Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

New Dimensions in Testimony: Digitally Preserving a Holocaust Survivor's Interactive Storytelling

David Traum[1], Andrew Jones[1], Kia Hays[2], Heather Maio[3], Oleg Alexander[1], Ron Artstein[1], Paul Debevec[1], Alesia Gainer[1], Kallirroi Georgila[1], Kathleen Haase[1], Karen Jungblut[2], Anton Leuski[1], Stephen Smith[2], and William Swartout[1]

[1] USC Institute for Creative Technologies, 12015 Waterfront Drive, Playa Vista CA 90094-2536, USA
[2] USC Shoah Foundation, 650 West 35th Street, Suite 114, Los Angeles CA 90089-2571, USA
[3] Conscience Display, 7155 Oakwood Ave, Los Angeles CA 90036, USA

> As survivors dwindle, what will this mean for memories of the Holocaust?
>
> The Independent [6]

Abstract. We describe a digital system that allows people to have an interactive conversation with a human storyteller (a Holocaust survivor) who has recorded a number of dialogue contributions, including many compelling narratives of his experiences and thoughts. The goal is to preserve as much as possible of the experience of face-to-face interaction. The survivor's stories, answers to common questions, and testimony are recorded in high fidelity, and then delivered interactively to an audience as responses to spoken questions. People can ask questions and receive answers on a broad range of topics including the survivor's experiences before, after and during the war, his attitudes and philosophy. Evaluation results show that most user questions can be addressed by the system, and that audiences are highly engaged with the resulting interaction.

Keywords: video • natural language dialogue • Holocaust survivor testimony

1 Introduction

The original method of interactive storytelling is still one of the most compelling: relaying of first-person experiences in face-to-face conversation. Analog and digital means have been used to preserve, and in some cases enhance, storytelling, but at the cost of the interactive aspect. Particularly with first person narratives,

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

it can be especially compelling to look the narrator in the eye, ask questions, and make a personal connection with the narrator and thus the narrative.

This is especially true for Holocaust studies. Over the years, personal conversations between Holocaust survivors, the public, and students have shaped the way in which the next generation has experienced the Holocaust in a visceral way. One person's direct conversation with another is an intimate and powerful means to educate, connect and inspire. Currently, Holocaust survivors' in-person testimonies and 'Question and Answer' sessions form a major component of Holocaust education at museums and in classroom studies. For example, Bar-On [3] writes, "Story-telling also has an emotional component of connecting. When survivors come to the classroom and tell first hand their personal experiences during the Holocaust, children feel what the survivors are going through again and again by telling their stories, and they appreciate this and are willing to listen to learn about this period in history first hand. They ask questions and are willing to read more and thereby enrich their knowledge about an era that to them lies far back in history. Suddenly, the figures and dates became alive in front of them." Lieberman [14] writes, "The actual testimony of witnesses provides us with a three dimensional life-breathing force, from which we cannot escape and which we cannot deny. When this testimony is presented first-hand to our young people, it becomes a mind shattering and mind-altering experience."

However, in a few short years, Holocaust survivors will no longer be with us to share their experiences first hand through such personal encounters. Especially given the 70th anniversary of the end of the war this year, there have been many recent newspaper articles discussing the imminent change in Holocaust education, transitioning from live interaction to second-hand narratives and recorded testimonies (for example [6, 16]).

We have created another alternative: live interaction with recordings of the survivor. Rather than hear the survivor's story in a linear way (as in a documentary film), future generations will be able to interact with the storytelling through conversation. Our hypothesis is that we can preserve much of the experience that students have with direct testimony and Q&A from survivors, using the following elements: a structured interview process to elicit answers to most of the questions visitors have, a high-quality recording process, immersive display of the recordings, and direct spoken language interaction to trigger contextually relevant recordings. What makes our project unique is the ability to connect on a personal level with a survivor, and the history, even when that survivor is not present.

In the next section, we review related work on computer systems that allow people to interact with and hear narratives with a historical or virtual character. In section 3, we describe the elicitation and recording protocols designed to collect engaging and immersive narratives from a survivor that can be used to support digital interaction with an audience using speech. In section 4, we describe the system architecture and functionalities. In section 5, we give an example of use of the system and describe preliminary evaluation results. Finally, we conclude in Section 6.

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

2 Related Work

One of the first systems that allowed spoken interaction with a historical character was the August system [8]. This was a 3D "talking head" fashioned after August Strindberg, and could give tourist information about Stockholm, as well as deliver quotes from and information about Strindberg.

In the late 1990s Marinelli and Stevens came up with the idea of a "Synthetic Interview", where users can interact with a historical persona that was composed using clips of an actor playing that historical character and answering questions from the user [15]. "Ben Franklin's Ghost" is a system built on those ideas and it was deployed in Philadelphia in 2005–2007 [17]. The system used speech recognition and keyword-spotting to select the responses.

Sergeant Blackwell [10] was a full-bodied virtual human, shown full-sized on a transparent screen, who told stories in response to unrestricted user questions. The character was fictional.

All of the above systems utilized writers to create the narratives, and actors or artists to create the visuals. An early system to allow interaction with recordings of a real person was "Ask the President" at the Nixon presidential library in the early 1990s [5], but users were only allowed to choose from a set of predefined questions. As far as we know, the first system to enable conversational interaction with elicited recordings of a real person was [2]. This system had only a small amount of content, and showed that it could be interesting to an audience, but did not demonstrate whether it could work with real users and their own questions, which is necessary for the type of engagement that people have in face-to-face interactions.

3 Recording "Future-Proof" Stories

Our goal is to create a system that can carry on direct interaction with students and others far into the future. This means considering not just current technology and interests, but also trying to anticipate the requirements for interaction in the future. In this section, we describe the process for deciding on interview prompts that can lead to answers and stories that will work in an interactive situation with a diverse collection of users. We also describe the recording process, to generate high-quality video that can be used for future display technologies.

3.1 Preparing for the Interviews

The preparation consisted of several activities, including preparing a list of prompts needed and answers and stories expected, but also preparing the survivor and interviewer for this kind of experience.

Our initial material was generated from experiences with Holocaust survivors giving their testimonies and answering questions. An initial list was drafted and sent to experts in the fields of Holocaust testimony preservation, history, genocide studies, trauma specialists, Holocaust museum education staff, and representative target audiences. Questions were also gathered from audience members

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

who had seen a film about the survivor and experienced a live question-answer session.

In addition to the collected questions, we crafted a set of questions designed to elicit specific stories and other bits of information, based on our prior familiarity with the survivor. This was done in recognition that any collected set of questions will have some gaps, and that a good story can often serve as a response to a question that did not ask for it specifically. Both the collected and the devised questions were categorized according to themes and arranged into a set of interview scripts for recording.

In addition to the narratives regarding the survivor's experiences, we also devised a set of prompts to further the interaction, including short factual biographical information, opinions, and a variety of non-answers that could be used when there was no content answer available. We also solicited multiple versions of some stories, so that a questioner could receive different levels of detail. After an initial set of survivor statements were recorded, new questions were collected through conversational interaction with the recorded stories, and used to devise a second elicitation script [1]. Overall, we devised prompts for over 2000 survivor contributions.

3.2 Recording Process

In order to preserve testimony for the future, our project targeted a wide range of display types including new types of displays that may be developed over upcoming decades. In recent years there have been two major trends in video production. The first trend is that display resolution has increased from standard definition (640×480 pixels) to high definition (1920×1080) and 4k content. This trend was motivated by rapid advances in digital camera sensors and thin-panel LCD displays allowing smaller and denser pixels. The second trend is a resurgence of 3D stereo content. This includes both glasses-based 3D movies as well as new immersive virtual reality headsets. Other technologies are emerging that could enable 3D perception without the need for glasses or headware. To span all these areas we focused on three delivery form factors: traditional 2D displays, glasses-based 3D stereo displays, and glasses-free 3D displays (also referred to as automultiscopic displays).

During production, we looked to achieve the dual goals of high-resolution and multiview capture for 3D displays. Our hybrid approach combined a pair of top-of-the-line digital cinema cameras, to capture a central stereo view, with low-cost HD consumer cameras to record additional viewpoints. The digital cinema cameras were RED Epic cameras with a 6k dragon sensor. These two cameras were mounted as a stereo pair using a mirror box to position the cameras close together and approximate the distance between the human eyes. The consumer cameras were Panasonic X900MK cameras, spaced every 6 degrees over a 180 degree arc (Figure 1). The cameras were chosen as they recorded HD footage with 3 sensors at 60fps. In addition, as the Panasonic cameras record data directly to SD cards with MPEG compression, we could record 12 hours of footage on a single 128GB card. The higher resolution RED cameras were limited to 60–75

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

minute continuous takes recorded on 512GB flash drives. In practice, this was not a problem, as it provided natural breaks in the interview. Scene illumination was provided by a LED-dome with smooth white light over the upper-hemisphere (Figure 2). This is a flattering neutral lighting environment that also avoids hard shadows.

Fig. 1. Panasonic cameras used to capture multiple views of the interview.

Fig. 2. Survivor in the lighting environment.

A key feature of natural conversation is eye-contact, as it helps communicate attention and subtle emotional cues. Ideally, future viewers will feel that the storyteller is addressing them directly. This could best be achieved if the survivor would maintain eye contact with the central stereo RED cameras. To create such eye contact, we placed the interviewer off to the side, so that their face was visible as a reflection in the stereo mirror box aligned with the central cameras. An opaque curtain, placed on the direct line of sight between the survivor and the interviewer, prevented the survivor from turning towards the interviewer directly.

A major consideration during the interviews was maintaining video and audio continuity. This is important as the interactive storytelling may jump back and forth between different takes and even different days of production. As much as possible, cameras were triggered remotely to avoid any unnecessary camera motion. We also prepared multiple identical outfits for the survivor to wear on successive days. Between interview sessions we would try to match body posture and costume. A video overlay was used to rapidly compare footage between sessions. Even with all these efforts, maintaining complete continuity was not possible. In particular, we noticed changes in how clothing would fold and hang as well as changes in the survivor's mood over the course of days. Both types of changes may be noticeable when transitioning between disparate contributions.

4 System Architecture

Figure 3 shows the architecture of the system. The user watches the survivor on a video monitor with speakers, and interacts by speaking into a microphone

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

and clicking a push to talk button to tell the system when to listen. Software components include:

- a speech recognition system that converts user speech into text;
- a natural language classifier that selects an audio/visual recording as a response, given the user utterance text and prior context (section 4.1);
- a video player that plays processed, high-quality videos, and manages transitions and idle behaviors between recordings (section 4.2).

A message bus supports communication between the software components.

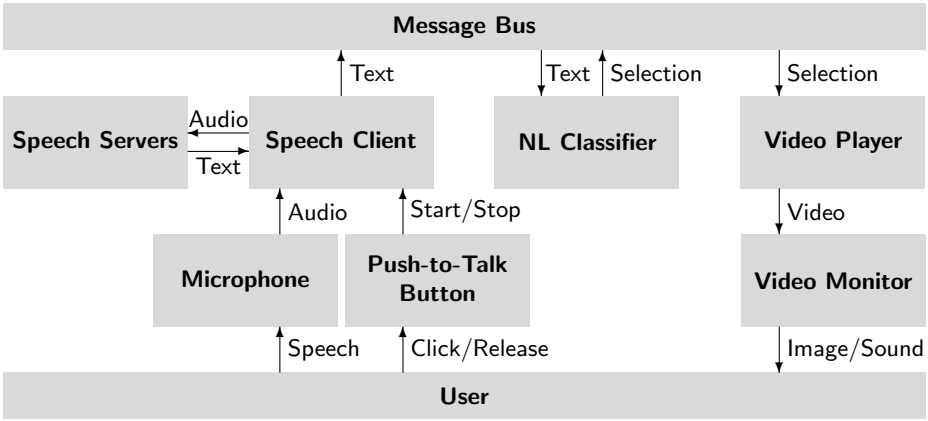

Fig. 3. System Architecture

4.1 Natural Language Processing and Interactivity

For speech recognition, we use both freely available general purpose recognizers and a custom-built language model, trained on our domain. Our NL classifier is built using NPCEditor [12], which supports both answer classification and a custom dialogue management policy. Trained on questions and their associated answers, a statistical algorithm builds a model that predicts words that are likely to appear in the answer, given the words that are seen in the question. Responses are ranked based on how closely they match the predicted answer words. This approach is fairly robust to speech recognition errors and to variant phrasings.

NPCEditor's dialogue management logic is designed to avoid cases where the top choice of the classifier is still not good. During training, NPCEditor calculates a response threshold based on the classifier's confidence in the appropriateness of selected responses: this threshold finds an optimal balance between false positives (inappropriate responses above threshold) and false negatives (appropriate responses below threshold) on the training data. At runtime, if the confidence for a selected response falls below the predetermined threshold, that response is

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

replaced with an "off-topic" utterance that asks the user to repeat the question or takes initiative and changes the topic [11]. Such failure to return a direct response (also called non-understanding, [4]) is usually preferred over returning an inappropriate one (misunderstanding). The dialogue manager also seeks to avoid repetition, so if the top ranked answer has been said recently, a different response (if any more are above the threshold) is chosen. Further details about the classifier and dialogue policy are available in [18].

4.2 Visual Processing and Presentation

Primary post-production consisted of segmenting the interview into stand-alone video responses. The initial rough edit points are marked during the interview transcription process. These in/out points are refined by automatically detecting the nearest start and end of the speech where the audio levels rose above a threshold. Occasionally, the detected audio start and end points will not exactly match the natural video edit points. For example, if the survivor made silent hand-gestures prior to talking these should be included in the video clip. In these cases we manually clean up the audio and video edit points.

We developed a custom video player that can instantaneously transition between multiple video clips as triggered by the NL engine. The video player generates different resolutions and video formats based on the display type. We use video from a single RED camera for traditional 2D playback, and footage from both RED cameras for glasses-based stereo displays. The video player alternates between playing video responses to specific questions, and an idle mode where several passive listening videos are looped. We explored several different visual transitions to connect video clips. We found that direct jump cuts between clips could be distracting, so for most applications we applied a simple half-second dissolve timed to coincide with the survivor's initial motion.

New display types such as automultiscopic displays or virtual reality headsets will require more sophisticated processing. We use a technique called *light field rendering* to interpolate arbitrary views around the survivor [13]. A light field is defined as the set of all possible light rays leaving a scene. Each digital camera records a slightly different subset of these light rays. If there are enough cameras, any new viewpoint could be synthesized by sampling the nearest light rays based just on existing image pixels. The advantage of this technique is that it directly samples the original video without reconstruction of 3D geometry. Traditional light field capture requires dense arrays containing hundreds of cameras to create smooth view interpolation [19]. In order to handle more sparse setups such as our Panasonic camera rig, we utilize pair-wise optical flow correspondences to more accurately sample pixels between adjacent cameras [7]. Flow-based light field rendering allows users to move seamlessly around the survivor, while generating any view over the front 180 degrees.

A constant desire throughout post-production was to maintain the authenticity of the survivor. In any documentary film, editing can play a large role in how a story is perceived. We endeavored to present all answers in their entirety, including, for example, lengthy pauses where the survivor would gather

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

his thoughts. Although it is theoretically possible to use 3D animation to adjust body posture, clothing, and even facial expression in post-production to create more seamless transitions, we chose instead to use video dissolves for transitions. While video dissolves are more noticeable, it is also clear where each answer ends and that all presented video frames are taken directly from the original interview.

5 Examples of Use and Evaluation

We are currently evaluating the usability of the system and the learning and emotional connection with the storytelling that the system can effectively facilitate. Initial evaluation focused on training data collection and analysis of the coverage of the recorded answers and the performance of the NL components of the system, in terms of the coherence of the answers to user questions. We review these results in Section 5.1. We are also conducting extensive testing of the system at the Illinois Holocaust Museum and Education Center in Skokie, in a setting where a museum docent relays questions from a large school group audience to the system. A detailed evaluation of two docent-mediated sessions (one from Illinois and one local) is described in section 5.2.

Figures 4 and 5 show an excerpt from a session in the museum. Audience members ask questions which are relayed (and possibly rephrased) by the docent to the system. Utterances 3, 6, and 9 show stories from the survivor delivered as interactive answers to audience questions.

5.1 Performance Evaluation of the System

We tested the system with individual museum visitors prior to the testing with docents and large classes described above. In both of these settings, 95% of user questions could be answered by the set of over 1700 survivor utterances [1]. Currently, the system does not find all of these ideal answers, due to either poor speech recognition, lack of adequate classifier training data, or other issues with classification or dialogue management (such as insufficient tracking of context or deep understanding and inference). However, the system does return an answer deemed appropriate over 2/3rds of the time, with most of the remainder being non-understanding answers [18]. Even most of the "incorrect" answers are topically related to the question, often including stories that are interesting in their own right, even though telling them seems to indicate mishearing or misunderstanding of the original question. We expect this level of accuracy to improve as we gather more data from actual user interactions with the system. However, even at this level of accuracy, we are already seeing that the conversational approach to interaction with media can have a profound effect on users, as we describe in the next section.

5.2 Educational Impact of Stories

The interaction between audiences and the system has been the subject of an external evaluation [9]; here we report highlights from this effort. Two groups of

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

1 Audience	Are there any other surviving members of your family?
2 Docent	Did any members of your family survive?
3 Survivor	Unfortunately from my immediate family nobody survived. In other words my parents, my mother, my father, and my sister, my twin sister, they were murdered the same day when we arrived at the death camp and concentration camp Majdanek. But after the war I found that out of over a hundred and fifty family members, extended ones, four cousins remained alive.
4 Audience	In the first years after you were liberated, what did you really do?
5 Docent	What did you do immediately after you were liberated?
6 Survivor	Well there's a very interesting story that I can tell you about the day of liberation, how I started. When we woke up on the eighth, actually on the eighth of May, the last day of the war, uh, when we woke up we found that the gates were open, the Czech gendarmes and the SS had disappeared, and we could just go out. And I was healthy enough to be able to run out and see what was going on. And I saw all kinds of things, but how I started my life is a story: I saw a wagon with two horses with their reins spanned in, and with the with the with the baytsh, you know, the the um um they were standing there ready, but nobody was standing with them. I was very fond of horses because before the war my grandfather had a farm, my mother's father, in Wieluń, and I was very much attuned to horses. In Hasidic homes we don't like cats, we don't like dogs, but horses to me were something special. I loved them, I loved to caress them, to clean them, the smell of horses, it all everything appealed to me about horses. So I went nearer – nobody. I looked around – nobody. Eventually I got courage and I climbed up, and I sat on the seat, and I pulled the reins to the right and the horse went right, so obviously I know what's what i'm doing. Then I pulled the other rein and it went left. So i took the whip which was there and I said, didn't hit them, I just said "Vio!" and they started. So i went to Theresienstadt with the horses and the cart and the wagon...

Fig. 4. Dialogue Fragment from Museum Interaction

students participated in the evaluation: high school students from the Camino Nuevo Charter Academy in Los Angeles interacted with the system at the USC campus; middle and high school students from the Boys and Girls Club of Chicago interacted with the system at the Illinois Holocaust Museum and Education Center. At both sites, students interacted with the system as a group, mediated by a docent as described above. Students also interacted with a live survivor: in Los Angeles, the students were split into two groups which differed in the order of interaction (one group experienced the system first, the other started with the live survivor); in Illinois, the students were split into two groups, and each had just one experience (system or live survivor).

The evaluation used a mixture of methods: pre/post surveys, participant observation, and focus group discussions after the interaction. In Los Angeles, pre/post surveys were taken before and after the first interaction, while the

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

7	Audience	What did you eat in the camps?
8	Docent	What did you eat in the concentration camps?
9	Survivor	That's another good question, uh, what did we eat. We were given starvation diets, so let me explain to you. Like in Skarżysko-Kamienna, in a working camp where we worked twelve hours, when we woke up in the morning we were given what they called coffee, but the coffee was made from acorns, from leaves, it was almost like hot water, it wasn't even sweet, it was just a brown muck hot water, but you had to drink it because you had to have some sustenance. And then lunch time there was a break of a half an hour, where they brought from the kitchen people who were, that that that was their job, they brought big vessels, kind of you know, kitchen, huge vats, kitchen vats, and there was, they were kind of, a kind of soup. And if you were lucky and the policeman, the Jewish policeman that was serving it out would go either to the middle or to the bottom, then you got some rotten potatoes, some rotten meat, because it was made with all the dregs of food, you know, they didn't feed you properly, and then you had some kind of sustenance, some kind of, you know, real food. But if he didn't like you or if he was cross or anything like that and he gave you from the top, then all you got was like hot water from the soup. And then when you came back to the camp in the evening, you again, sometimes you got another bowl of soup of the same kind, or, and you got a piece of bread. And depending how the guards, the Ukrainian guards were the ones that brought the bread, and they would steal some of it on the way, sometimes they did not. So depending how many breads came into the camp it was either you got one, from one bread, one piece of divided by eight, sometimes divided by twelve, but that was your food for the whole day. It was very little and people were starving and dying from hunger.

Fig. 5. Dialogue Fragment from Museum Interaction (Continued)

focus group discussion took place after the second interaction. The post-surveys showed that the system gave students a connection to the survivor, kept their attention, and had a positive impact (Table 1; note that these data do not show whether there was a meaningful difference between the conditions).

Live observation of the system interactions provides insight into the dynamics between the students and the system. The observer in Illinois noted that a few students waved hello and goodbye, and that hands were still raised at the end of the 50-minute session; these observations suggest that students felt a connection with the survivor and were engaged with the interaction. The same observer noted that questions got more detailed as time went on; this indicates a measure of depth to the storytelling, as each story generates additional interest. The observer also reports that students were impressed when the survivor responded to a request to say something in Polish by singing a Polish song, suggesting that the interactive abilities exceeded their expectations.

The Los Angeles focus groups allowed students to compare interaction with the system to interaction with a live survivor. In comparing the experiences,

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

Table 1. Percent of students rating the statements as "Strongly agree" or "Agree"

	Live survivor ($N = 28$)	System ($N = 25$)
I felt that I could connect with the story of the survivor.	57	72
I felt that the activity kept my attention.	100	80
I think that my experience in this activity will have a positive impact on me.	86	92

students tended to talk about the different survivors and their stories, rather than concentrate on the delivery methods. Students were aware, of course, of the difference; when probed about it specifically, one participant noted that sometimes the survivor in the system would answer a different question than he was asked (though he qualified this statement by noting that the response was still interesting); another student noted that this affected the way that they asked questions – they tried to ask less specific questions to the system.

6 Conclusions

We have presented an interactive system that allows first person audio-visual narratives to be presented to an audience in an interactive fashion, where people can engage in spoken, conversational interaction and ask questions to trigger narratives. The system has "new dimensions" including interactivity and high-fidelity recording, to allow an engaging and emotional experience somewhat comparable to direct interaction with a Holocaust survivor. Methodical recording preparations and processes and digital technology support this experience. Ongoing evaluation results suggest that the system can be accurate enough to support conversational interaction and timely delivery of narratives, and have the desired impact on the target learner population.

Acknowledgments The New Dimensions in Testimony prototype was made possible by generous donations from private foundations and individuals. We are extremely grateful to The Pears Foundation, Louis F. Smith, and two anonymous donors for their support. We thank the Los Angeles Museum of the Holocaust, the Museum of Tolerance, New Roads School in Santa Monica, and the Illinois Holocaust Museum and Education Center for offering their facilities for data collection and testing. We owe special thanks to Pinchas Gutter for sharing his story, and for his tireless efforts to educate the world about the Holocaust.

This work was supported in part by the U.S. Army; statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

Published in *Interactive Storytelling – 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30–December 4, 2015, Proceedings*, edited by Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute (*Lecture Notes in Computer Science* 9445), pp. 269–281. Springer, Heidelberg, 2015. The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-27036-4_26

References

1. Artstein, R., Leuski, A., Maio, H., Mor-Barak, T., Gordon, C., Traum, D.: How many utterances are needed to support time-offset interaction? In: Proceedings of FLAIRS-28. pp. 144–149. AAAI Press, Hollywood, Florida (May 2015)
2. Artstein, R., Traum, D., Alexander, O., Leuski, A., Jones, A., Georgila, K., Debevec, P., Swartout, W., Maio, H., Smith, S.: Time-offset interaction with a Holocaust survivor. In: Proc. IUI '14. pp. 163–168. Haifa, Israel (February 2014)
3. Bar-On, D.: Importance of testimonies in Holocaust education. Dimensions Online: A Journal of Holocaust Studies 17(1) (2003)
4. Bohus, D., Rudnicky, A.I.: Sorry, I didn't catch that! – An investigation of non-understanding errors and recovery strategies. In: Proceedings of SIGDIAL. pp. 128–143. Lisbon, Portugal (September 2005)
5. Chabot, L.: Nixon library technology lets visitors 'interview' him. Los Angeles Times (July 21 1990), http://articles.latimes.com/1990-07-21/news/mn-346_1_richard-nixon
6. Dejevsky, M.: As survivors dwindle, what will this mean for memories of the Holocaust? The Independent (January 5 2014)
7. Einarsson, P., Chabert, C.F., Jones, A., Ma, W.C., Lamond, B., Hawkins, T., Bolas, M., Sylwan, S., Debevec, P.: Relighting human locomotion with flowed reflectance fields. In: Proceedings of the 17th Eurographics Symposium on Rendering. pp. 183–194 (2006), http://dx.doi.org/10.2312/EGWR/EGSR06/183-194
8. Gustafson, J., Lindberg, N., Lundeberg, M.: The August spoken dialogue system. In: Proceedings of Eurospeech 99 (1999)
9. Kim, K.A.: New Dimensions in Testimony: Findings from student pilots. Internal report, USC Shoah Foundation (August 2015)
10. Leuski, A., Pair, J., Traum, D., McNerney, P.J., Georgiou, P., Patel, R.: How to talk to a hologram. In: Proc. IUI '06. pp. 360–362. Sydney, Australia (2006)
11. Leuski, A., Patel, R., Traum, D., Kennedy, B.: Building effective question answering characters. In: Proceedings of SIGDIAL. Sydney, Australia (July 2006)
12. Leuski, A., Traum, D.: NPCEditor: Creating virtual human dialogue using information retrieval techniques. AI Magazine 32(2), 42–56 (2011)
13. Levoy, M., Hanrahan, P.: Light field rendering. In: Proceedings of SIGGRAPH '96. pp. 31–42 (1996), http://doi.acm.org/10.1145/237170.237199
14. Lieberman, L.: Using testimony in the classroom. Dimensions Online: A Journal of Holocaust Studies 17(1) (2003)
15. Marinelli, D., Stevens, S.: Synthetic interviews: The art of creating a 'dyad' between humans and machine-based characters. In: Proceedings of MULTIMEDIA '98. pp. 11–16 (1998), http://doi.acm.org/10.1145/306774.306780
16. Schechter, D.: After the survivors: Facing a future without Holocaust witnesses. Atlanta Jewish Times (April 9 2015)
17. Sloss, E., Watzman, A.: Carnegie Mellon's Entertainment Technology Center conjures up Benjamin Franklin's ghost. Press release, Carnegie Mellon Media Relations (June 28 2005), http://www.cmu.edu/PR/releases05/050628_etc.html
18. Traum, D., Georgila, K., Artstein, R., Leuski, A.: Evaluating spoken dialogue processing for time-offset interaction. In: Proceedings of SIGDIAL. pp. 199–208. Prague (Sept 2015), http://www.aclweb.org/anthology/W/W15/W15-4629.pdf
19. Wilburn, B., Joshi, N., Vaish, V., Talvala, E.V., Antunez, E., Barth, A., Adams, A., Horowitz, M., Levoy, M.: High performance imaging using large camera arrays. ACM Trans. Graph. 24(3), 765–776 (Jul 2005), http://doi.acm.org/10.1145/1073204.1073259