Books with Voices: Paper Transcripts as a Tangible Interface to Oral Histories

Scott R. Klemmer [1] srk@cs.berkeley.edu
Jamey Graham [2] jamey@rii.ricoh.com
Gregory J. Wolff [2] wolff@rii.ricoh.com
James A. Landay [1] landay@cs.berkeley.edu

[1] Group for User Interface Research
Computer Science Division
University of California
Berkeley, CA 94720-1776, USA

[2] Ricoh Innovations
2882 Sand Hill Road, Suite 115
Menlo Park, CA 94025-7064 USA

ABSTRACT
Our contextual inquiry into the practices of oral historians unearthed a curious incongruity. While oral historians consider interview recordings a central historical artifact, these recordings sit unused after a written transcript is produced. We hypothesized that this is largely because books are more usable than recordings. Therefore, we created Books with Voices: bar-code augmented paper transcripts enabling fast, random access to digital video interviews on a PDA. We present quantitative results of an evaluation of this tangible interface with 13 participants. They found this lightweight, structured access to original recordings to offer substantial benefits with minimal overhead. Oral historians found a level of emotion in the video not available in the printed transcript. The video also helped readers clarify the text and observe nonverbal cues.

Keywords
tangible interface, interactive paper, video retrieval, reading, augmented reality, handheld, oral history

INTRODUCTION
"Oral history is primary source material obtained by recording the spoken words—generally by means of planned, tape-recorded interviews—of persons deemed to harbor hitherto unavailable information worth preserving" [25]. The discipline began in 1948 when Allan Nevins founded the Columbia University Oral History Research Office.

Our experiences reading, doing contextual inquiry, and conducting oral histories led us to develop Books with Voices: bar-code augmented paper transcripts enabling fast, random access to digital video interviews on a PDA. Members of the Regional Oral History Office (ROHO) at UC Berkeley [1] participated in our design process. We showed nine different mock-ups of bar-code augmented oral histories to twelve members of ROHO. They were excited to hear as they read and encouraged us to make our paper book design more usable.

We evaluated Books with Voices (see Figure 1 and 2) with 13 users: eight oral historians from ROHO, one oral historian from the University of San Francisco, and four members of the Computer Science Graduate Student Book Club. The response from participants was overwhelmingly positive. The calm interface allows users to concentrate on the task and stay in the flow. Participants repeatedly accessed recordings while reading. They did so to: a) get a sense of the personality of the interviewee, b) hear the tone of a particularly compelling passage, and c) verify the accuracy of the transcript. We found several fixable usability issues, most notably that the PDA must be rotated into an awkward vertical position to scan. Motivated partially by our system, ROHO has begun transcribing their oral histories digitally.

CHI 2003, April 5–10, 2003, Ft. Lauderdale, Florida, USA.

FIELDWORK
We undertook a three-pronged approach to better understanding the discipline of oral history. First, we read oral historians' reflections on practice (e.g., [4, 8, 10, 25]). Second, we conducted a contextual inquiry at ROHO. Third, we experienced the process by conducting oral histories with two well-known computer science professors.

Through our fieldwork and literature review, we discovered a curious incongruity. While oral historians consider the audio or video recording a central historical artifact, these recordings sit unused after a written transcript is produced. Louis Starr wrote about why transcripts dominate use:

> "This is not so much because those who favor the transcript have the better of the argument of theoretical grounds as because of practical convenience: to most researchers, a written document that carries page numbers, and an index to them, is vastly preferable. Tapes, no matter how carefully indexed, are awkward to use, particularly if the memoir is a massive one. … A consensus emerges: tapes are more suitable for some purposes, transcripts for others; but so far as possible both should be preserved, allowing researchers to choose for themselves" [25].

Mackay also found people's preference for paper, writing:

> "Contrary to what many believe, users are not Luddites, clinging to paper as a way of resisting change. On the contrary: most are excited by the benefits offered by computers… Their resistance is, in fact, extremely practical. New computer

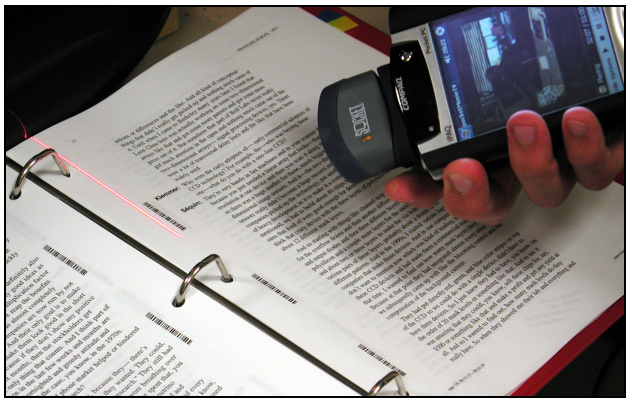

Figure 1. Accessing digital video by scanning transcripts.

> systems are either less efficient or simply cannot perform many required tasks" [17].

Suchman writes about how participants understand conversation, stating "contextualization cues by which people produce the mutual intelligibility of their interaction consist in the systematic organization of speech prosody, body position and gesture, gaze, and the precision of collaboratively accomplished timing" [27]. These contextualization cues are more available in a video recording than a written transcript.

The written transcripts of these interviews are a wonderful artifact. But they lack a humanity that is available in the videos through the interviewee's body language and prosody.

Contextual Inquiry at ROHO

In our research, we spent time at ROHO, the third oldest oral history office, and one of the largest. Interviewers are generally knowledgeable about the area in which they interview. For example, ROHO is documenting the "Disability Rights and Independent Living Movement" (DRILM) [3]. One member told us, "All of the project interviewers have personal experience with disability." In other cases, the interviewer is well-versed in the history of the area but is not a direct participant. Interviewers generally prepare with two to three weeks of background reading on the subject area and on the writings by or about the interviewee.

A typical oral history consists of between four and twenty hours of oral interviews. A rule of thumb is that each interview session should last roughly an hour and a half. Some historians bring a still camera to photograph the interviewee or objects important to them. Afterwards, interviewees are generally sent a preliminary transcript for review.

Isolated interviews are rare. Usually, a small group works together to create a set of 12 to 20 interviews about a subject. With DRILM, seven interviewers and two managers conducted oral histories with more than forty interviewees.

Oral historians in general are highly enthusiastic about video and other digital technologies. Traditionally, interviews have been recorded on analog audio cassette. At

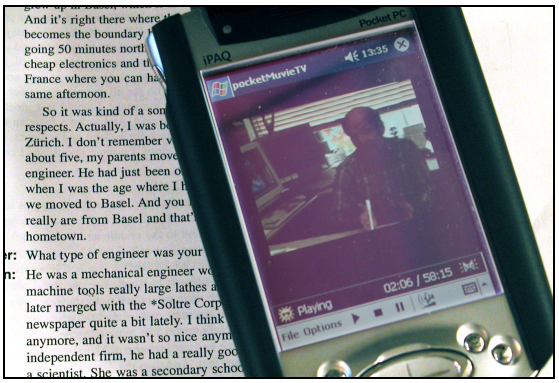

Figure 2. PDA playing a video of a recorded oral history.

ROHO, digital MiniDisc is rapidly replacing cassettes. Digital video is gaining popularity in the oral history community at large (as judged by traffic on the main oral history mailing list [4]), and ROHO is beginning to use it in their work.

The main barrier to adoption is that few historians feel they have the time to learn a new technology. One particularly compelling aspect of our system is that it augments, rather than replaces paper. Books with Voices paper transcripts offer the same familiar affordances as current transcripts because both *are* ordinary paper. The user population is not required to change how they read. As users recognize the benefits of our system, they *can* adopt it, but they are not *forced* to give up their current practices.

Conducting Oral Histories

To better understand oral history practice, the first author attended an oral history training workshop, and then conducted oral histories with computer science professors David Patterson and Carlo Séquin. He conducted three interviews with Prof. Patterson (four hours and 77 pages worth) and four interviews with Prof. Séquin (six hours and 114 pages worth). These interviews were recorded on digital video and converted to MPEG-2. The digital files were then professionally transcribed with time-stamps corresponding to each utterance.

RELATED WORK

We now discuss research with related technologies (paper and tangible interfaces), domain (oral histories), or tasks (reading and video retrieval).

Paper as a Tangible Interface

Wellner's seminal Digital Desk [29] inspired our interest in paper-based interfaces. Books with Voices differs from the Desk and much current research in that it employs paper as an *archival,* rather than *ephemeral,* artifact.

PARC: Experiments in the Future of Reading

The PARC Experiments in the Future of Reading project offers several insights. Most related to our work is the Listen Reader [5], a system combining the look and feel of a real book with an interactive soundtrack. Each RFID-tagged page has a unique soundtrack modified by the user's hand position. Hand tracking is accomplished via capacitive sensing. This coordination of reading and listening is highly compelling.

Mackay and colleagues have built several excellent paper interfaces [17]. Video Mosaic is a Digital Desk system where the user edits and rearranges digital video using paper storyboards. The Caméléon project recognized that attempted computerization of air traffic control flight strips had failed. Mackay and colleagues approached the problem in a radically different and empathetic way, augmenting these paper flight strips with RFID tags.

The Tangible Media Group at MIT has been a leader in tangible interfaces. Most relevant to our work is the mediaBlocks system [28], where RFID tagged wooden blocks serve as pointers to electronically stored video segments.

Bar-codes and Glyphs

Bar-codes, often the cheapest and easiest method for tagging physical objects, "have been used in packaging since 1974, when the first item, a pack of chewing gum, was scanned at a supermarket in Ohio" [18]. Reappropriating that technology for HCI, Johnson and colleagues introduced the idea of a "paper interface" [13]. Their system prints a 2D bar-code (in this case, a glyph) onto paper printouts, such as order forms, enabling users to "complete the document loop" and electronically trigger workflow actions. This work illustrates the utility of bar-codes as a lightweight augmentation mechanism.

Glyphs are a highly flexible 2D bar-code. In their most recent incarnation [2] they can even be embedded into text or images. However, they require a high quality image scanner. CyberCode is a similar type of 2D bar-code. CyberCodes [20] are coarser than glyphs, enabling recognition by low-cost imaging hardware such as the cameras on mobile devices. We may use CyberCode in our next prototype because cameras are smaller, cheaper, and more common than bar-code readers.

Recently, there have been several research and commercial systems that use bar-codes as links to web site URLs (e.g., Cooltown [14]). WebStickers [16] are shareable physical handles (Post-it notes with bar-codes) to electronic resources. Users can associate a note with a web URL, and later scan the note to retrieve the URL.

Palette [19] provides bar-code-tagged index cards to give users control over their presentation's slide order. Users print their presentation onto index cards—one slide per card. An evaluation [7] found the main drawback to be that "People worked on presentations 'until the last minute' and did not 'have time to print cards.'" In general, paper interfaces are most successful when the electronic content is stable (e.g., in Books with Voices), or when a new electronic version does not require a new paper version.

Technology Support for Oral Histories

The topic of the May 2002 issue of *Forum: Qualitative Social Research* [10], an online journal, is "Using Technology in the Qualitative Research Process." It offers an excellent overview of current work in the field and shows social scientists' strong interest in using technology in their work practices.

Palaver Tree [9] is web-based educational software enabling middle school students to conduct email oral histories with elders (such as civil rights leaders). This work inspired us to research oral histories. While Palaver Tree's contribution is primarily pedagogical, our contribution is primarily technological.

Steven Spielberg's *Survivors of the Shoah Visual History Foundation* has "collected more than 50,000 eyewitness testimonies in 57 countries and 32 languages" from Holocaust survivors [24], a huge collection. Researchers at the University of Maryland are addressing the information retrieval issues in this archive (particularly multilingual access), and have conducted ethnographic studies detailing the needs of researchers using this archive [12]. Many other oral history web sites have some of their oral histories online; [6] has an excellent list in the appendix.

Reading, Listening, and Video Browsing

One of the most compelling electronic book systems is XLibris [22]. XLibris, and e-books in general [23], enable more rapid searching of text and more fluid sharing of annotations. However, "there is a tension between the advantages provided by computation and the advantages provided by paper: the choice depends on the reader's goals" [22]. We chose paper over e-books for our domain because "People clearly prefer reading on paper to reading on their PCs" [23].

Audio Notebook [26] and Dynomite [30] inspired our interest in using text as an interface to streamed media. Audio Notebook is a paper notebook that sits on top of an ink and audio capture device. As note-takers write, ink is time-associated with audio recorded at that time. Dynomite offers similar functionality, with an electronic notebook as opposed to a paper notebook. These projects compellingly show the use of one modality (written notes) as a query interface to another modality (recorded audio).

One aspect of our work is providing an interface to pre-recorded video. In an evaluation of a digital video browsing system [15], the authors found that "the most frequently used features were time compression, pause removal, and navigation using shot boundaries." While these interaction techniques become less important in the presence of a full transcript, it would certainly benefit our system to incorporate them.

PAPER PROTOTYPE OF A PAPER INTERFACE

In our design, bar-codes augment the paper transcript, providing a physical affordance for fast, random access to digital video recordings. As our first design step, we produced nine different paper mock-ups, brainstorming the visual design of these augmented paper reading artifacts. We based these mock-ups on the print format that ROHO currently uses, adding bar-codes and video stills to the sides of the text and metadata to the header and footer (oral history title, interview date, and time-code information).

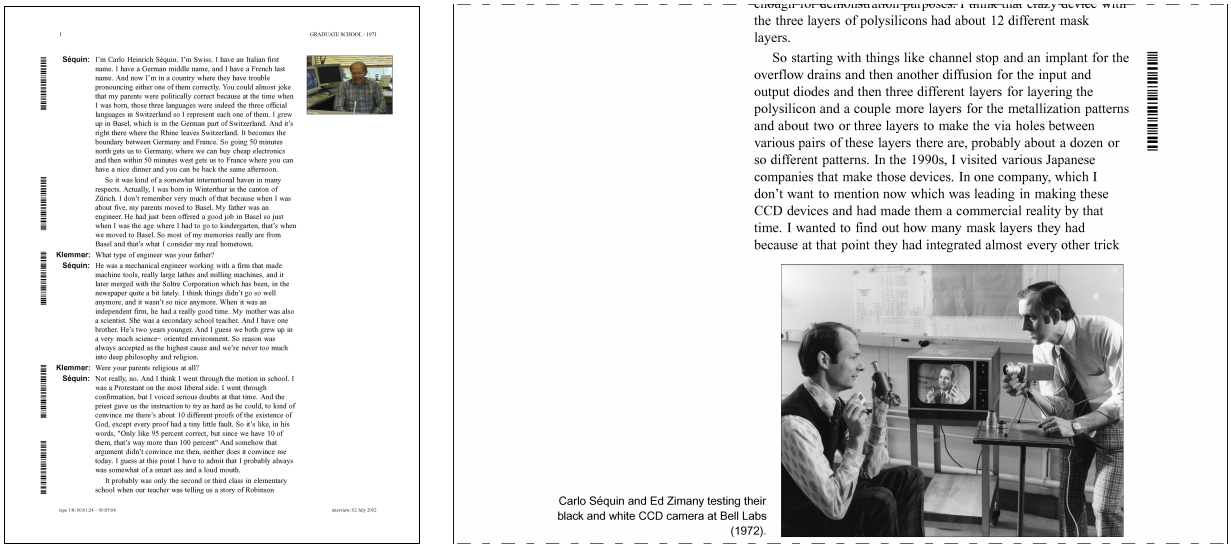

Figure 3. Books with Voices augmented paper transcripts: a complete first page (left), and detail of an internal page (right).

We showed these mock-ups to twelve members of ROHO: the director, the head transcriber, several interviewers, the technology director, and two historians that use but do not produce oral histories. People spoke passionately about the importance of hearing the original voice. (With one exception: One woman felt that she was a reader only, and not a listener.) This echoes the feelings of other oral historians such as this H-OralHist list posting, "Each speaker's voice is so distinctive! I've found that sometimes even when two speakers seem to be saying something very similar, their intonations can indicate subtle differences in meaning that can complement each other. … The unedited tapes are much more lively and interesting although of course it takes much longer to listen to them than to skim transcripts" [21]. The participants appreciated the metadata as well.

The participants encouraged us to address two issues in the next iteration:

1) For this domain, people did not want a lot of video stills in the printed book. These "talking head" stills are nearly identical; repeated printing would offer little value. One person suggested that one video still at the beginning of each section would be about right, and others agreed that this made sense. People also encouraged us to incorporate photographs and other media.
2) People really liked the bar-code access, and about 3 bar-codes/page seemed to make sense. Most of our mock-ups had bar-codes that were evenly spaced (e.g., top, middle, and bottom). The users did not like this; they specifically asked that bar-codes be visually aligned at speaker turn and paragraph boundaries (see Figure 3).

INTERACTIVE PROTOTYPE

We used the feedback we received on the mock-ups to build an interactive Books with Voices system. We built it by modifying the Video Paper software [11]. Video Paper is a paper-based interface for browsing, retrieving, and viewing pre-recorded video. We used three parts of the Video Paper system: 1) Creating an MPEG-2 video (at 20fps, 208x160 pixels) and making JPEG thumbnails from a video source; 2) Creating a paper layout (see Figure 3) from a time-stamped transcript; and 3) Pocket PC software that reads the bar-code and plays the corresponding video. We modified the rendering engine (2) to produce a document more suitable for our domain.

Hardware

Books with Voices requires a PDA with: 1) a display screen, 2) audio output, 3) a scanning device (such as a bar-code reader or digital camera), and 4) access to a video store (we use a 2GB PC card hard drive on the iPAQ; this could also be achieved by wirelessly streaming video to the device.)

There are four categories of technologies that are appropriate for recognizing paper:

1. Passive electronic tags (e.g., RFID tags)
2. Active electronic tags (e.g., motes)
3. Bar-codes (includes 2D variants such as Glyphs)
4. Vision-based content analysis (e.g., OCR)

Our inquiry found photocopying to be an important historical research practice, making tagging inappropriate. Tagging would also make book production more expensive and time-consuming. OCR and bar-codes work with photocopies and require no additional materials or time. Bar-codes are preferable for our domain because they are more reliable and the interaction is simpler.

INTERACTIVE PROTOTYPE EVALUATION

To understand both professional and amateur use, we evaluated the utility of Books with Voices as an augmented reading tool with thirteen users: eight oral historians from ROHO (two history professors, three editors, one interviewer/editor, and two transcribers), one oral historian from the University of San Francisco (a history professor), and four members of the UC Berkeley Computer Science Graduate Student Book Club.

Figure 4. The trigger button initiates bar-code scanning.

We conducted this study to learn a) if and when oral historians find video access valuable, and b) how suitable our interaction techniques are for oral histories.

Study Design

Each session lasted between 90 minutes and two hours. The studies were conducted in the participants workplaces. We videotaped the sessions and took time-coded handwritten notes.

First, we showed participants the printed transcripts, and demonstrated using the trigger button (see Figure 4) to scan a bar-code, invoking the corresponding video. We then gave the users a few minutes to practice using the system.

In their oral histories, Professors Patterson and Séquin both talk about graduate school; this was 15 pages (42 minutes of video) for Patterson's, and 10 pages (28 minutes of video) for Séquin's. For our main task, we asked users to spend about 30 to 45 minutes reading these sections and then write a short summary of what they read. To watch all of the video would take 70 minutes; we designed the task to have more video than could be watched in the allotted time.

After completing the main task, we asked the professional oral historians to complete two short editing tasks (these were not relevant for the book club members). Oral history transcripts are nearly always edited for clarity (e.g., removing "um," "like," and false starts), and sometimes edited for flow. Hypothesizing that strongly correlating transcript and recording benefits editing tasks, we asked participants to edit one page of the paper transcript and compare an already edited page of the paper transcript with the recorded video. At the end of the study, we asked the participants to fill out a 35 question questionnaire, addressing their background, current practices, and opinions on our system. 20 questions were multiple choice; the remaining 15 were free response.

Results

The participants in our study (see Figure 5) took between 26 and 58 minutes to complete the main task (mean 46.9), successfully accessing between 2 and 21 media clips (mean 10.9, median 10). Eight users experienced a total of 31 failed scans (mean 2.8). One user accounted for 12 of the failed scans; she also had 11 successful scans. We believe six of the failed scans were because users did not hold the device fully orthogonal to the bar-code, and 25 were because of hardware errors (e.g., the bar-code expansion pack was not properly seated) or software errors (e.g., we printed a faulty bar-code).

Benefits of Paper for Fast, Direct Video Access

Our study shows that users frequently and fluidly access recorded interviews when paper books are the interface. When asked "Do you feel that Books with Voices would change your usage practices?" participants reported responses such as, "Would be easier to select and listen to portions of an interview - easy to find excerpts - this is what is way too hard and time-consuming with analog tapes" (#10); "I think the video/audio would frequently be useful for confusing and interesting passages" (#3); and "Yes - accuracy, deeper meaning" (#5).

Participants had several motivations for accessing recordings: to get a sense of the personality of the interviewee, to hear the tone of a particularly compelling passage, and to verify the accuracy of the transcript. #9 watched the video extensively at the beginning of the main task. At the end of the study, we asked her about this. She responded that watching the video helped her understand Prof. Séquin's character.

All 13 users indicated interest in using Books with Voices for historical research. On a five-point Likert scale, half the users reported they would be "Very Likely" to use the system and half reported they would be "Somewhat Likely." None responded "Neither Likely nor Unlikely," "Somewhat Unlikely," or "Very Unlikely." When asked, "What aspects of the system do you particularly like?", several participants complimented the system for its ease of use and for, "Being able to sense tone of voice, context." More than half specifically appreciated the direct access, e.g., "I like how easy it is to access specific points in the audio" (#2). Aspects of the system they particularly disliked included a lack of fast forward and rewind, the audio was "hard to hear" (#4), the "image hard to see" (#6), or the "scanner is bulky" (#12).

Nine of the 13 users reported that, "the number of bar-codes" was about right. Three reported, "I'd like bar-codes a bit more often," and one reported, "I'd like bar-code a bit less often." This indicates to us that bar-codes offered enough value that their visual addition did not bother the participants.

Books with Voices proved effective for the editing task. With analog tapes, the editing process is often compromised by the difficulty of accessing the appropriate media segment. While transcribers sometimes verify transcripts against the tapes, it is rare that the interviewer or the interviewee has time. Users fluidly integrated video watching into their editing process. The participants responded enthusiastically to "What role, if any, do you think Books with Voices could play in editing transcripts?" writing "Very helpful to determine context, expression, figuring out certain words" (#1), "I think it could make the process more efficient - in time for searching for tape sections" (#2), and "Fantastic, accuracy, nuanced" (#5).

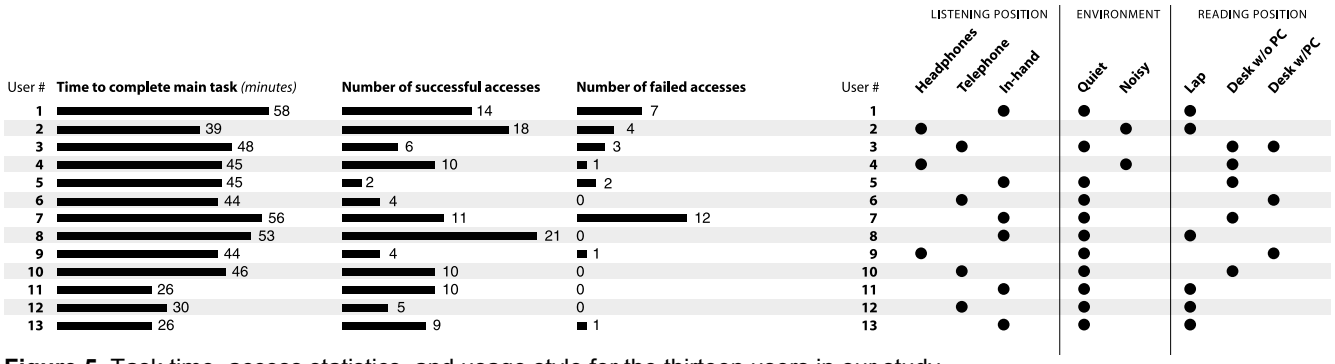

Figure 5. Task time, access statistics, and usage style for the thirteen users in our study.

Our transcripts contained the word "*inaudible" when the transcriber could not clearly hear what was said, and he prefaced transcribed words with a "*" when he was not confident the transcript was correct. Most of the participants referred to the video to clarify these inaudibles, and/or a sentence that appeared incorrectly transcribed to the participant.

Richer Practice, Minimal Overhead

We often look to new technologies, like Books with Voices, to save labor. Our participants accomplished with just a few button presses what would have otherwise consumed hours. More importantly, our system makes a *richer* practice *tractable*. Books with Voices augments reading with an audiovisual experience previously unavailable. However, this experience takes slightly more time than just reading. In the user study, some participants (e.g., #4) stopped using the software when they felt rushed late in the task.

Listening and Watching

We found that people held the device in especially different ways, depending on their age, body size, personal style, and working environment (see Figure 6).

Usability of audio

Participants were free to listen using the iPAQ's built-in speaker or to use headphones. Hearing the audio clearly was of great importance to the users. Three used headphones; #2 and #4 worked in noisy offices; #9 worked in an open, but quiet, office cubicle space. Of the ten speaker users, four held the device to their ear like a telephone, and six held it in their hand. Five of the six subjects under 30 in a quiet work environment held the device in their hand. Only one of the six participants over 30 held the device in their hand; two used headphones and three held it like a telephone.

The speaker volume on the iPAQ (like most PDAs) is soft. In a quiet environment, it may be acceptable, but generally speaking, we believe that once accustomed to the device, users will use headphones or plug the PDA into external speakers. In the questionnaire, one user lamented that because of the quiet audio, "I had to hold it up to my ear and not see the video" (#4).

Usability of video

When listening in a telephone style, as four users did, it is not possible to watch the video. The nine non-telephone users sometimes watched the video, but often did not. Several participants explicitly complained that the video was either too small or not bright enough. Increasing the resolution and brightening the videos (either with software or a physically brighter device) would help. Our prototype's 2GB hard drive holds 20 hours of video in its current form. If file size becomes an issue, it is worth lowering the frame rate or using higher compression in exchange for increasing the resolution. Although sound is much more important than image for this domain, we believe a well-presented video is still valuable.

Reading styles

All four book club readers and two oral historians read with the transcript on their lap. Four historians read on a desk without a computer, and three read on a desk with a computer. Our transcripts were bound in three-ring binders. We never would have expected binder style to become a usability issue. The physical world brings both physical benefits and physical problems. Séquin's binder was like a hardcover book; it had a cardboard spine and covers, and just plastic in between, allowing it to remain closed or open. Patterson's binder was made from one piece of shaped plastic; it would stay closed, but not open. Participant #4 promptly got rid of the binder, and several other users responded by weighting the binder to keep in open. After six users, we fixed the problem, replacing the "paperback" binder with a "hardcover" binder.

Multitasking: reading while listening

Because people read three times faster than they listen, it can be difficult to do both simultaneously. However, we found that several users comfortably listened to one section while reading another. A few users (e.g., #5) also listened while writing their summary. Generally, it seemed that younger users were more comfortable with this multi-sensory approach.

The Bar-code Scanning Process

Perhaps the largest usability issue is that the user has to rotate the PDA somewhat awkwardly to scan bar-codes. Currently, the device must be held *perpendicular* to the paper and *oriented* so the scanner scans vertically. We noticed that most participants had difficulty with this; three asked for horizontally oriented bar-codes in the free-response section of the questionnaire (#5, #7, and #9), and a few more asked about this during the study.

The *perpendicular* requirement is an intrinsic property of infrared bar-code scanners. The vertical *orientation* is a

Figure 6. Video stills from our evaluation: participants watching and listening to oral histories on the Books with Voices PDA.

design flaw of our system which could be remedied either by horizontal bar-codes or by presenting video horizontally on the PDA. Moving from an infrared bar-code scanner to a PDA with a rotatable integrated camera, such as some Sony CLIÉs would remedy both rotation problems. This way, the device could be oriented for viewing, and the sensor could still be oriented for capture.

Making and recovering from errors

When there were scan errors, the device notified users through a sonic chirp and a visual error dialog. However, about half of the participants scanned with their hand covering the device (because of the required orientation and the device's bulk). Therefore, they did not see the error dialog, and were unaware an error occurred. Switching to an orthogonal sensing technology would likely remedy this; speech output of errors might provide additional clarity.

We found the buttons need to be easier to access. The hardware trigger button should be larger. A few users had to be told where the GUI stop playback button was, and many had difficulty pressing it. It would be much better if stop was mapped to one of the four physical buttons on the lower section of the iPAQ.

Slow start latency

It takes about 3 seconds for users to orient the device and execute a scan, and 5 seconds to wait for the video to begin playing. Several users effectively adapted to this latency by pre-scanning the media, and putting the device aside while it loaded. One unexpected problem with the start latency is that some users did not realize that it had successfully scanned, and would re-scan the same bar-code. This can be fixed in three ways: by better visual or auditory feedback, by ignoring re-scans, and by lowering the latency.

Visual Design

The user study encouraged us to improve the visual design of the system in two ways. First, page numbers should be on the outside margins as opposed to the inside. Second, we should produce an index (as is currently done) and augment it with bar-codes. Participants were very excited about a bar-code augmented index. Three asked for it in the questionnaire's free-response section and a few more mentioned it during the study.

Requested Features

Nearly everyone asked for fast forward, rewind, and/or backspace. Backspace is a technique available on transcription systems, letting transcribers auto-rewind a few seconds. A physical jog dial would be a good way to offer fast forward, rewind, and backspace.

There is great potential for Books with Voices as a CSCW technology. In our questionnaire, we asked "What are your thoughts on using Books with Voices to help you keep track of important parts of an oral history?" and would this "functionality be useful for sharing? (e.g., email.) How might it work?" Participants responded enthusiastically, saying, "It could be pretty useful if you could keep a list of scanned sections and then choose to keep or delete chosen sections from the list" (#11). With regard to email, #1 wrote, "It would be helpful to email portions to others. Esp. for editing, since many people work on any single transcript," and #8 wrote, "There would have to be some way to add your annotations, but [email] seems useful."

General Remarks

Participants used the device differently depending on their job. One transcriber mostly used the device for editing and fact-checking, and could not help but make edit marks on the transcript as she read. A senior history professor accessed the video the least (twice). She took care to finish the study in a timely manner as she had another meeting immediately afterward.

While the current system makes text a fluid interface from text to video, there is no facility to go from video to text. Occasionally, participants got lost. Subtitles on the video indicating page and paragraph number might remedy this.

One success of the system in our study was that, while there were usability problems, none of the users found the system conceptually difficult. The concept of using paper transcripts as an interface to original recordings seemed perfectly "natural."

CONCLUSIONS AND FUTURE WORK

Reading is a highly evolved practice. Our evaluation showed that Books with Voices effectively enables active reading by scaffolding new technologies on paper, which is

highly familiar, cheap, and usable. Our technology support for oral histories differs from much current research in that it employs paper as an *archival,* rather than *ephemeral,* artifact. After seeing paper prototypes and participating in our study, oral historians asked us how they could switch to digital transcription tools. They saw many benefits to this, including the ability to use Books with Voices. As we move toward long term deployment of this software at ROHO, we are eager to see how use patterns evolve.

ACKNOWLEDGEMENTS

We thank Jennie Park, Lin Hsu, and Alan Liu for their help in preparing and analyzing the user studies; Jonathan Hull and Ame Elliott for their stimulating conversations; and Marti Hearst and our other GUIR colleagues for insightful feedback on drafts of this paper.

REFERENCES

1 The Bancroft Library Regional Oral History Office, UC Berkeley, 2002. http://bancroft.berkeley.edu/ROHO/

2 DataGlyphs: Embedding Digital Data. Palo Alto Research Corporation, 2002. http://www.parc.xerox.com/solutions/dataglyphs/

3 Disability Rights and Independent Living Movement. Regional Oral History Office, 2002. http://bancroft.berkeley.edu/collections/drilm.html

4 H-Oralhist. http://www2.h-net.msu.edu/~oralhist/

5 Back, M., J. Cohen, R. Gold, S. Harrison, and S. Minneman, Listen Reader: an electronically augmented paper-based book. CHI Letters, 2001. 3(1): p. 23-9.

6 Brewster, K. Internet Access to Oral Recordings: Finding the Issues, October 25, 2000. http://www.uaf.edu/library/libweb/collections/apr/internet.oralhist/

7 Churchill, E.F. and L. Nelson. Tangibly Simple, Architecturally Complex: Evaluating a Tangible Presentation Aid. In Proceedings of CHI, Extended Abstracts: ACM Press. pp. 750-751, 2002.

8 Dunaway, D.K. and W.K. Baum, Oral History: An Interdisciplinary Anthology. 2nd ed. ed. Altamira Press. pp. 440, 1996.

9 Ellis, J.B. and A.S. Bruckman, Designing Palaver Tree Online: supporting social roles in a community of oral history. CHI Letters, 2001. 3(1): p. 474-81.

10 Gibbs, G.R., S. Friese, and W.C. Mangabeira. Using Technology in the Qualitative Research Process. Forum: Qualitative Social Research, May, 2002. http://www.qualitative-research.net/fqs/fqs-eng.htm

11 Graham, J. and J.J. Hull. Video Paper: A Paper-Based Interface for Skimming and Watching Video. In Proceedings of International Conference on Consumer Electronics. Los Angeles, CA, June 16-18, 2002.

12 Gustman, S., D. Soergel, D. Oard, W. Byrne, M. Picheny, B. Ramabhadran, and D. Greenberg. Building and using cultural digital libraries: Supporting access to large digital oral history archives. In Proceedings of ACM Joint Conference on Digital Libraries. Houston, TX: ACM Press. pp. 18 - 27, 2002.

13 Johnson, W., H. Jellinek, L. Klotz, R. Rao, and S. Card. Bridging the Paper and Electronic Worlds: The Paper User Interface. In Proceedings of CHI: Conference on Human Factors in Computing Systems. pp. 507–512, 1993.

14 Kindberg, T. Implementing physical hyperlinks using ubiquitous identifier resolution. In Proceedings of 11th International World Wide Web Conference. Honolulu, 2002.

15 Li, F.C., A. Gupta, E. Sanocki, H. Li-Wei, and R. Yong, Browsing digital video. CHI Letters, 2000. 2(1): p. 169-76.

16 Ljungstrand, P., J. Redström, and L.E. Holmquist. WebStickers: using physical tokens to access, manage and share bookmarks to the Web. In Proceedings of DARE: Designing Augmented Reality Environments. Elsinore, Denmark: ACM Press. pp. 23 - 31, 2000.

17 Mackay, W.E. and A.L. Fayard. Designing interactive paper: lessons from three augmented reality projects. In Proceedings of IWAR. pp. 81-90, 1999.

18 Murphy, K., Bigger Bar Code Inches Up on Retailers, The New York Times, 2002. http://www.nytimes.com/2002/08/12/technology/12CODE.html

19 Nelson, L., S. Ichimura, E.R. Pedersen, and L. Adams. Palette: a paper interface for giving presentations. In Proceedings of CHI: Conference on Human Factors and Computing Systems. Pittsburgh, PA: ACM Press. pp. 354-361, May 15-20, 1999.

20 Rekimoto, J. CyberCode: Designing Augmented Reality Environments. In Proceedings of DARE: Designing Augmented Reality Environments: ACM Press. pp. 1-10, 2000.

21 Riley, J. human voice. H-OralHist Mailing List, 2002. http://h-net.msu.edu/cgi-bin/logbrowse.pl?trx=vx&list=h-oralhist&month=0209&week=c&msg=DKYtxLCtdA8Uwy3IP52fLA&user=&pw=

22 Schilit, B.N., G. Golovchinsky, and M.N. Price. Beyond paper: supporting active reading with free form digital ink annotations. In Proceedings. pp. 249-56, 1998.

23 Schilit, B.N., M.N. Price, G. Golovchinsky, K. Tanaka, and C.C. Marshall, The reading appliance revolution. Computer (USA), Computer, 1999. 32(1): p. 65-73.

24 Spielberg, S. Survivors of the Shoah Visual History Foundation. http://www.vhf.org/

25 Starr, L., Oral History, in Encyclopedia of Library and Information Sciences, P.R. Penland, Editor. Marcel Dekker: New York. p. 440-63, 1977.

26 Stifelman, L., B. Arons, and C. Schmandt, The audio notebook: paper and pen interaction with structured speech. CHI Letters, 2001. 3(1): p. 182-189.

27 Suchman, L.A., Conversation as "ensemble" work, in Plans and Situated Actions. Cambridge University Press: Cambridge. p. 72, 1987.

28 Ullmer, B., H. Ishii, and D. Glas. mediaBlocks: physical containers, transports, and controls for online media. In Proceedings. pp. 379-86, 1998.

29 Wellner, P., Interacting with Paper on the DigitalDesk, Communications of the ACM, vol. 36(7): pp. 87–96, 1993.

30 Wilcox, L.D., B.N. Schilit, and N. Sawhney. Dynomite: a dynamically organized ink and audio notebook. In Proceedings of Conference on Human Factors and Computing Systems. Atlanta, GA. pp. 186–193, March 22–27, 1997.