals each scene in the story like a storyboard (Figure 4) so that children may view and think about the storybook in its entirety. Children can drag the recorded videos from the repository into any of the spaces indicated by the "empty" characters' faces in each scene. The text of the story can be edited directly by clicking on the text in each scene. If the children wish to record additional facial expressions, they may go to back to the video booth where they can record as many additional expressions as they wish.

Design and Implementation
The Stories
StoryFaces currently includes three stories, designed to involve a variety of emotional expressions. For example, "Expressions with an Alien" involves a curious friendly alien asking a child about human facial expressions
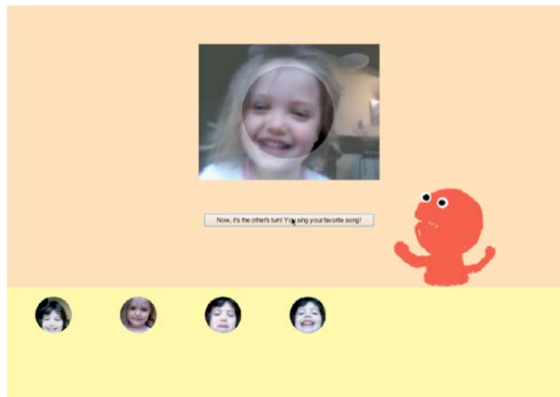

Figure 2: Video booth
Ema, the character asks children for different expressions

Figure 3. The Story World
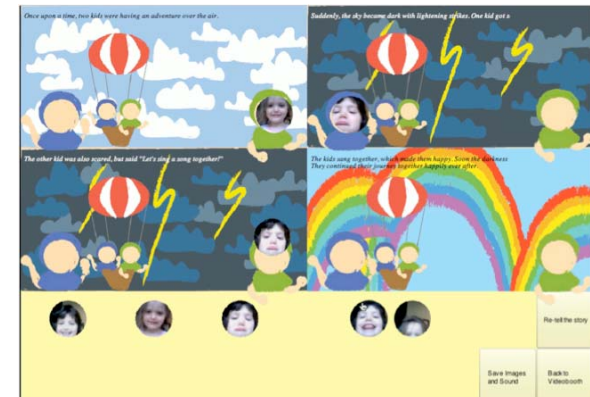showing "In the Balloon" story

Figure 4: Backstage
showing "In the Ballon" story scenes

(Figure 5). In this story, the alien asks, "What do you do when you are happy?" The child says, "We smile." In the next scene, the alien asks, "What do you do when you are sad?" The child says, "We cry." And so on. In the Video Booth for this story, Ema asks the children to make "Happy," "Sad," "Laugh," and "Angry" faces.

"In the Balloon" is a story that involves two children going on an adventure in the hot air balloon (Figure 3). The story characters suddenly encounter a storm! One child gets scared and starts to cry. The other child is also scared but suggests that they sing a song together. In the end, both girls are singing and the storm passes. In preparation for this story, Ema in the video booth asks two children, who take turns recording, to pretend to "Sing," "Smile," and "Cry."

"Walking in the Woods" is an adaptation of the classic fairy tale "Little Red Riding Hood" with the twist that
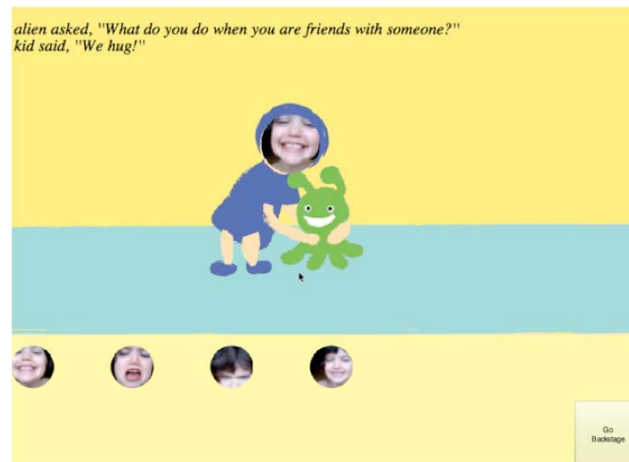
the child unwittingly acts out the expressions both for the little girl and for the trickster Wolf (Figure 6). In preparation for this story, Ema in the video booth asks two children, who take turns recording, to pretend to "Smile," "Be Angry," and "Cast a spell." For "casting a spell," Ema asks the children to swing their arms big in the video booth.

Technical Implementation
The software was written in *Flash* and *ActionScript 3*. Within the video booth, all visual and audible input is processed from the webcam/microphone. To simplify certain processing operations, still frames are captured from the camera at 10 fps. As soon as the recording is finished, all images are masked and aggregated again to one video object, which includes synchronizing the recorded audio stream. During image masking the rectangular webcam images get transformed to circular frame shapes, which are better suited for the

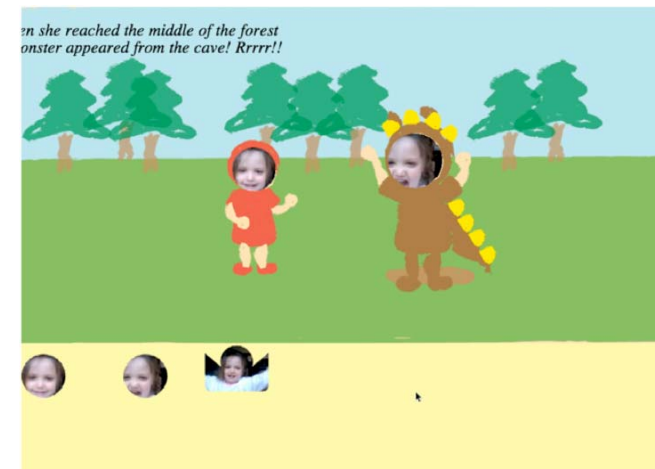Figure 5: "Expressions with an Alien"

Figure 6: "Walking in the Woods"

illustrations that accommodate round faces. A scene handler is responsible for displaying the story text, playing the audible narration, as well as animating children's recorded videos at the right time and place.

In order to edit the story, children drag and drop the video thumbnails from the bottom of the screen to the according characters in the storyboard (Figure 4). To make sure children can identify their recordings, backstage allows them click on a video thumbnail to preview the recording. Children can also alter the story by directly editing its text with standard text-editing functions. Finally, children may also re-record videos by directly moving from backstage to the video booth for re-takes of the original recordings.

Related Work
We are inspired by a trend in digital tools for children to create original compositions, reflect on those creations and learn through the creative process. KidPad [4] is a collaborative storytelling tool that supports children to create hyperlinked stories in a large two-dimensional zoomable space. Jabberstamp [6] is a tool for children to embed audio recordings into their drawings created on paper. This tool allowed children to compose and arrange their recordings on a graphical canvas, but did not provide tools for children to manipulate or play with their recordings nor to record video. Picture This! [11] is a video editing and capturing device designed for young children to craft movies with physical toys and artifacts. As children play with the toys to act out a story, the system analyzes their gestures and play patterns so that they can alternate between the role of the characters and cameramen.

Various projects in our community have looked at how interactive agents can facilitate children's learning. The interaction design of our character in the video booth was inspired by the puppet in Family Story Play, which was a video conference based system mediated by an animated Elmo character that invites grandparents to read books together with their grandchildren [7]. While Family Story Play focuses on dialogic reading over distance, StoryFaces focuses on children exploring emotions and creating stories with them.

Animated online greeting card services such as "Elf Yourself" [3] allow children and adults alike to insert photos of themselves into animated e-cards. Such greeting cards do not generally elicit a variety of emotional expressions from the users, nor do they support narrative creation or editing. Our work illustrates how such online tools may be adapted to support children's emotional and literacy learning.

Evaluation
A one-hour formative pilot study was conducted by a researcher with two girls (sisters called "Ani" and "Paula" in this paper) ages 4 and 6 in their home. Our objective was to assess the overall system design and identify strengths and weaknesses of the concept. A laptop computer with a built-in webcam was set up in the girls' living room on a low table, so that they could use the computer while sitting on the floor. The table was arranged against a window to assist with camera lighting, and the study was documented with screencast software and an external video camera.

"In the Balloon" was the first of three stories to be introduced to the girls by the researcher. "In the Balloon" was scripted specifically for two children to

Figure 7: Paula (6) and Ani (4) see their recorded expressions bring the story to life.

author together, and this story was chosen as an introduction so that the researcher could immediately gauge how both girls comprehended the system features and concepts.

The researcher explained that the girls would be able to record themselves making different faces into the camera, and that they would each get to take turns making video recordings. Basic UI elements were explained, such as how to stay in frame for the camera, and how to start a recording. The researcher did not explain that the recordings would become part of a storybook.

Then the children began making recordings. Both children were observed to spontaneously make faces before pressing the "record" button, in direct response to Ema's prompts. After needing to remind both children to press "record" several times, the researcher

helped the children press the GUI buttons. Both children recorded several videos that were prompted by Ema, such as "make a happy face," and "sing your favorite song." Then, they watched the story come to life, with their own videos as part of the storybook images. Ema the narrator read, "One girl got scared and started to cry," and the video of Paula crying came to life in the illustration (Figure 7). Both girls laughed hysterically at seeing Paula pretend to cry in this story, and lamented when the story was finally over. After the story Ani commented, "I want a different one," asking to try a new story.

The children began "Walking in the Woods," a second StoryFaces narrative. Neither child wanted to be left out, both trying to get in the camera view to record themselves with Ema. The researcher asked, "Why don't you take turns. Or you can both cram your faces in there together." Laughing, the girls said "Yeah!" and they did this for the many video recordings in this story

Figure 8: Paula and Ani squeeze into frame to record together, suggesting a need to consistently support shared use in the UI.

(Figure 8). After watching this story retold, Ani asked, "Can we do this one again?" She wanted to do it herself, and they started over. The researcher pressed the "record" button for Ani on every movie, because it seemed to allow Ani to improvise more easily. When watching the story later, Ani commented on one page, "Those are the same people," seeing that both the princess and the monster had the same face - her own!

At the end, Paula exclaimed, "My turn Ani, it's my turn!" and pushed her out of the way. Paula performed very exaggerated expressions, and when viewing the story later, giggled and smiled at herself, mirroring the expressions she saw herself doing in the storybook images. The researcher suggested another recording session with both girls taking turns for the different characters, but they instead asked to do a new story.

Story 3, "Expressions with an Alien" featured the children showing an alien how earth people look when they feel different ways. At the end of the story, Ani seemed to quickly grasp the meaning of the narrative. Paula asks, "why is (the alien) making a very, very smiley face?" Ani said, "They're friends!"

After Paula experiments with "Expressions with an Alien," the researcher offered, "Do you want to try mixing up the movies in the story and putting them in different places?" Paula asked, "Right now?" The researcher introduced the Backstage features to Paula, in which she could drag-and-drop video clips into the storybook illustrations, or edit the story text (Figure 4). Paula excitedly set out to craft a new story with the "wrong" faces on the pages. Due to some software bugs with the Backstage features Paula was not able to finish reordering the movies. She noted, "Everyone makes

mistakes. Even computers," and then devised her own mechanism to achieve her goal. She started the whole story over, re-recording all movies, but performing the opposite of what Ema told her to do. When Ema asked her to "make a happy face," Paula recorded herself crying. When Ema asked her to laugh out loud, Paula made an angry face. And so on... When Paula viewed the final story, she laughed with delight as her expressions so obviously mismatched the story line.

At the end of this mismatched story, the researcher commented, "I think you can also change the way the story is written if you want." Although she had ignored this idea earlier in the study, Paula excitedly repeated this news to her little sister and dove in with rewriting the storyline to match her opposite faces. For the next 4 minutes, Paula (with the researcher's help) wrote new words for the story, now changing the text to match her formerly opposite expressions. When replaying the story, Paula read the new words she had written (the researcher having muted the narrations which would have now mismatched the text). In re-reading, Paula found some parts of the story that did not make sense. "Can we go back (a page)?" Paula asked. The pages auto-advanced during replaying, but she wanted to inspect her work more closely. Alas, she had to wait. When the story had played through, Paula spent an additional 9 minutes backstage, rewriting more of the story, first herself and then with the researcher typing. At the end, Paula commented, "I like making my own words and watching. I like all of it. The part I don't like so much is when I have to drag it (the movie) on. It's so hard."

In total, the children played with StoryFaces for 55 minutes, with the session being terminated by the researcher for scheduling reasons.

Discussion

The girls' reactions to the UI design are somewhat consistent with our earlier findings regarding age-range and UI. Ani (4) experienced the system in the moment: when prompted by the character Ema to create a happy face, she did so immediately, before clicking on the GUI button, "Record a happy face." The researcher commented, "you have to push the button before you make the face." "I know!" said Ani. But after her initial performance, the researcher needed to help her press the button to make the movie—and the recorded happy face was much less sincere than the original reaction to Ema. This finding suggests that for younger children, the UI would allow children to be more spontaneous and improvisational if the recordings started automatically, removing modal buttons such as "record" or "play" as much as possible. Similarly, the editing features ("Go Backstage") were not captivating Ani's attention. In line with our earlier observations with different research projects, the symbolic nature of an editing tool is likely too complex to appeal to a 4 year old, who is at a quite concrete developmental level [5]. However, this age is rich with pretend play, and Ani seemed to relish in the images of herself pretending, and seeing them take context in a pretend story.

StoryFaces features seemed to be more fully understood by Paula, who at 6 years old is in a notably different developmental stage. Children of this age are entering middle childhood, with a greater grasp of symbolic abstraction and ability to hold procedural and operational knowledge in their minds while executing a task [2]. Further, Paula was a beginning reader, and could work more fully with all of the story elements, including text. Paula spent a great deal more time with StoryFaces than her little sister did. This was due, in part, because she was quite interested in the backstage (editing) feature, spending a good deal of her time exploring the possibility to alter the story. She seemed to quite enjoy "playing with the elements of the story" and using the editing features to eventually create her own story and accompanying videos for the illustrated book.

Paula's significant time spent with Backstage (14 minutes total, about 25% of both children's time with the system) suggest that this feature should be more fully developed. As well as making the tools more reliable, the space may benefit from more complete text editing features, ability to record oneself reading the narration, and more integrated video recording and arranging capabilities, such as an ability to re-record the videos directly into the pages that appear on backstage. Several simple UI limitations emerged through the study as well, such as Paula asking for a way to view the pages larger, and a need for text wrapping in the editor.

Despite these minor technical and design problems, Paula used the Backstage tools in meaningful ways, editing videos and then the story itself, suggesting Backstage was facilitating her understanding of how the story elements came together in the story, and how they might be rearranged to alter the story meaning. This kind of understanding of story elements and context has been shown to be a core literacy skill [10],

and these early findings indicate a path for future investigations.

For both the younger and older sisters, the system seemed to be a successful mechanism for the children to cast their voice into a short narrative. With a minimal amount of prompted video recording - which was a quite engaging performance activity on its own - the children were able to make a story their own. For Paula, this was a point of entry into story creation and editing, and for both children, the system allowed them to personalize a story that follows narrative and cinematic conventions, and reflect on how personal emotional expressions can bring meaning to a traditional narrative.

We believe that technologies for children need to be adopted by the whole family to make an impact (and often to exist at all). As we noted, Ani and Paula are not regular users of technology. Neither was comfortable using the trackpad without help, and the mother, who described herself as "not interested" in technological tools for children, commented that this system was "really good." When the mother commented with some surprise, "I like the feminine quality of the project," she may have been referring to it being appropriate for her girls, or perhaps that the project highlighted an emotional dimension that is rarely highlighted among technological tools. Our further studies will address these possibilities.

Future Work
Our pilot study highlights improvements to the basic UI design, including a modeless (timer based) video booth, better scaffolding for multiple children to author together, and a more robust Backstage. Following the

successful exploratory work outlined in this paper, these improvements will be implemented and tested on a larger scale.

The study also suggested that a touch-based tablet with a built-in webcam might be a better platform for young children to engage in audiovisual creation than a standard laptop. A touch screen could give young children more direct and intuitive controls for basic UI operations such clicking and dragging movies into storybook illustrations, or advancing through story pages with the flick of their fingers.

Touch interactions point towards a more significant opportunity: to get more physical with StoryFaces. Future investigations should include support for body-scale performances like "jumping up and down with happiness" or "stomping your feet in disgust," etc. as well as support of gestural and full body interactions with the UI. Unencumbered full-body interactions around large displays have been shown to engage children's social and emotional play [9] and may bring the entry level for this type of activity to even younger children.

Another path for development is to expand the social possibilities for children's creative storytelling. Could StoryFaces support children's work and play over a distance, so that two friends who are in different locations might create and play with stories together, in the spirit of a "video playdate? [12]" Once tools like StoryFaces are online, we can investigate how the stories could become artifacts a whole family can enjoy, for example through online sharing of the creative process, and well as sharing of finished stories with friends and family who are far away.

Conclusion

We have presented our prototype of StoryFaces, a tool to support young children's exploration of emotional expressions and reflection in storytelling play. We reported a formative evaluation with two children ages 4-6. Through recording and replaying performances, arranging and rearranging those performances, StoryFaces invited the children to explore a variety of emotional expressions and play with these expressions as part of their story elements. Our preliminary results suggest that a digital authoring can give young children an opportunity to play and reflect on their pretend emotions, and that when these emotions are cast into the context of a traditional narrative, children can engage with and meaningfully manipulate elements of those stories. This research suggests that multimedia authoring tools can leverage children's performative play and more traditional forms of content like storybooks to support the social-emotional and literacy development of young children.

Acknowledgements

References

[1]   Ananny, M (2002). Supporting children's collaborative authoring: practicing written literacy while composing oral texts. *Proc. of the CSCL*, Boulder, CO.

[2]   Bruner, J. (2004) *Towards a Theory of Instruction*. Belknap Press.

[3]   Elf Yourself. http://www.elfyourself.com/

[4]   Juan Pablo Hourcade, Benjamin B. Bederson, Allison Druin, and Gustav Tax. KidPad: collaborative storytelling for children. In *CHI '02 extended abstracts on Human factors in computing systems*.

[5]   Piaget, J. (1976). *The grasp of consciousness.* Cambridge: Harvard University Press.

[6]   Hayes Raffle, Cati Vaucelle, Ruibing Wang, and Hiroshi Ishii. 2007. Jabberstamp: embedding sound and voice in traditional drawings. In Proc IDC '07. ACM, New York, NY, USA, 137-144.

[7]   Hayes Raffle, Rafael Ballagas, Glenda Revelle, Hiroshi Horii, Sean Follmer, Janet Go, Emily Reardon, Koichi Mori, Joseph Kaye, and Mirjana Spasojevic. 2010. Family story play: reading with young children (and elmo) over a distance. In *Proc. CHI '10*. ACM, New York, NY, USA, 1583-1592.

[8]   Singer, Jerome L. (1998). Imaginative play in early childhood: A foundation for adaptive emotional and cognitive development. *International Medical Journal.* 1998 Jun; Vol 5(2): 93-100.

[9]   Scott S. Snibbe and Hayes S. Raffle. Social immersive media: pursuing best practices for multi-user interactive camera/projector exhibits. *Proc. CHI '09.*

[10] Snow, C. E. (1983). *Literacy and language: Relationships during the preschool years.* Harvard Educational Review, 53, 165-189.

[11] Cati Vaucelle and Hiroshi Ishii. 2009. Play-it-by-eye! collect movies and improvise perspectives with tangible video objects. *Artif. Intell. Eng. Des. Anal. Manuf.* 23, 3 (August 2009), 305-316.

[12] Svetlana Yarosh, Kori M. Inkpen, and A.J. Bernheim Brush. 2010. Video playdate: toward free play across distance. In *Proc. CHI '10*. ACM, New York, NY, USA, 1251-1260.