

Using Dialogue System Technology to Support Interactive History Learning

David Traum

USC Institute for Creative Technologies
12015 Waterfront Drive, Playa Vista CA 90094-2536, USA
traum@ict.usc.edu

Abstract

We describe the use of spoken dialogue technology to enhance informal history learning. We describe several uses for this technology, including allowing learners to engage in natural interactions at a historical site, allowing learners to talk with recreations of historical figures, and using oral history recordings of a witness to create a dialogue experience. Two projects are highlighted, one to give a guided experience of a historical location, and another, *New Dimensions in Testimony*, that allows an experience similar to face to face conversation with a Holocaust survivor. These techniques allow many of the benefits of an intimate connection to historical places and people, through direct interaction and user initiative, but can also be delivered to a mass audience, formerly only reachable by broadcast, non-interactive media.

1 Introduction

While professional historians are used to sifting through artifacts, written records, and other evidence of the past, most people learn history through direct narrative accounts by experts. Mass media, such as printing, audio recording, and film allow wide dissemination of narrative history, but the most compelling means of connecting to the material is often direct, interactive, face-to-face conversation with an expert, or even someone who experienced the events in question. This “informal” setting allows the learner to help guide the experience and ask questions when they don’t understand, or would like to delve further into a particular topic. Moreover, it allows the expert to be able to sense the level

of knowledge and interest of the learner, and tailor the length and focus of the testimony to the appropriate levels. Interactive narrative history is a common presentation means in history classes, in tours of historical sites and in history museums, and also part of informal conversations with elders. Especially powerful is the opportunity to directly talk with someone who experienced the events and can provide first person testimony. However not everyone can physically travel to a historical location, and very few people will be able to have a conversation with direct witnesses to important historical events.

In this article we outline how dialogue system technology is being used to re-create the experience of interacting with first-person history on a scale that was formerly possible only for non-interactive materials. In section 2, we introduce the type of technology used and related uses. In section 3 we describe use of this technology to support interactive dialogue relating to a visit to a historical site (in this case a simulation of the site). In section 4, we briefly review prior work on interactions with (authored) simulation of dialogue interaction with historical people. In section 5 we give an overview of a project that allows conversations with directly recorded material of an important witness to important historical events, specifically a Holocaust survivor. We conclude in section 6, with prospects for future use of dialogue system technology for historical preservation and education.

2 Conversational Technology for Informal Education

One of the major successes of AI over the past half century has been in conversational systems. While the idea of talking to a machine was once the realm of science fiction, now there

are many systems that allow interactive spoken dialogue, on a variety of topics. Systems can disseminate information at call centers 24 hours a day for low cost. Mobile phone-based and in-car assistants can perform simple tasks, given oral commands. There are also systems that engage in more social dialogue, including socially assistive robots, aimed at long-term care and companionship. See (Jokinen and McTear, 2009) for many examples of dialogue systems and alternative approaches to their construction.

There are also numerous systems that are used for education. Some systems allow people to engage in conversation to practice a conversational skill (such as learning a new language, or how to engage in an interview). Others act as teachers or tutors, instructing users and helping them form correct cognitive models of target material. There are also systems that allow a user to collaborate with a peer in constructive problem solving.

For informal education, that is not part of a formal curriculum, there are additional demands for a system to be successful. Here, the student is not a “captive audience”, and can leave if they are bored or frustrated with the activity. In this case, it may be important to give the user some measure of control - not just a fixed narrative or top-down directed dialogue, but an ability for the user to guide the interaction or redirect it when they want, for example by asking questions.

The NPCEditor (Leuski and Traum, 2011) has been used to create several informal education systems, including guides for an exhibit hall in a science museum (Swartout et al., 2010). The system combines a cross-language retrieval approach to utterance classification with an authoring environment that allows creation or importation of annotated training data. The classification algorithm tokenizes the text of the utterances, computes separate language models for the questions and for the answers, and learns how to translate a question into an answer language model. The system computes the likelihood of observing each token in the answer given a set of tokens in the question, and uses this model to rank the answers. This approach takes advantage of the content of both questions and answers, and

has been found to be fairly robust to imperfect speech recognition.

The authoring environment allows authors to create a set of states, each of which can be associated with a subset of available responses and training data. Dialogue management can be done either inside the NPCEditor, using information state and a scripting language, or as a separate component (e.g. (Gandhe et al., 2011)), in which case the NPCEditor is just used as a natural language understanding component (Leuski and Traum, 2008). This system has also been used for enabling informal history-learning dialogues, as outlined in the following sections.

3 A Virtual Tour Guide

One very popular way of learning about history is visiting the scene of historical events. This can sometimes allow a more concrete sense of what happened, particularly for incidents like military battles, in which the terrain and landscape can play an important role in how events unfolded. Many historical sites offer guided tours, where an expert can guide a group through the events as they walk around the location, and can narrate incidents, but also answer questions that the visitors have, or pose challenges that can deepen the visitors’ understanding.

This kind of tour has been extended and formalized within the context of military education as a “Staff Ride”. (Robertson, 2014) characterizes a staff ride as consisting of “systematic preliminary study of a selected campaign, an extensive visit to the actual sites associated with that campaign, and an opportunity to integrate the lessons derived from each.” Thus it includes study before the visit, the visit itself, and then discussion afterwards.

Not everyone will have the opportunity to travel to all important historical sites, so we can use technology to bring as much of this experience as possible to “virtual” visitors. Virtual environments such as “second life” can be used to re-create the visual features of a historical location, and serve as the basis for a virtual guided tour. Situated in such an environment, we have created a virtual human tour guide who can act as the staff ride guide, and carry out the three parts of the staff ride.



Figure 1: Pre-visit Briefing

The initial part consists of a briefing in a virtual classroom, about the location. Figure 1, shows a snapshot of the briefing, in this case a check-point at a traffic-control road block, where miscommunication led to civilian casualties. Shown in the figure are the guide, introducing the incident and asking the students (e.g., “Dusan Jan”, whose avatar is shown in the foreground) questions that can help focus the tour and after action review parts.

After the briefing, the guide takes the students on a tour of the virtual world itself, pointing out aspects such as the point of view from various locations. Figure 2 shows a snapshot from the tour. Students can ask questions, and also the guide will ask questions to challenge students. Students can answer either publicly to the whole class, or privately, just to the guide. The guide is fully automated, following a script for the main lesson plan, but also reacting to user behavior, following the design from (Jan et al., 2009).

The final segment consists of an “after-action review”, that is a tutorial dialogue with the group, but in which individual sub-conversations can also be crafted to the knowledge and gaps of each student. Different pedagogical strategies are used for each student, based on the model of their knowledge built during the tour. In this way, language technology has the potential to surpass physical group tours, in allowing a higher percentage of

tailored individual instruction than is possible with a single human tour guide. The after-action review component is described in more detail in (Roque et al., 2011).



Figure 2: Visit to Virtual Site

4 Conversations with Historical Characters

Just as seeing the location of a historical event can be beneficial for learning history, it can be equally or even more powerful to talk face to face with someone who was involved in the events. Hearing the tone of voice, facial expressions, and personal reminiscences can provide a much more powerful insight into the human impact of historical events. Unfortunately, this is only fully possible for the lifetime of the person involved. Even without technology, there have been attempts to replicate this experience, by having actors portray a famous individual, in plays and for educational experience, often using the recorded words to re-create the experience.

Technology can be used not just to represent places in a virtual world, but can also be used to create a conversational experience with a historical individual. One of the first systems that allowed spoken interaction with a historical character was the August system (Gustafson et al., 1999). This was a 3D “talking head” fashioned after August Strindberg, a famous Swedish playwright and author. August could give tourist information

about Stockholm, as well as deliver quotes from and information about Strindberg.

In the late 1990s Marinelli and Stevens devised a “Synthetic Interview”, where users can interact with a historical persona that was composed using recordings of an actor playing that historical character answering questions from the user (Marinelli and Stevens, 1998). “Ben Franklin’s Ghost” is a system built on those ideas and was deployed in Philadelphia in 2005–2007 (Sloss and Watzman, 2005). The “Ben Franklin’s Ghost” system used a menu selection or typing, although (Marinelli and Stevens, 1998) report on use of speech recognition with keyword-spotting to select the responses. These systems are similar to actors portraying historical characters. They used writers to create the narratives, and actors or artists to create the visuals. While this approach can create a similar sense of talking to the historical person, it is more of a dramatic than historical experience. The same kind of experience can be created for fictional characters who have never lived, and thus questions arise of authenticity and whether the experience really reflects the characteristics of the subject as opposed to the author, actor, or producer of the experience.

Another approach is to use primary “oral history” recordings of the actual person to create the interactive experience. 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 (Chabot, 1990). Here, visitors could hear President Nixon’s actual response to selected questions rather than see the performance of an actor or synthetic character. Unfortunately, this system did not create a full spoken conversational experience, since a conventional menu-selection approach was used, rather than spoken language technology.

As far as we know, the first system to enable conversational interaction with elicited recordings of a real person was (Artstein et al., 2014). This system had only a small amount of content, and showed that it could be interesting to an audience, but did not make clear 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 in-

teractions. This system was a proof of concept for the *New Dimensions in Testimony* project, described in the next section.

5 New Dimensions in Testimony

One historical area that is currently moving from the realm of direct personal experience to more distant history are the events surrounding the second world war, which has been called the “deadliest conflicts in human history” (Sommerville, 2008). This multi-year period of conflict involved over 100 million people from many countries around the world. In addition to the direct military actions, there were many atrocities committed against civilian populations, leading to hopeful cries of “never again”. Santayana famously wrote “Those who cannot remember the past are condemned to repeat it” (Santayana, 1920), and these events will soon be beyond the passage of direct human memory.

One case in point is the attempted genocide committed by the Nazi regime against Jews living in Europe, known as the “Holocaust”, or “Shoah” in Hebrew. In the latter part of the Twentieth century and early part of the Twenty-first century, direct testimony from the survivors of this experience has been one of the most powerful ways to teach the next generations of these events, with a hope toward preventing similar occurrences in the future. The Shoah Foundation was created in order to collect oral histories of the Holocaust and other episodes of genocide, to preserve these stories. However, as these stories move from direct first-person interactions, where learners can ask questions, to recorded oral history, some of the immediacy and direct human connection may be lost.

The USC Institute for Creative technologies has partnered with the Shoah Foundation and Conscience Display to create “new dimensions” of testimony, in which people can continue to have spoken, face-to-face interactions with survivors’s stories - this time using technology to present the preserved material in an interactive fashion. The hypothesis behind the project is that much of the experience that students have with direct testimony and Q&A from survivors can be preserved using the following:

- a structured interview process to elicit answers to most visitor questions
- a high-quality recording process
- immersive display of the recordings
- direct spoken language interaction to trigger contextually relevant recordings.

More details of the overall project can be found in (Traum et al., 2015b). In the rest of this section, we outline some of the challenges facing the dialogue processing aspects of the system.



Figure 3: Pinchas Gutter, as Portrayed in the NDT System

While we knew from projects such as those described in Sections 2 and 3 that conversational interaction with compelling characters could work using technology such as (Leuski and Traum, 2011), there were several unknown challenges at commencement of the project. These are

1. Can this work with real (octogenarian) people telling their own stories, not actors and authors creating the content?
2. Can we know what to record in one or a small number of recording sessions to reach adequate coverage?
3. Can automated speech recognition and natural language understanding and dialogue management work well enough to facilitate a good user experience?

The first challenge was addressed by building a small “proof of concept” system. This involved the creation of recorded interview material with a Holocaust survivor, Pinchas Gutter, (shown in Figure 3) who has been telling his story for years at museums, classrooms, and public forums. The system, described in (Artstein et al., 2014) enabled a knowledgeable user to demonstrate an interactive conversation. While there was broad flexibility in which questions were asked, the order of questions, or the way the user would phrase the questions, there was not enough material for naive users to just approach the system and expect their own questions to be covered.

The second challenge was potentially more daunting. Since visitors can literally ask anything, including questions never heard before or even thought of by Mr Gutter or the system designers, there is a potentially infinite set of material that might be required. Potentially the problem is “AI complete”, in that a full solution would require passing the Turing test and creating a system that would answer any question in exactly the manner that Mr Gutter would. On the other hand, in practice, many people ask similar questions, which are mainly tied to the events and thoughts under discussion. Likewise, often during conversation, people (especially politicians) will address the question on their mind rather than what someone actually asked. Thus a more limited amount of material, along with suitable “off-topic” or “out of domain” policies might suffice. The approach to this question is addressed in (Artstein et al., 2015), and involved focus groups and “wizard of oz” testing with a human-controlled version of the system, as well as a multi-part interview process, so that the most critical gaps could be filled.

The wizard system illustrated that sufficient material was available not just for a demo, as with the proof of concept system from (Artstein et al., 2014), but that people could actually have compelling interactive experiences with the recorded material, asking their own spontaneous questions.

The third challenge again invites a comparison to AI-complete capabilities: could the automated system learn to replicate the behavior of the human “wizards” closely enough to provide the same quality of experience? The process to create and evaluate this experience is described in (Traum et al., 2015a). Supervised learning techniques were used, taking the corpus of questions from the wizard system and other sources, and annotating them with correct answers. Iteration also with initial versions of the automated system provided additional data that is directly pertinent to the level of performance. While the automated system is currently not quite as accurate as human wizards, it is faster, and the overall experience of engaging in spontaneous interactive conversation is quite compelling to the audience.



Figure 4: Visitors to IHMSC interacting with NDT System

The system has been deployed at the University of Southern California (both the Shoah Foundation and ICT) and also in the Illinois Holocaust Museum and Education Center for over a year. It was also tested for three months at the US Holocaust museum. An example of the system being used with a classroom visiting the museum is shown in Figure 4. Museum staff have reported on the impact of the system on visitors, which is quite similar to the impact from hearing survivors in person.¹ Ini-

¹See for example this letter to Mr. Gutter from a staff member at the Illinois museum: <https://sfi.usc.edu/blog/doris-lazarus/meeting-pinchas>

tial testing has also borne this out with comparative surveys of visitors. Figure 5 shows an example of dialogue excerpt from interaction with a visitor at the US Holocaust museum.

After the system with Mr Gutter, 11 other Holocaust survivors have been recorded and systems for them are currently under development and will be on display starting in 2017. Plans are also underway to record people that have other important experiences but are not specifically related to the Holocaust.

6 Conclusions

As we have seen, AI technology, and specifically conversational dialogue systems, can contribute much to historical dissemination, and combine the twin strengths of direct involvement with people and places with mass dissemination, to people who are not able to have direct experiences with these places and individuals involved in historical events. Future work using these and other technologies may bring further opportunities to not just see and hear about historical events, but to “participate” in them, which is still the realm of science fiction involving time travel or a “Holodeck”.²

Acknowledgments

I would like to thank first, the many individual and institutional partners in the VIGOR and New Dimensions in Testimony projects, described in sections 3 and 5, respectively. 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

²<https://en.wikipedia.org/wiki/Holodeck>

Visitor tell me about your childhood

Pinchas my life before the war i was born in lodz in Poland and i was eight years old when the war started so the first eight years i lived in a home a very religious one a hasidic home i had a twin sister my father was a winemaker and we lived a normal life. . . .

Visitor what was it like in a concentration camp

Pinchas i had one friend in the camp it wasn't a question of whether it was easy because the type of existence that we had in the camps working twelve hours and then given very little food and trying to make ends meet and and trying to get rid of all the vermin and the sicknesses and the persecution and the torture there was very little time to actually have but people helped you i had in one camp i had a very close friend and his name was Yakov and let me tell you a story about that . . .

Visitor what do they have for food

Pinchas that's another good question uh what did we eat we were given starvation diets so let me explain to you like in skarzynsko-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 up would go either to the middle or to the bottom and 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 it from the top then all you got was like hot water from the soup . . .

Visitor were they able to take showers

Pinchas eh to to wash simply as soon as you did your ablutions there was there were some uh uh places where there was running water and you'd quickly wash your face and try and make yourself as clean as possible um but as far uh but you but what happens when you came back from work in the camp in in skarzynsko if you were diligent enough and i was you made sure that you went and found enough water to wash yourself especially after going to the toilet because there was no uh there were no um there was no paper or anything like that to wipe yourself and to try and keep you as clean as possible and that also helped you survive

Visitor how many camps were you in

Pinchas i was sent to actually i went through five concentration camps . . .

Visitor were you ever sick and can't work?

Pinchas yes i had typhoid and typhoid is a disease uh it's it's a disease which is a um very virulent and it is distributed by contact uh generally it's through lice and dirt and you get very ill and the chances of recovery are not very great it becomes like it's like an epidemic and you you get very high fever and comes a day where there is a crisis where the fever is so high that you you either die the next day or that day or you kind of you start deteriorating or you get over it and then you recover the um the only problem was is that in the camps there was absolutely nothing to knock down the fever the only thing that you could do while we were while i was not hidden is to take cold water and and and and try and cool yourself off because the fever was so was so horrible but it's a contagious disease you it's a it's a disease that you that you can get from from touching others and from getting lice from others and this is how it spread and it spreads like wildfire

Figure 5: Sample Dialogue between a visiter to US Holocaust Museum and the Pinchas Gutter system

not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

- Ron Artstein, David Traum, Oleg Alexander, Anton Leuski, Andrew Jones, Kallirroi Georgila, Paul Debevec, William Swartout, Heather Maio, and Stephen Smith. 2014. Time-offset interaction with a Holocaust survivor. In *Proc. IUI '14*, pages 163–168, Haifa, Israel, February.
- Ron Artstein, Anton Leuski, Heather Maio, Tomer Mor-Barak, Carla Gordon, and David Traum. 2015. How many utterances are needed to support time-offset interaction? In *Proceedings of FLAIRS-28*, pages 144–149, Hollywood, Florida, May. AAAI Press.
- Lucy Chabot. 1990. Nixon library technology lets visitors ‘interview’ him. *Los Angeles Times*, July 21.
- Sudeep Gandhe, Alys Taylor, Jillian Gerten, and David R. Traum. 2011. Rapid development of advanced question-answering characters by non-experts. In *Proceedings of the SIGDIAL 2011 Conference, The 12th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 347–349, Portland, Oregon, USA, 06/2011. The Association for Computer Linguistics, The Association for Computer Linguistics.
- Joakim Gustafson, Nikolaj Lindberg, and Magnus Lundeborg. 1999. The August spoken dialogue system. In *Proceedings of Eurospeech 99*.
- Dusan Jan, Antonio Roque, Anton Leuski, Jacquelyn Morie, and David Traum. 2009. A Virtual Tour Guide for Virtual Worlds. In *Intelligent Virtual Agents Conference (IVA)*, Amsterdam, The Netherlands.
- Kristiina Jokinen and Michael McTear. 2009. Spoken dialogue systems. *Synthesis Lectures on Human Language Technologies*, 2(1):1–151.
- Anton Leuski and David Traum. 2008. A statistical approach for text processing in virtual humans. In *Proceedings of the 26th Army Science Conference*.
- Anton Leuski and David Traum. 2011. NPCEditor: Creating virtual human dialogue using information retrieval techniques. *AI Magazine*, 32(2):42–56.
- Donald Marinelli and Scott Stevens. 1998. Synthetic interviews: The art of creating a ‘dyad’ between humans and machine-based characters. In *Proceedings of MULTIMEDIA '98*, pages 11–16.
- W. G. Robertson. 2014. *The Staff Ride*. United States Department of Defense.
- Antonio Roque, Dusan Jan, Mark Core, and David Traum. 2011. Using virtual tour behavior to build dialogue models for training review. In *Proceedings of the 11th International conference on Intelligent Virtual Agents (IVA 2011)*, Reykjavk, Iceland, September.
- G. Santayana. 1920. *The Life of Reason: Or, The Phases of Human Progress*. C. Scribner’s sons.
- Eric Sloss and Anne Watzman. 2005. Carnegie Mellon’s Entertainment Technology Center conjures up Benjamin Franklin’s ghost. Press release, Carnegie Mellon Media Relations, June 28. http://www.cmu.edu/PR/releases05/050628_etc.html.
- D. Sommerville. 2008. *The Complete Illustrated History of World War II: An Authoritative Account of the Deadliest Conflicts in Human History with Analysis of Decisive Encounters and Landmark Engagements*. Anness Publishing.
- William Swartout, David Traum, Ron Artstein, Dan Noren, Paul Debevec, Kerry Bronnenkant, Josh Williams, Anton Leuski, Shrikanth Narayanan, Diane Piepol, Chad Lane, Jacquelyn Morie, Priti Aggarwal, Matt Liewer, Jen-Yuan Chiang, Jillian Gerten, Selina Chu, and Kyle White. 2010. Ada and grace: Toward realistic and engaging virtual museum guides. In *Intelligent Virtual Agents*, pages 286–300. Springer Berlin / Heidelberg.
- David Traum, Kallirroi Georgila, Ron Artstein, and Anton Leuski. 2015a. Evaluating spoken dialogue processing for time-offset interaction. In *Proceedings of SIGDIAL*, pages 199–208, Prague, Sept.
- David Traum, Andrew Jones, Kia Hays, Heather Maio, Oleg Alexander, Ron Artstein, Paul Debevec, Alesia Gainer, Kallirroi Georgila, Kathleen Haase, et al. 2015b. New dimensions in testimony: Digitally preserving a holocaust survivors interactive storytelling. In *International Conference on Interactive Digital Storytelling*, pages 269–281. Springer.