The Conversational Role of Electronic Guidebooks

Allison Woodruff, Margaret H. Szymanski, Paul M. Aoki, and Amy Hurst[1]

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304 USA

Abstract. We describe an electronic guidebook prototype and report on a study of its use in a historic house. Visitors were given a choice of information delivery modes, and generally preferred audio played through speakers. In this delivery mode, visitors assigned the electronic guidebook a conversational role, e.g., it was granted turns in conversation, it introduced topics of conversation, and visitors responded to it verbally. We illustrate the integration of the guidebook into natural conversation by showing that discourse with the electronic guidebook followed the conversational structure of storytelling. We also demonstrate that visitors coordinated object choice and physical positioning to ensure that the electronic guidebooks played a role in their conversations. Because the visitors integrated the electronic guidebooks in their existing conversations with their companions, they achieved social interactions with each other that were more fulfilling than those that occur with other presentation methods such as traditional headphone audio tours.

1 Introduction

Visitors often go to cultural heritage locations, such as museums, with companions. Many seek what has sometimes been called a "learning-oriented" experience [8]. To facilitate learning, institutions typically present information through guidebooks and prerecorded audio guides as well as through labeled exhibits and docent-led tours. However, sharing the experience with companions is often a higher priority than learning, particularly for infrequent visitors [12]. Unfortunately, existing presentation methods interfere with the interaction among visitors. For example, visitors frequently complain that audio tours with headphones isolate them from their companions, and visitors have few opportunities to interact effectively with each other while docents "lecture" to them.

   We are interested in identifying electronic guidebook designs that facilitate rather than hinder social interaction. To this end, we conducted a qualitative study of visitors using an electronic guidebook. We first constructed a prototype, designing it to provide a range of options for information presentation and sharing. In particular, we provided a mechanism for visitors to hear each other's audio selections. We then observed fourteen visitors using the guidebook in a self-guided tour of a historic

---

[1] Work performed during an internship from the College of Computing, Georgia Institute of Technology.

house and conducted semi-structured interviews. Visitor actions were captured using audio recording, video recording, and device logging.

In this paper, we report an analysis of how different properties of the device helped visitors to interact with their companions while preserving their ability to engage in independent activity. A previous analysis focused on issues of attentional balance, drawing on content from the interviews supplemented by informal observations of visitor use of the guidebooks [27]. The work reported in this paper focuses on social interaction and is based on a detailed analysis of video taken while visitors used the electronic guidebooks.

This analysis demonstrates that visitors integrate the electronic guidebooks into their conversations, thereby achieving their goal of having a social experience with their companions. Specifically, visitors possess a strong desire to have conversations in which they share information and responses. A natural way for them to include information in conversation is to simply assign the information source (the electronic guidebook) a conversational role. When the guidebook plays a conversational role, the visitors are naturally positioned to have shared responses to its statements, which leads to gratifying social interactions. Further, directly incorporating the electronic guidebook in conversation is an elegant solution, since it does not require visitors to develop new skills for coordinating their interaction with the guidebook; they can simply use their existing conversational conventions.

In this paper, we show that visitors in our study assigned a conversational role to the electronic guidebook, e.g., the electronic guidebook was granted turns in the conversation, the electronic guidebook introduced topics of conversation, and visitors verbally responded to the electronic guidebook. To demonstrate this role, we show that discourse with the electronic guidebook followed the conversational structure that occurs in traditional human storytelling. We also present behavioral evidence that visitors made many efforts to listen to the same content at the same time so they could better incorporate the guidebook in conversation. For example, they chose to listen to audio through speakers so they could listen simultaneously, they negotiated about which descriptions to listen to, and they physically positioned themselves and their electronic guidebooks so they could be shared effectively.

Because the guidebook was assigned a conversational role, it enhanced social interaction between visitors. For example, visitors shared reactions to stories told by the electronic guidebook, e.g., they offered opinions or stories of their own. These shared reactions are key social interactions that do not occur with traditional presentation methods such as audio tours with headphones.

The transcription shown in Example 1 is helpful in understanding what we mean by "storytelling" and "reactions." (Table 1 summarizes the transcription notation used in all of our examples.) Two women, A and S, are interacting with each other and the guidebook. A looks at the view of the room shown in the guidebook and finds that same view in the physical room.[2] A announces this information to S so they will have a shared orientation. A then tells S which object she is about to select, thereby alerting S that a description is about to begin. Next A and S listen to the description of the object. Following this, they have a brief interaction about the contents of this description.

---

[2] The design of the guidebook is discussed in the next section.

Example 1: The story of the deer head.

A:	Okay, so here's over here, looking this way. (A takes step towards fireplace wall)
A:	We can see about the heh heh deer head. (A selects description of deer head)
A-PDA:	*This deer was shot by the Bourns' son-in-law at the Irish estate the Bourns purchased as a wedding gift for their daughter.*
S:	Oh…
A:	An Irish estate.
S:	Wouldn't you want that for a wedding gift?
A:	Eh hm.

Table 1: Summary of transcription notation.

X: X-PDA:	Visitor X is speaking. *Visitor X's guidebook is speaking.*
(*n*)	A conversational pause of *n* seconds.
[Speech of first speaker [Speech of second speaker	Overlapping speech.

In the next section, we describe our guidebook prototype. We then describe our research methods. We next describe and analyze the visitors' behavior. We then focus in detail on the conversational structure of storytelling interactions, demonstrating that the electronic guidebook has a conversational role. We discuss these findings, and then review related work and conclude.

2 Prototype

In this section, we describe the electronic guidebook used in the study. The design rationale and many of the details discussed here have been reported elsewhere [3,4,27]. However, it is important to review these details because some of the observations reported in the analysis sections are best understood with our specific design in mind.

The electronic guidebook application runs on a Casio Cassiopeia™ E-105 personal digital assistant (PDA), a small device weighing 255g (9 oz.). Its display is a color touch-sensitive screen. A user generally holds the PDA in one hand and a stylus in the other hand, touching the stylus to the screen to interact with the device.
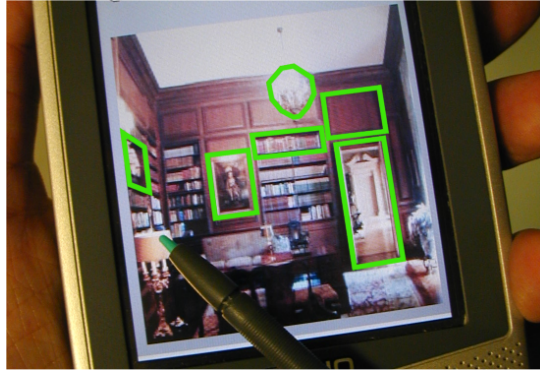
Figure 1: Electronic guidebook prototype with outlines visible.

Visitors obtain information about objects in their environment using a visual interface. The interface is akin to a set of Web browser imagemaps, but has many refinements [4] that simplify operation on a handheld device. Our prototype presents visitors with one of a collection of photographs. Each of these photographs was taken facing one wall of a room in a historic house. Visitors change the viewing perspective (i.e., display a different photograph) by pressing a button on the device. When visitors tap on an object in a photograph, the guidebook gives a description of that object, if one is available. Many, but not all, of the objects visible in a given photograph have descriptions. (These objects with associated descriptions are known as *targets*.) Because the historic house environment is complex, many different kinds of objects may be targets. Figure 1 shows a photograph with a number of targets, including a wood panel and a doorway. To help visitors identify targets, the guidebook displays outlines around each target, triggered when the user taps on the photograph but does not "hit" a target.

The visual selection design is motivated by the principles described in [3]. We learned through a combination of observation, informal interviews, and professional study [5] that system designs that seem plausible in a museum are not workable in a historic house. Most notably, location-aware systems that use sensors to select content automatically are not feasible in historic houses for a number of reasons (e.g., barriers often prevent visitors from approaching objects). Usability testing of the prototype by thirteen users, conducted prior to the study reported in this paper, confirmed that visual selection is a viable alternative that allows visitors to select objects that interest them quickly and easily.

The prototype gives visitors several choices with regard to the presentation of the descriptions. Visitors have the option of seeing a text description of an object or hearing an audio clip with identical content read by a female voice. Visitors can change the choice of textual or audio presentation at any time. Audio clips can be played at a low volume through speakers on the device or at a user-controlled volume through headphones.

The descriptions themselves are typically two or three sentences (40 words) long but do vary, with the audio duration ranging from 3 to 23 seconds. For example, the 20-second description of a portrait of the Duchess of Richmond and Lennox reads:

> "This 17th century portrait shows the Duchess of Richmond and Lennox, about whom Pepys said in his diary, 'Never had a woman more beauty nor less wit.' The portrait was done by Sir Peter Lely, who was the first of the great English portrait painters. Lely created the distinctive look of British portraiture, including the three-quarter pose and the emphasis on beautiful clothing."

3 Method

In this section, we describe the study participants, the procedure (including the setting) by which we collected the observational data, and the methods we used to analyze the data.

3.1 Participants

The study participants were members of the Xerox PARC community (not necessarily employees), accompanied by friends or relatives with whom they would normally attend a museum. For example, a grandmother attended with her 7-year-old granddaughter and a husband attended with his wife. The visitors comprised a total of seven couples and ranged in age from 7 to over 60 years of age. Two of the couples were adult-child pairs; all other visitor pairs consisted only of adults. Eight of the visitors were female and six were male. Visitors were instructed to bring glasses if they used them for reading (to our knowledge one visitor who wore glasses forgot to bring them). One visitor used hearing aids. Even with them, he was still slightly hard of hearing. Many of the visitors were non-technical and/or had not previously used a PDA. Two of the visitors had used a previous version of our prototype. Most of the visitors had not previously visited the study site. Half of the visitors described themselves as frequent museum visitors (visiting museums three or more times a year) and half described themselves as infrequent visitors (visiting museums fewer than three times a year).

3.2 Procedure

Participants were observed during a private visit to Filoli, a Georgian Revival house in Woodside, California.[3] Each visit consisted of three phases: a partial tour using a paper guidebook, a partial tour using the electronic guidebook, and an interview.

---

[3] http://www.filoli.org/

Figure 2: Observation of visitors using the guidebook.

The visitors went through the first several rooms of the house with a paper guidebook, accompanied by a docent who was available to answer questions. In some cases the docent was one who worked at the house regularly, and in other cases the docent was an escort from the research team (the research escort was always present, even when the regular docent was accompanying the visitors). During this phase, the visitors' comments and conversation were recorded using wireless microphones.

The visitors used the electronic guidebook in the next two rooms of the house. One of these rooms contained security barriers and the other did not. The visitors received brief instructions from the research escort in the use of the guidebook. They were then asked whether they would each like their own guidebook or if they would prefer to share. They were also offered headphones. They were told that they could change their decisions at any time and that the research escort could answer questions about the use of the device (this was rarely necessary). The visitors then toured the two rooms, referencing their electronic guidebooks as desired (the electronic guidebook contained a total of 42 descriptions of objects in the two rooms). The visitors' comments and conversation were recorded using wireless microphones, the visitors were videotaped by a camera placed in a corner of each room (Figure 2), the visitors were directly observed by the research escort, and the visitors' actions in the electronic guidebook were logged by the device for future reference.

After the visitors were finished in these two rooms, the research escort conducted a semi-structured interview to elicit visitors' reactions to their experience with the electronic guidebooks.[4]

No time limits were imposed on the visitors during any portion of the study. Visitors spent approximately 20-30 minutes using the electronic guidebooks in total and approximately 10-15 minutes in the interview. The entire procedure took approximately 75 minutes.

---

[4] The results of these interviews (which indicate that visitor response to the prototype was generally very positive) are discussed at length in [27].

Figure 3: Composite video of visitors and screens of their electronic guidebooks.

3.3 Analysis

The analysis reported in this paper primarily uses conversation analytic methods. A key goal of conversation analysis is to examine social interaction to reveal organized patterns or practices, under the fundamental assumption that interaction is structurally organized.

A conversation analytic research program involves analyzing a collection of interactive encounters that contain a specific practice that is visibly relevant for the participants producing it [23]. The analysis is twofold. First, each encounter is described in the detail of its moment-by-moment, turn-by-turn unfolding. Then these encounters are comparatively analyzed to reveal the practice's generalizable orderliness. The ultimate goal of a conversation analysis inquiry is to describe a practice in a way that is more general than the particulars of any one specific occurrence, yet specific enough to render "how to" instructions for the way it is accomplished.

In this study, our goal is to describe, in an empirical way, visitors' systematic practices as they use an electronic guidebook to tour a historic house with a companion. To identify these systematic practices, we examine in detail the data collected during selected visits. Specifically, we create a video that includes the audio and video recordings of the visitors, as well as video of the screens of each visitor's electronic guidebook (these screens are re-created based on the activity logs of the visitors' guidebooks). See Figure 3. The resulting data are transcribed and analyzed.

4 Patterns of Electronic Guidebook Use

In this section, we describe the usage patterns of the visitors. In general, the typical adult-adult visit consisted of two adults operating separate electronic guidebooks,

playing audio through the speakers on these guidebooks, chatting with each other, and frequently looking at objects in the room. The typical adult-child visit consisted of the child operating an electronic guidebook, playing audio through the speakers on the guidebook while the adult watched and listened, the adult and child chatting with each other, and the adult and child frequently looking at objects in the room. However, there were some variations in these patterns, so in this section we discuss visitors' usage patterns in more detail. We begin by discussing visitors' preferences for separate versus shared electronic devices and for delivery mode. We then discuss visitor attention and visitor-visitor engagement.

4.1 Separate versus shared electronic guidebooks

Each member of the five adult-adult pairs chose to use their own electronic guidebook. (One of these pairs began by sharing, but quickly decided that they would each like to use their own device.)

One adult-child pair shared a single electronic guidebook during the entire visit, with the child operating the device (playing audio descriptions through the speakers) and the adult looking on and making suggestions. The second adult-child pair began by using separate electronic guidebooks, but gradually evolved to the same shared model as the other adult-child pair. This evolution was initiated by the adult, who gradually stopped using his own electronic guidebook and began to listen to the audio from the child's electronic guidebook.

4.2 Delivery mode

Four of the five adult-adult pairs predominantly used audio played through the speakers of the electronic guidebooks. These visitors would frequently stand right next to each other playing different audio clips through the speakers. Visitors had a high tolerance for these overlapping audio clips. (Members of one pair discussed the audio, one member asking the other if it was bothering them, and both concluded that it was not a disturbance.) Visitors would also play audio clips for each other. The fifth adult-adult pair began by using audio played through the speakers of the device, but early in the session one of the members of the pair switched to headphone use (he stated this was because the audio from his companion's electronic guidebook was distracting him). His companion continued to listen to audio through the speakers of the device.

As mentioned above, both adult-child pairs used audio played through the speakers of a single device.

Visitors stated they preferred audio because, unlike text, it allowed them to look at objects while they learned about them. They also said the shared audio environment was more social (possible reasons for this are discussed further below). Visitors did say that they used text in a limited way for specific purposes, e.g., to see the spelling of a word, or to check whether they had heard something correctly. Additional discussion of visitor choice of delivery mode appears in [27].

Audio was also preferred in some cases because it seemed to be a more effective means of getting a companion's attention than reading text aloud. One visitor (J) initially read aloud information from text descriptions to her companion (L) but then resorted to audio mode when she realized that playing audio clips was a more effective way to get her companion's attention. When J switched to audio mode, she said, "[I] have to play the audio to get your attention, huh?" A similar phenomenon occurred with an adult-child pair (W and V): on multiple occasions the child began to read a text description to his father and was interrupted by his father's playing the corresponding audio description. This same pair briefly coordinated the use of two devices to simulate a combined text/audio mode, the father instructing the son, "You get the text and I'll get the audio."

4.3 Visitor attention

Visitors did interact heavily with the electronic guidebooks. Visitors selected an average of 37 descriptions, some selecting up to 69 (the latter indicating that visitors viewed or played some of the 42 descriptions more than once, e.g., either to listen to the content again individually, or to share it with their companion). While visitors did attend to the electronic guidebooks, they did not generally dominate visitors' attention. Visitors spent a great deal of time interacting with their companions and looking at objects in the room. More detailed discussion of visitor attention and the user interface appear in [4,27].

4.4 Visitor-visitor engagement

While we saw surprisingly little variation in areas such as the use of separate versus shared devices and the choice of delivery mode, we observed an extremely wide range of degree of interaction within the pairs. At one extreme, visitors stayed continuously engaged with each other throughout their visit, having a shared electronic guidebook experience, e.g., negotiating to choose a description and commenting on the descriptions once they heard them. At the other extreme, visitors rarely engaged with each other, focusing most of their attention on the electronic guidebooks and the objects in the room. Other visitors represented intermediate positions along the continuum, e.g., some visitors displayed a "rubberband" behavior in which they cycled between being engaged and disengaged. For example, J and L would look at objects and listen to descriptions independently for a while, but then one of them would share something with the other and they would become re-engaged. Further variation occurred because single pairs would often have a number of different behaviors, e.g., W and V started by using individual devices and then moved to a shared device.

The electronic guidebook helped visitors move fluidly from states of disengagement to states of re-engagement. For example, when a visitor's companion played audio through the speakers, that visitor could see/hear when their companion was listening to a description and therefore that visitor could identify or anticipate a moment when talk might be readily received by their companion. Further, audio

played through the speakers created a shared information context in which each visitor knew what the other had heard.

Within a given pair, the members would often have different behaviors. For example, one visitor would often be more likely to select a description, or one visitor would be more likely to orient their guidebook physically to their companion (so that their companion might easily hear it).

Recall that we observed visitors using paper guidebooks and then using electronic guidebooks. We have not yet conducted a formal analysis of visitor interaction while using the paper guidebooks. However, our preliminary observations indicate that electronic guidebook use generally preserved many of the pair's behavioral properties as compared with the paper guidebook use. The one exception was the pair in which one of the participants used headphones; while they interacted extensively when using the paper guidebooks, they had minimal interaction when one member used headphones with their electronic guidebook. This suggests that the use of shared audio does not interfere with the visitors' social interaction in the way that other technologies such as headphones do.

5 Storytelling Model

In this section, we show that the visitors' interaction with the electronic guidebooks had a storytelling structure. Visitors followed the conversational conventions of traditional storytelling in their interactions with the electronic guidebooks, treating the descriptions as stories. By following these conventions, visitors assigned the guidebook a conversational role. Further, they used the storytelling structure as a resource to physically and contextually *align* themselves so they could participate in fulfilling sharing episodes.

We now introduce the three sequentially structured phases of traditional storytelling: preface, telling, and response. Stories in conversation traditionally adhere to the following structure [22]: first, there is a *preface* for the story. In the preface, the storyteller requests and is granted the right to take an extended speaking turn while telling the story, since ordinarily a speaker would not be allowed to hold the floor for as long as a story requires. For example, a speaker may say something like, "I have a funny story" and a listener may say, "uh huh" or "go ahead." Second, there is the *telling* of the story itself. During the telling, the listeners will often make utterances that encourage the speaker to continue, or utterances that indicate a sympathetic response. Third, listeners make some appropriate *response* at the conclusion of the story, e.g., laughing.

In the following subsections, we discuss how the three phases are paralleled in visitor practices of guidebook use. Note that each phase can be located in the examples as follows: the conversation before the electronic guidebook description is the preface, the electronic guidebook description and any concurrent comments from the visitors comprise the telling, and conversation after the electronic guidebook description is the response.

In this section and the one that follows, it will be important to keep in mind that we are using the word "storytelling" to refer to a sequential interactional structure that is

Example 2: The story of Mrs. Roth's painting.

R: I want to know about that.
                (R points to the painting on R-PDA display)

G: About the painting?

R: Yeah. (R points to the painting in the room)

G: Yeah, okay. (G points to the painting in the room)

R-PDA: *This is a portrait of Mrs. Roth painted by her friend Lloyd Sexton, a well-known Hawaiian painter. In the photograph from which Sexton created the picture, she was holding daffodils. Sexton replaced the daffodils, which Mrs. Roth did not care for, with Queen Elizabeth roses of which she was very fond.*

G: (laughs) Diju- diju get that? They had a picture with daffodils so they made a painting and put roses in her hand, (laughs) that's neat, (laughs) I personally prefer daffodils.

well-defined in the context of conversational analysis. Storytelling often has other meanings in the HCI and cultural heritage communities; we discuss this further in Section 7.

5.1 Preface

During this phase, visitors attempted to come into alignment with each other's activities and to create a conversational place for the storytelling to occur. One of the key parts of this phase was deciding what object description[5] (story) to listen to. In this section, we begin by discussing verbal coordination around the choice of description. We then discuss how the displays of the devices are used as tools for coordinating object choice.

Visitors verbally coordinated choice of description

Coordination relating to choice of description ranged from minimal to elaborate. In some cases, coordination would occur before a particular description had been listened to by either participant, or in fact even chosen. One common occurrence was the joint choice of a description. For example, one visitor might ask another what they would like to listen to next. Often this would lead to the selection of a description that both visitors would listen to. Another common occurrence was one visitor announcing that they were about to listen to a description of a given object. Example 2 shows one such interaction, in which the grandchild (R) tells her grandmother (G) that she is about to select a painting. Note that coordination over choice of description often includes identifying the subject of the description in the

---

[5] Recall from Section 2 that visitors actually choose an object in the interface; this choice triggers a description associated with that object.

Example 3: The story of the false books.

	(V displays description of books in text mode)
W:	Wonder if there's a story on this ta- on the desk. (W switches PDA to fireplace wall)
V:	False books.
W:	Hm?
	(0.4)
V:	False books, heh heh.
	(1.2) (W selects an object on his PDA, but no information is available for that object)
W:	That's the chair. (W switches PDA to reception room wall)
V:	Oh yeah, these are false books, eh huh.
W:	Eheh, are they? (W walks around V towards bookcase on reception room wall)
V:	Yeah.
W:	Really?
	(0.4) (V switches his electronic guidebook to audio mode)
W:	Let's hear about it. (As W is speaking, V selects the description of the books)
V-PDA:	*Many of the top shelves contain false books. They are lighter than normal books, so they reduce the stress on the bookcases. Many are made of greeting cards, clothing, fabric, et cetera.*
W:	Eh hah, that's a riot. (W looks at V and smiles)
	(0.2) They're just for looks.

room, e.g., in Example 2, R and G both point to the physical painting that is about to be described.

In some cases, coordination of description choice would occur when one visitor (the initiator) encouraged their companion to listen to or read a description the initiator had just completed. Sometimes the initiator would share text snippets from the description. Other times they would make more general statements like, "Check out over here." Often the initiator would play the audio description for their companion, beginning the storytelling phase.

Example 3 shows a particularly strong effort to share a specific description. Visitor V reads a description of false books and then makes multiple attempts to share information from the description with his father (W). Although W is acting independently, searching for a description for a different object, V does eventually succeed in getting W's attention and plays the audio clip. As an aside, note that W refers to the guidebook's description as a "story."

In visitor interaction, the electronic guidebook did get a turn in conversation. If visitors were speaking during the preface phase, and the description began (signaling

the start of the telling phase), visitors generally stopped speaking, effectively allowing themselves to be interrupted by the guidebook.

Visitors used electronic guidebook displays to coordinate choice of description

The display of the device was a useful tool for coordinating the selection of descriptions, e.g., specifying the object that was about to be selected. Visitors most commonly demonstrated object selection or other activities on their own electronic guidebooks, rather than on their companion's. For example, if visitors were trying to get their companion to select an object, they would point to that object on their own electronic guidebook (or in the room) rather than on their companion's electronic guidebook. Sometimes this involved quite a bit of work.

For example, V was reading the textual description for an object. When his companion, W, wanted to get the same description on his device, V closed the window containing the textual description to bring up the object selection screen, pointed to the object, and then reinvoked the textual description. In other words, he opted to perform a series of relatively cumbersome operations on his device rather than simply pointing to the object on W's device. This was a fairly common interaction. By contrast, pointing to, touching, or otherwise operating a companion's device was extremely uncommon. For example, we only observed one or two instances in which a visitor touched their companion's device.

Presumably to facilitate these types of interactions, visitors would hold the electronic guidebooks and stand so that they could see each other's devices, e.g., a common pose was for one visitor to stand looking over the shoulder of another visitor. We also believe that largely differing heights may impact certain visitors, e.g., the child in the parent-child pair was apparently rarely able to see the screen on his father's device.

5.2   Telling

Often, an explicit preface for the story did not occur, or did not result in the participants being fully aligned at the beginning of the electronic guidebook's storytelling. This is entirely natural, since full coordination can be inefficient and perhaps uncomfortable in that it may preclude independent activity. Therefore, preface-related activities sometimes occurred during the telling. In this subsection, we first discuss how alignment attempts fail and discuss repairs that are done (or not done) to bring visitors into alignment. We then discuss other behaviors that occur during the telling, specifically focusing on physical behaviors for sharing audio and visitor comments that are interleaved with the guidebook's telling of the stories.

Launched stories prompted visitors to align

In many cases, visitors were fully aligned when the telling began. However, as mentioned above, sometimes visitors would choose a description and start to play it

without full agreement from their companion. In some cases, the companions would stay nearby and listen to this description, and in other cases, they would continue with independent activity.

In another type of phenomenon, visitors would often observe their companion listening to or reading a description and attempt to participate. For example, visitors might walk across the room to examine an object while listening to a description. In this case, their companion would sometimes follow them to listen to the description and comment. Visitors would also eavesdrop on each other's devices. Sometimes when visitors heard an object described by their companion's guidebook, they would select that description on their own device. Visitors would also explicitly ask other visitors what they were doing, e.g., J approached L (who was listening to a description) and said, "Okay, what'd you find?" J then chose the description of the same object on her electronic guidebook. Visitors generally seemed open to all these types of approaches or mimicry of their behavior. We did not observe any behaviors or verbal exchanges that suggested privacy issues.

Additionally, once an object had been selected for description, we often saw further attempts to establish a shared understanding of which object was being described, e.g., the person who selected the description would often point at the corresponding object in the room during the description.

Visitors physically shared electronic guidebooks during the telling

The desire to hear the audio on a companion's electronic guidebook (or to have a companion hear the audio on one's own electronic guidebook) affected visitor behavior in several ways.   (Recall that visitors had a choice of reading text descriptions or listening to them through headphones or through speakers on the device at a low volume, and that in the study most descriptions were played through speakers on the device at a low volume.) To share descriptions, visitors generally stood close together.

If either visitor attempted to move away while the description was playing, visitors had to coordinate their positions to maintain a shared audio context. In some cases, one visitor would attempt to move away while the description was still playing, e.g., to investigate the object being described. If the visitor who moved away was not the one holding the guidebook playing the audio, they would often quickly return to the side of the visitor playing the description. If the visitor who moved away was the one playing the description, the other visitor would employ one of a number of behaviors to stay involved in the activity. For example, we saw adults put their hands on children's shoulders to prevent them from moving away. One parent (W) gave a general instruction to his child: "If you're gonna play it, [then] stay near so I can hear it." (This was rare – almost all communication around this topic was non-verbal.) As mentioned above, people also frequently followed their companions who were playing audio descriptions.   Some visitors also showed an awareness of their companion's needs while playing audio descriptions.   For example, while a description was playing, a parent walked from one side of his child to the other so that he could more clearly see the object being described; while he was moving, the parent held his electronic guidebook to his child's ear so that his child's listening would not
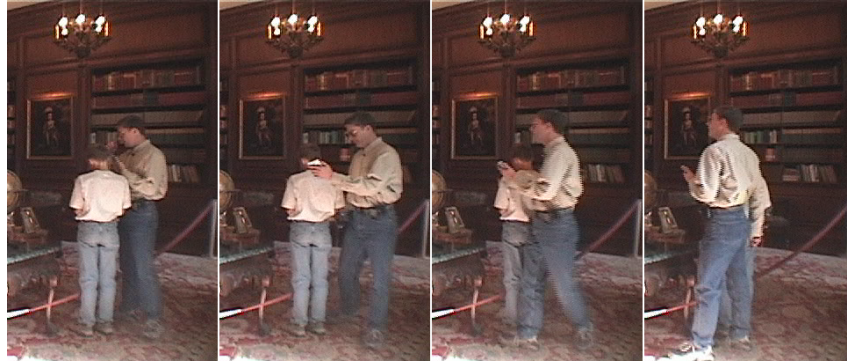
Figure 4: W moves from one side of V to the other to get a better view of the object being described. As W moves, he brings the guidebook near V's ear so V's listening will not be interrupted.

be interrupted. See Figure 4. Similarly, on one occasion V moved away from W to approach an object. As V walked away, he switched the electronic guidebook from his left hand to his right hand, which brought it closer to W (who was on his right side).

Visitors interleaved comments with the story

While the audio was playing, visitors would respond to the electronic guidebook as they might to a storyteller, e.g., they would utter affirmatives such as "uh huh" or brief responses such as "interesting." Such interleaved comments were made at "appropriate" points in the electronic guidebook's utterances, e.g., pauses between sentences. Visitors did not generally make lengthy comments while the guidebook was delivering a description, i.e., they did not generally interrupt it. Visitor tolerance may have been improved by the fact that the descriptions are quite short (the longest being 23 seconds).

In Example 4, A and S listen to a description of the stain pattern on the floor. Notice how they interject brief comments while the description is playing. Also note that these comments are neatly aligned with sentence boundaries, as they would be for a human speaker.

5.3 Response

The crucial moment in the story is the moment of response. Visitors went to a great deal of effort during the preface and telling phases to make sure they were positioned to share a response to the story. At the conclusion of a description, visitors very frequently shared reactions. For example, listeners made comments like "wow" or offered an opinion or made a joke. In the examples in this paper, we see some of the possible behaviors. In Example 1, A and S discuss the extravagance of the wedding

Example 4: The story of the floor.

S: (0.4) There, floor again, that's floor again.

S-PDA: *The floor is made of oak wood carved with gouging planes.* [*After the gouges were made, stain was*

A: [yeah

S: [that's, that's interesting,

S-PDA: *applied and puddled in the holes.*

A: Wow.

S: Hm, [that's interesting.

S-PDA: [*This technique was popular in the early nineteen hundreds when Filoli was built. The effect wears away with use, as can be seen in the floor by the door to the room.*

S: Ehm uh huh, see here it's darker there, kind of interesting.

A: (Points towards corner) By you.

S: Heh.

gift. In Example 2, G retells the story, presumably because R does not laugh. This type of retelling or elaboration in the absence of an appropriate response is another common phenomenon in storytelling. In Example 3, W laughs and comments that the information conveyed is "a riot." In Example 4, A and S study and discuss features of the object described.

The guidebook descriptions also prompted visitors to tell their own stories. For example, a discussion of how the family spent Christmas at Filoli led J to remark on an upcoming Thanksgiving celebration of her own. When studying charcoal sketches of previous owners of Filoli, a child (R) volunteered that her own sketch had been done by an artist at the mall.

Regarding physical space, after the description was completed, one of the visitors would often step away from the other, sometimes to go investigate the object which was being described (this investigation would sometimes be followed by a reaction), or sometimes apparently simply to create slightly more space between the two visitors (while listening to a description, visitors often stood very close together).

6 Discussion

Having described the fit between the observed behavior and the storytelling model, we now summarize the key points resulting from the application of the model to the data. First, we discuss how visitors became aligned so they could share the storytelling experience. We then discuss how the electronic guidebook functioned in a conversational role. We next talk about ways in which visitors benefited from the shared storytelling experience. Finally, we discuss the implications of the visitors' desire to share.

6.1 Alignment

In many cases, visitors went to a great of effort to become aligned so they could hear the same content at the same time:

- Visitors chose a delivery mode that facilitated a shared audio experience (audio through speakers). This delivery mode also facilitated re-engagement by revealing conversation availability and by making it easy to eavesdrop.
- Visitors negotiated the choice of descriptions. Some of these negotiations were fairly mutual. Others were *push* behaviors in which one visitor would encourage their companion to join them in an activity (e.g., "Check out over here."). Still other behaviors were *pull* behaviors in which one visitor would try to join another visitor (e.g., "What did you find?"). Push and pull behaviors could be verbal or non-verbal, e.g., pointing to an object in the room was a frequently observed push behavior.
- Visitors oriented their electronic guidebooks to their companions and oriented themselves to their companion's devices, e.g., they held their guidebooks so their companions could see/hear them, and they stood so they could see/hear their companion's devices.

Note that visitors did not want to be fully aligned at all times. Visitors often disengaged and conducted activities independently. Visitors were also very articulate about wanting to use separate electronic guidebooks. Visitors' desire for control over their experience is discussed further in [27].

6.2 The conversational role of the electronic guidebook

By analyzing the structure of the storytelling interactions, we see that the electronic guidebook was allocated a role in the conversation, e.g., the electronic guidebook was granted turns in the conversation, it was allowed to introduce topics of conversation, and visitors verbally responded to it. This affordance of the electronic guidebook (that it may be treated as a conversational interactant) allowed visitors to integrate the electronic guidebook in their naturally occurring conversations. By including the electronic guidebook in their existing interaction, visitors were able to incorporate it without fundamentally changing the way that they interacted with each other, e.g., the degree to which they engaged with each other was similar whether they were using a paper guidebook or an electronic guidebook. This preservation of fundamental interaction did not occur when headphones were used.

6.3 Shared response

By integrating the electronic guidebook in their conversation, visitors became aligned so they could have shared responses to electronic guidebook content. The shared response was particularly enabled by the conversational structure of the storytelling; this structure by definition includes a place for a response, a participatory moment for

the visitors. Visitors used this moment to share their reactions to information, and they also used the content of the stories as a branching-off point for their own stories. These are gratifying social interactions that are not possible with all designs, e.g., headphones preclude them. Further, the shared responses positioned the visitors to launch new storytelling sequences with the electronic guidebooks, thereby leading to more social interaction.

6.4 Drive to share

As seen in this study, visitors found ways to incorporate an electronic device in their social interaction. This gave them a more fulfilling experience. Perhaps designing devices that simulate properties of conversational interaction (e.g., real human voices, short turns to promote opportunities for humans to turn-take) can help users of mobile devices integrate these devices into their existing interactions with other people.

Further, visitors manipulated themselves and their devices so they could see and hear the contents of each other's devices, e.g., they held them in unnatural positions and positioned themselves so they could more easily see their companion's devices and vice versa. This suggests that device design should facilitate sharing of audio and visual content.

7 Related Work

Our goal in this project has been to improve visitor experience as measured against essentially self-perceived motivations. This is along the lines of those who assess the quality of learning-oriented leisure activity [8,12]; our evaluation did not focus on, e.g., the learning environment [10] or the aesthetic experience [11] per se. (Of course, one of our design goals was to minimize the amount of work the visitors needed to do to gather information in order to maximize their opportunity to both learn about and appreciate the objects.)

With that in mind, our work can be compared to specific previous efforts in the engineering domain, the museum studies domain, and their intersection. We discuss each in turn. We then (briefly) relate our work to more general discussions of the social aspects of technological artifacts.

7.1 Electronic guidebook design

Electronic guidebook products include a wide variety of systems from industry leaders Acoustiguide and Antenna Audio, as well as from other vendors such as Ameritech (smARTour), JVC (Audio Guidance System), Organic (eDocent), Visible Interactive (iGo), and Vulcan Northwest (Museum Exhibit Guide). Many research systems have also been built (see, e.g., [6-7,13,16-19,21]).

Our system differs from previous systems in its reliance on a lightweight visual interface based on photographic images [4]. It also differs in its use of independent

navigation mechanisms for different stages of the object selection task [3]. Finally, it provides the option of either text or audio presentation (through headphones or speakers) of identical content.

7.2 Museum studies

Some of the relative advantages of audio and text presentation are well understood in the cultural heritage community. For example, Serrell notes that audio allows simultaneous use of eyes and ears but tends to isolate the listener [25]. Acoustiguide's marketing literature states that research "based on a series of surveys at client sites... proves that visitors who access Acoustiguide interpretations learn more about exhibitions – and enjoy them more" [1], and some of this is borne out in the academic literature (e.g., [24]). However, we believe that our observations about short, conversationally compatible audio clips are novel and are not obviously predictable from studies of, e.g., short text labels [25]. Our results on visitor usage of individually controlled, sharable audio are new as well.

In a study of exhibit label reading, McManus observed high rates of "text echo," the inclusion of label text in conversation [14]. The widespread uses of audio sharing in our study (both deliberate and eavesdropped) demonstrate that technology can be used to help visitors introduce label content into conversation directly. McManus also suggested that visitors process – and are inclined to treat – exhibit labels *as conversation*. Again, our findings indicate that technology can bring visitor experience even more in line with this inclination, particularly if the audio descriptions are short enough to easily integrate in existing conversations with companions.

Vom Lehn *et al*. have examined visitor interaction in museums using a qualitative methodology similar to the one used in this paper [26]. While they are studying interactional patterns, their work does not have a personal technology component analogous to an electronic guidebook.

7.3 Electronic guidebook studies

A variety of research systems have been designed, built, and deployed but few have resulted in in-depth studies. For example, HyperAudio [16] was deployed but only results of pre-design studies have been reported [19]. Similarly, Hippie [17] was deployed and received initial feedback, but the results of user evaluations are not available [18]; the same is true of Plantations Pathfinder [21]. We are aware of only two electronic guidebook studies resembling ours. In both cases, the methodology was, like ours, based on a combination of interviews, observation, and device activity log analysis. A University of Salford team evaluated the design of a tablet computer guidebook prototype at the Museum of Science & Industry in Manchester [7], and a Lancaster University team evaluated the design of another tablet computer guidebook prototype in historic Lancaster [6].

Unlike previous work, our study takes an applied conversation analytic approach to examining how visitors interact with electronic guidebooks, and how visitor-visitor interaction is impacted by electronic guidebook design.

7.4 Social properties of technological artifacts

Reeves and Nass argue that individuals' interactions with computers, television, and new media are fundamentally social [20]. Our focus is somewhat different, since we argue that integration of the electronic guidebook into conversation facilitates interaction among participants. Further, we are studying the process by which this interaction becomes social. Finally, where Reeves and Nass conducted their studies in laboratory conditions, we are studying natural behaviors in largely unrestricted environments.

The Campiello project designed and evaluated a community information service for the use of both tourists and local residents in Venice [9]. In a month-long deployment study, careful design around social use aspects (e.g., support for group use of kiosks) proved to be a critical factor in system acceptance [2].

Depending on the context, "storytelling" can mean many things. For example, it is widely accepted in the museum studies community that visitors retain more information if the information is related as a cohesive story [8], which is why so many exhibits are structured in this way. As another example, much of the work on children and computing has focused on mechanisms for the creation and effective social use of stories. Our emphasis here is on storytelling as a framework of conversational roles and sequentially organized acts, on the ways that our visitors found to work an artifact (the guidebook) into this framework, and on the aspects of the artifact that facilitated this act.

8 Conclusions and Future Work

Our electronic guidebook prototype had a substantive impact on visitors' ability to interact with each other. Visitors predominantly chose audio played through speakers, which allowed them to include the electronic guidebook in their conversation with their companions. We found that these interactions followed the structure of storytelling. As a result, visitors were positioned to share moments in which they responded to stories told by the electronic guidebooks. These interactions led to experiences that were significantly more social than those that occur with traditional headphone audio tours.

Our current work includes further analysis of the data collected in the course of this study as well as application of our findings. For example, we are studying what drives a visitor's inquiry into a particular object (e.g., independent observation of the object in the room, a companion's interest in the object, or the presence of a description in the guidebook). We are also applying the iterative design approach [15], using the lessons learned about sharing to design our next prototype.

9 Acknowledgments

References

1. "Why an Audio Tour?" Acoustiguide Corp., New York, Dec. 2000. http://www.acoustiguide.com/why/

2. Agostini, A., G. De Michelis and M. Divitini, "Ubiquitous Access to Community Knowledge Via Multiple Interfaces: Design and Experiences," in *Universal Access in HCI* (Proc. 1st Conf., New Orleans, LA, Aug. 2001), C. Stephanidis (ed.), Lawrence Erlbaum & Assoc., 2001, to appear.

3. Aoki, P.M. and A. Woodruff, "Improving Electronic Guidebook Interfaces Using a Task-Oriented Design Approach," *Proc. 3rd ACM Conf. on Designing Interactive Sys.*, New York, Aug. 2000, 319-325.

4. Aoki, P.M., A. Hurst and A. Woodruff, "Tap Tips: Lightweight Discovery of Touchscreen Targets," *ACM SIGCHI Conf. on Human Factors in Comp. Sys. Extended Abstracts*, Seattle, WA, Mar. 2001, 237-238.

5. Beard, G.W., *Attingham: The First Forty Years*, Attingham Trust, London, 1991.

6. Cheverst, K., N. Davies, K. Mitchell, A. Friday and C. Efstratiou, "Developing a Context-Aware Electronic Tourist Guide: Some Issues and Experiences," *Proc. ACM SIGCHI Conf. on Human Factors in Comp. Sys.*, den Haag, the Netherlands, Apr. 2000, 17-24.

7. Evans, J.A. and P. Sterry, "Portable Computers & Interactive Museums: A New Paradigm for Interpreting Museum Collections," *Proc. 5th Int'l Cultural Heritage Informatics Mtg.*, Washington, DC, Sep.1999, 93-101.

8. Falk, J.H. and L. D. Dierking, *Learning From Museums*, Altamira Press, Walnut Creek, CA, 2000.

9. Grasso, A., D. Snowdon and M. Koch, "Extending the Services and the Accessibility of Community Networks," in *Digital Cities* (Proc. Kyoto Meeting, Kyoto, Japan, Sep. 1999), T. Ishida and K. Isbister (eds.), Springer Verlag, Berlin, 2000, 401-415.

10. Hein, G.E., "The Constructivist Museum," *J. Educ. in Museums 16* (1995), 21-23.

11. Hein, H.S., *The Museum in Transition*, Smithsonian Institution Press, Washington, DC, 2000.

12. Hood, M.G., "Staying Away: Why People Choose Not to Visit Museums," *Museum News 61*, 4 (Apr. 1983), 50-57.

13. S. Long, D. Aust, G. D. Abowd and C. Atkeson, "Cyberguide: Prototyping Context-Aware Mobile Applications," *ACM SIGCHI '96 Conference Companion,* Vancouver, BC, Canada, April 1996, pp. 293-294.

14. McManus, P.M., "Oh, Yes They Do: How Museum Visitors Read Labels and Interact with Exhibit Texts," *Curator 32*, 3 (1989), 174-189.

15. Newman, W.M. and M.G. Lamming, *Interactive System Design*, Addison Wesley, Reading, MA, 1995.

16. Not, E., D. Petrelli, O. Stock, C. Strapparava and M. Zancanaro, "Person-Oriented Guided Visits in a Physical Museum," *Proc. 4th Int'l Cultural Heritage Informatics Mtg.,* Paris, France, Sep. 1997, 69-79.

17. Oppermann, R. and M. Specht, "A Nomadic Information System for Adaptive Exhibition Guidance," *Proc. 5th Int'l Cultural Heritage Informatics Mtg.*, Washington, DC, Sep. 1999, 103-109.

18. Oppermann, R. and M. Specht, "A Context-Sensitive Nomadic Exhibition Guide," in *Handheld and Ubiquitous Computing* (Proc. 2nd Int'l Symp., Bristol, UK, Sep. 2000), P. Thomas and H. W. Gellersen (eds.), Springer Verlag, Berlin, 2000, 127-142.

19. Petrelli, D., A. De Angeli and G. Convertino, "A User-Centered Approach to User Modeling," in *User Modeling* (Proc. 7th Int'l Conf., Banff, Alberta, June 1999), J. Kay (ed.), Springer Verlag, Berlin, 1999, 255-264.

20. Reeves, B. and C. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, CSLI Publications, Stanford, California, 1996.

21. Rieger, R. and G. Gay, "Using Mobile Computing to Enhance Field Study," in *Computer Supported Collaborative Learning* (Proc. 2nd Conf., Toronto, Ontario, Dec. 1997), R. Hall, N. Miyake and N. Enyedy (eds.), L. Erlbaum & Assoc., Mahwah, NJ, 1997, 215-223.

22. Sacks, H., "An Analysis of the Course of a Joke's Telling in Conversation," in *Explorations in the Ethnography of Speaking*, R. Bauman and J. Sherzer (eds.), Cambridge University Press, Cambridge, 1974, 337-353.

23. Sacks, H., "Notes on Methodology," in *Structures of Social Action: Studies in Conversation Analysis*, J.M. Atkinson and J. Heritage (eds.), Cambridge University Press, Cambridge, 1984, 21-27.

24. Screven, C. G., "The Effectiveness of Guidance Devices on Visitor Learning," *Curator 18*, 3 (1975), 219-243.

25. Serrell, B., *Exhibit Labels*, Altamira Press, Walnut Creek, CA, 1996.

26. vom Lehn, D., C. Heath and J. Hindmarsh, "Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries," *Symbolic Interaction 24* (2001), to appear.

27. Woodruff, A., P.M. Aoki, A. Hurst, and M.H. Szymanski, "Visitor Attention and Electronic Guidebooks," *Proc. 6th Int'l Cultural Heritage Informatics Mtg.,* Milan, Italy, Sep. 2001, to appear.