Sonic Gems: Exploring the Potential of Audio Recording as a Form of Sentimental Memory Capture

Gerard Oleksik
Instrata Ltd
12 Warkworth St
Cambridge CB1 1EG
+44 1223 479 700

goleksik@instrata.co.uk

Lorna M. Brown
Microsoft Research
7 JJ Thomson Ave
Cambridge CB3 0FB
+44 1223 479 700

lornab@microsoft.com

ABSTRACT

This paper presents an exploratory study exploring the potential of audio recording as a form of sentimental memory capture. Drawing on data from four family households, we spotlight participants' attitudes towards sounds and audio recording, their existing recording practices and the types of sounds that they record when given a digital sound recorder. The findings indicate that a variety of different sounds were important, for a diversity of reasons. The paper considers participants' experiences of listening back to sounds, illuminating some of the unique affordances of audio recordings as a form of sentimental memory capture. The paper finishes by exploring the design challenges which the study raises for the capture, playback and archiving of sentimental audio.

Categories and Subject Descriptors

H.5.5 [Information Interfaces and Presentation]: Sound and Music Computing.

General Terms

Design, Human Factors.

Keywords

Audio capture, sound, audio, memories.

1. INTRODUCTION

In the 1980's the audio cassette was a popular musical format. Audio tape differed from the musical formats which had preceded it and, indeed, many which came after, in that it allowed the user to easily make and save their own sound recordings and play them back. However, as the audio cassette declined in popularity, to be replaced with the compact disc (which offered no scope for self recording), so too did the potential to easily capture audio. Although audio recorders are still available in the form of digital sound recorders, such devices lack the popularity and appeal of cassette recorders, which could be equally used for playback and recording. The option to record audio now also exists on many mobile phones, but the quality is generally rather poor. And, thus, the practice of audio recording appears largely to have been lost. One might argue that audio recording has been made redundant by the

ability to capture video; after all why only record sound, when you could capture sound and vision? However, it is possible that audio capture alone may have other unique affordances, and this opens up an interesting area for exploration.

Indeed, findings from our previous work in the Sonic Interventions [10] study suggest that this may be the case. Sonic Interventions explored the domestic soundscape through interviews and self recording exercises. Although the recording exercises were primarily geared towards the documentation and understanding of everyday sounds, they nevertheless uncovered certain sounds which had huge emotional resonance to the listener. These sounds were referred to as "sonic gems" – short, precious sounds which people wanted to keep.

The present study explores the concept of "sonic gems" in more depth. This study aims to understand how people feel about sounds, and the idea of capturing and storing special sounds ("sonic gems") in the same way that people currently capture and store other media, such as photos and video. Several studies have been published exploring how people capture and store photos and video [8, 9], and even audiophotos [3]. However, there has not, to date, been any study of the use of audio recording alone for sentimental memory capture. Indeed, the majority of research on sound recording has focused almost exclusively on the functional uses of sound [5, 6, 12], while research on the value of sentimental audio recordings is scant [4, 11]. This study aims to help redress this balance.

Our study was designed to explore the potential of sound as a form of sentimental memory capture. In doing so, we provide a fine grained account of users' experiences of capturing and replaying audio, examining the affordances of audio recording and playback and how this compares to other more established forms, e.g. photography or video. Here, we spotlight what sounds people most remember and consider to be special, what audio recording and archiving practices people currently engage in and the kinds of sounds that people will capture when again given the opportunity to do so. By gaining an understanding of these issues, we can uncover a number of challenges and opportunities in designing for audio capture, storage and playback, which open up new areas to be explored by the HCI community.

2. METHOD

Four families with young children (under the age of 6) were recruited to participate in this study (See Table 1), through personal contacts and via an online community forum. We chose to focus on families with young children as our previous work on the Sonic Interventions study [10] had shown that most of the treasured sounds captured in the study featured the sounds of young children, and this, therefore, represented a fruitful starting point from which to explore this area further.

While the sample used in this study is small and homogenous, which has implications for the generalisability of our findings, we would like to emphasise that this is a preliminary study to explore this new area. Future work could follow up this study with a larger sample and wider range of households, to understand the different types of audio that might be important to different demographics.

Table 1: The four households who participated in the study

Household	Description
H1	Couple with young granddaughter who visits daily
H2	Couple with 2 children under the age of 6
H3	Couple with 2 children under the age of 6 (only the wife, Sarah, took part in the study)
H4	Couple with baby aged 5 months.

The study comprised two phases. Phase 1 took the form of a two hour home visit investigating the kinds of sounds that people remember and consider to be special, and any existing practices of recording, archiving and listening back to sounds. We opened this area by asking participants to consider ‘all the sounds they had heard over their life’, and to tell us about any sounds which they particularly liked or remembered warmly, for whatever reason. This overarching theme was supplemented by sub-questions, regarding special everyday sounds and sounds from special occasions, with a further enquiry investigating whether participants wished they had recordings of any sounds from the past. The second part of the visit was devoted to exploring the range of existing audio recording practices, and we asked participants whether they had any recordings of sound or devices on which to record sound. Subsidiary questions asked where any such recordings were stored, why they had been made and if and when they were listened to.

For Phase 2, we gave each family a handheld digital voice recorder (Sony ICD-SX67) and asked them to record any sounds they felt they would want to keep, over the following two weeks. After two weeks, we collected the voice recorders and listening to and categorised the participants’ self recordings. We returned to the homes the following week to discuss the recordings. Household specific interview schedules were prepared, based around the recordings. On arriving for the home visit, we would set up speakers and play the recordings back to participants, pausing after each to ask questions.

For each recording we asked the following questions, along with more specific questioned tailored to the particular sounds recorded: Why did you record this sound? What does this sound evoke? how does it feel listening back to this sound? How do you think the experience might differ if this was a video, or was accompanied by a photo? After listening back to all of the recorded sounds, we asked a number of general questions regarding what they might want to do with these sounds, which they would want to keep, which they would want to share, and their experience of performing the audio recording task.

To explore how people might interact with recorded sounds in new ways, we also edited some of the sounds, and included the sound fragments in a prototype design concept to show to the families. The Sonic Gems Bowl was proposed in [10], as a display concept for Sonic Gems. We implemented this concept using a large bowl with an RFID reader attached to the side (Figure 1). Objects (in this case, plastic eggs) with RFID tags attached to them are placed in the bowl, and each tag is associated with a short sound. As an object is lifted out of the

bowl, past the tag reader, the sound is triggered. We extracted a range of short clips, which we felt might be “sonic gems” from the participants’ recorded sounds and put these into the bowl. At the end of the home visit, participants spent time interacting with the bowl and we gathered their reactions to it.

Figure 1: The Sonic Gem Bowl. The objects in the bowl contain RFID tags (top left) and the bowl has an RFID reader attached to the side

3. FINDINGS

The following sections present the findings from both phases of the study. Sections 3.1 and 3.2 report the findings from Phase 1. 3.1 discusses the sounds that people described as special or which they remembered warmly for some reason, while 3.2 presents the findings regarding participants’ current and past audio recording practices. The findings from Phase 2 are presented in Sections 3.3-3.7, which discuss the types of sounds recorded, the experience of listening back to them, the experience of capture, and participants’ views on playing back, archiving and sharing sounds. Finally Section 3.8 presents an overall discussion of the findings from the study.

3.1 Special/Remembered Sounds

Interview data from the first home visit revealed a wide range of sounds which participants described as being their favourite, or which they particularly liked or remembered warmly. These sounds fell into the following broad categories: natural/ambient sounds, voices/phrases/spoken word, special music, pets and everyday domestic sounds. Below, we briefly outline the main findings under each of these categories, whilst also considering past sounds which participants wished they had a recording of. We then begin to explore *why* these sounds are special to participants, unpicking the many complex meanings and associations bound up with them.

Natural/Ambient sounds: Natural and ambient sounds were amongst the most commonly mentioned as favourite or special. Participants referenced a range of sounds, including the trams in Brussels, traffic noise in London, crickets, the sea, thunder, wind blowing through eucalyptus leaves and rain on the roof.

Voices/phrases/spoken word: Several participants mentioned the sound of peoples’ voices as being amongst their favourites. In Household 1 in particular, participants talked about the importance of voices, accents and catchphrases of friends and family members. During the interview, they often mimicked such voices and phrases, and talked about why these were important to them. In the following, Sue from household 1 provides one such example:

“It’s quite nice watching The Last of the Mohicans, the way that the father, the Scottish father says to his girls

'hello girl'. It sounds, although he's Scottish and my Dad was Irish, it sounds exactly like him. There are certain words that come out the same. And it always brings a tear to my eye, it's nice to hear it all the same.'

The sound of children was commonly mentioned as a special sound, with Sarah (Household 3) mentioning the sounds of her children laughing and Amy and Pete (Household 4) mentioning the sounds of their new born baby *"babbling"*.

Special Music: When talking about favourite sounds, participants sometimes talked about particular pieces of music. Such pieces of music were often considered special because of their associations with particular times, places or events.

Pets: Another special sound was the sound of pets, particularly those who were no longer alive. Household 1 were the only family to discuss this, and gave several examples from various pets (including dogs and budgies).

Everyday Domestic Sounds: A number of everyday domestic sounds were mentioned as being particularly liked by people in the study. Pete (H4) said that he liked the sound of a hammer drill, which he described as *"the sound of a Sunday afternoon"* as it reminded him of his Father, who would use a drill during this time when he was a child. Sue (H1) mentioned that she likes hearing the key in the door of her home because she knows that Pete is home safely. This sound now has a special association for her was although in the past the same sound had a negative association when she had been in an abusive relationship. Mark and Angie (H2) mentioned the sound of the front room grandfather clock, while Sarah mentioned the sound of coffee brewing.

When we asked people about sounds from that past that they wished they had recorded, they mentioned many of the types of sounds described above. One of the most common types of sound mentioned at this point, which had been mentioned less frequently before, was sounds from special occasions, usually parties, wedding and in one case a funeral. Several households mentioned that it would be nice to have a recording of their wedding vows, or the conversations at their wedding. Sue (Household 1) mentioned a recent family gathering she wished she had recorded, and a number of participants wished to have recordings of sounds they had heard whilst travelling or being on holiday.

People, places and special events all appear to have their unique sonic signatures, which participants often valued and remembered with warmth. Such sounds range from the aesthetically pleasing sounds of nature, to a familiar voice or accent, through the sounds of traffic noise, to the sound of a hammer drill. We are now in a position to explore *why* these particular sounds were important to people. In most cases, it is not the sound itself that is important, but rather the web of associations within which the sound exists. For example, the sound of a hammer drill reminded Pete (H4) of his father, the voice in *Last of the Mohicans* reminded Sue (H1) of her father, the sound of the tent zip reminded Amy (H4) of being warm and secure and the sound of the ocean reminded Mark (H2) of being on holiday. Associations with people, places, events and eras all helped to determine the importance of a sound to an individual.

The findings regarding the importance of loved ones voices are also significant. Although the things that people say are important (their catch phrases and comments) it is often the way in which these words are said that holds value. The tones, timber, inflections and accents of the human voice are instantly recognisable and often held dear to participants.

These data have shown that there are a number of sounds which members of our sample remember fondly or consider special for some reason. However, had our participants ever tried to harness the emotional power of sound by capturing it?

3.2 Past and Existing Audio Related Practices

The interview data revealed that various practices of audio capture, sharing and storage were already established. Such audio recordings had been made on a number of different technologies for a range of reasons, and in a variety of ways. These technologies included traditional tape recorders, answer machines, Dictaphones, and mobile phones. Below we shall discuss these recordings further.

In all four households, participants reported that they had in the past made audio tape recordings. These had often been made many years earlier, when audio tape had been a popular format, on cassette recorders, or, in one case, a Dictaphone. Participants reported making recordings of a variety of people and events, including singing, children putting on a staged performance and special events, such as Christmas day. In the following example, Sue (H1) speaks very fondly of a recording her daughter made on a Dictaphone as a child:

"Oh, Lucy's little garden programmes, I wish we'd kept them. She used to walk round the garden when she was about seven reading all the plant labels and talking about them as if she was a garden presenter. Oh, it was priceless because she used to read them wrong. So the Magnolia tree became the Mangolia and all sorts of mispronunciations. It was so sweet, it was so nice to listen to it. I've lost them."

While this recording was of children, many participants remember using tape recorders *while* children themselves. Pete (H4) talked about his creative use of a reel-to-reel tape recorder as a child:

"One of my earliest memories was me and my sister recording ourselves going, Mum, I've finished, [laughing] and trying to trick my mum into coming upstairs. "

Many audio tape recordings which participants had made in the past were now lost or deleted, or their exact whereabouts unknown. Even when the location of recordings was known, participants often no longer had the means to listen back to them. In Household 3, a large collection of tapes had been disposed because the family no longer had a tape player, even though Sarah said it would be *"awesome"* to hear the recordings again. Household 1, however, had a large archive of audio recordings, which they found incredibly precious and emotive, and a cassette player on which to play them. Some additional problems participants reported with audio tapes, was that they tended to be stored rather haphazardly, and that it was difficult to locate significant moments on a given tape.

One surprising finding was that a few of the recordings participants had made, and kept, had actually been captured accidentally. This accidental capture had probably caused annoyance at the time, but now these sounds had become more precious than the sounds originally intended for recording. The two examples of this are as follows (both from Sue, H1).

"...my dad had set up a tape to record a radio programme, back in the old days when you had tapes. And it hadn't worked, so it had just recorded the kitchen instead, so you could hear my mum and me making a cake and you could hear her sort of saying, you know, 'oh, careful, hang on, oh, okay, well get another egg out; okay, fine', you know. And me sort of chatting in the background

and things. So that was an accidental half an hour, you know."

"One old cassette, a music compilation tape, and in the middle of it, suddenly you hear Fiona and Lucy when they were little children. Fiona had obviously pressed the record button and you hear this little voice going 'oh, no!'. And then Lucy, who was even littler, saying something like, you've got to press this one or something like that, and it stopped. So it's just, like, a couple of seconds but it's so precious."

Another way in which audio had been captured in the past was through answer phone messages. Although these are generally recorded for functional purposes, the resulting messages can sometimes be very precious to the recipient. Sue (H1) spoke very emotionally about an answer machine message she had wanted to save of her father shortly before he died. Unfortunately she could not afford to replace the answer machine tape, and the message was deleted, much to her distress. Although this is a particularly emotional example, it illustrates that audio messages can be precious.

In recent years, mobile phones had been used occasionally by two households (H3, H4) to capture sounds. The resulting sounds were either used as messages, or shared with others. For example, Sarah (H3) used a recording of her son saying *"you've got a message"* as her message alert tone, and had used this same recording for two years. In Household 4, both Amy and Pete had used phones to capture environmental sounds and had sent them to others as messages. In a particular case, which occurred in the week prior to the study, Amy had recorded the sound of a very distinctive hotdog seller in the town in which they live and sent it to a friend via MMS. This was a sound which is very familiar to most residents, and which Amy knew that her friend would recognise. Thus, a shared memory was embodied in the sound recording

One further example of sharing sounds arose in Household 1, although through very different means. In this case, Phil talked about an occasion when an audio tape itself had been copied and physically sent to its recipient:

"I gave those recordings of John Smith to his family, a chap that I was quite friendly with who committed suicide. I didn't actually know his family but he'd been interviewed by me about his year that he spent on a sort of aid project in Nicaragua... I had the tape so, I did a copy for the family and it was quite a shock to them really, although they appreciated it."

As well as sharing sounds, participants also on occasions listened back to their recordings. There were a number of different reasons for this. For example, Sue (H1) sometimes chose to listen back to a recording of her mother and sisters for its cathartic and restorative properties:

If I'm feeling really bad, I'll put it on. It makes me cry too because my mother's gone now, but it's actually a wonderful, wonderful thing."

For some participants tapes were not listened to on a regular basis, but rather when they were found. Angie (H2) had a number of tapes containing recordings from her childhood, and said that she listened to the recordings a few years ago when she had come across them, but had not listened to them since.

The findings considered above reveal the existence of a number of audio recording and sharing practices which participants had in the past, or still did, engage in. Such recordings had been made for a variety of reasons, be it for playfulness (e.g. children's use of recorders), to share with others (e.g. hot dog

seller), to personalise an object (e.g. mobile phone message alerts), for functional reasons (e.g. answer machine messages), to playback or just keep (e.g. sounds of family members having fun), or simply by accident (e.g. the recording of a room instead of a radio programme). Many of these recordings had been kept, but many had been lost or deleted. Regardless of how or why they were captured, participants talked about them with emotion and seemed to remember recorded sounds fondly even when the actual recordings were long since lost.

The findings also show that, for these families, recording was more common in the past when cassette recorders were commonplace, and that audio recording is rarer now, occurring only occasionally via mobile phone. The importance of existing audio recordings to participants suggests that there is potential scope for sentimental audio capture to be revisited. The rest of the paper explores what people will record when given the opportunity to capture sounds again with the ease previously offered by cassette recorders, and how they react when listening back to these sounds. Reflecting on these findings, we shall go on to consider the ways in which we might design for sentimental audio capture and display.

3.3 Responses to the recording task

Participants responded to the recording task in a wide variety of creative ways. Each family recorded a different number of sounds: Household 1 recorded 33 (although several of these were old recordings from their audio archive which they re-recorded for us), Household 2 recorded 9, Household 3 recorded 17, and Household 4 recorded 13. The main types of audio captured were as follows¹.

Recordings of children / babies: In Households 2 and 3, the majority of the recordings made were of the children from those households (6 of the 9 recordings from Household 2 and 9 of the 17 recordings from Household 3). These two households had children of similar ages, and the recordings featured the children singing, talking to the recorder and in conversation. In Household 4, five of the 13 recordings made were of sounds associated with their baby son, including the sounds of him burping and squeaking, and the sound of the mobile that hangs in his room. Household 1 only made one recording of their granddaughter during the study, in conversation with the grandmother. However, they also re-recorded some sounds from their audio archive which included three recordings of their children when they were young.

Singing: Recordings of family members singing were also very common. It was mentioned above that some of the recordings of children featured them singing. For Household 2 this was the only occurrence of music making that was captured, but for Households 1 and 3, there were also recordings of adults singing. In Household 1 the participants were very musical, with Sue having worked as a singing teacher, and four of the self recordings were of the participants singing in staged performances for the recording. In addition they included five recordings from their audio archive of them or members of their families singing. In Household 2 the recordings of adults singing were of the adults and their friends playing the "Singstar" game on Playstation 2.

Conversations: All of the households made recordings of conversations. This was most common in Household 1, who made 7 recordings featuring conversations. These ranged from ambient recordings of conversations about practical things like diets and food, to chats about the recording task itself, to a discussion with a market seller, to an in depth debate about a

¹ It should be noted that some of the recordings fall into more than one of these categories.

book. These conversations took place with both family members and friends. The other families recorded fewer conversations (1-2 per household). These were generally ambient recordings which also happened to capture conversation.

‘Ambient’ recordings: Another common recording was that of ambient sounds, which were recorded by all Households (with 2 sounds being the minimum and 6 the maximum). These included ambient recordings of everyday life in the home (such as watching TV together), the sound outdoors at 5am, the sounds of a trip to the zoo, the sound of driving and the sound of walking down a particular street. Household 1 also provided some recordings of canned ambient sounds, downloaded from the internet.

Everyday domestic noises: Household noises were actively recorded by H3 and H4. The majority of these were sounds of food preparation (pot boiling, pan sizzling). In addition H3 recorded the sound of a bath running and H4 recorded the sound of a curtain being opened.

Recorded music: Household 4 provided recordings of 13 of their favourite pieces of music. No other household specifically recorded any music, except in the cases where the music was put on for a child to sing along to.

‘Special Occasion’ recordings: As the recording phase only took place over only 1-2 weeks, most families did not attend any special events during that time. Therefore, there were fewer recordings of special events than one might expect. Nonetheless, H3 made several recordings of a holiday, and H4 made one recording of a surprise birthday party.

3.4 Liked/disliked sounds

Participants placed varying degrees of value upon the sounds which they had recorded. While there were certain sounds, or certain aspects of recordings, which participants tended particularly to like and value, there were other aspects which were less valued and sometimes actively disliked.

In terms of the content of recordings, sounds of children / babies were almost universally liked and prized the most highly of all recordings. These recordings were often liked in themselves, but were also valued because they captured a certain point in the development of the child, capturing words or phrases typical of that time. Sounds of family members being together (for example, in conversation or watching TV together) were also valued because they provided an authentic ‘slice of life’ (Angie, H2). Sounds of singing made specifically for the recording were often considered to be fun and enjoyable to listen to. ‘Canned’ ambient sounds were liked for their aesthetic properties and also for their associations. Household sounds, while liked primarily for their associations, were not particularly valued or precious to participants.

While recordings of children were the most prized by participants, they could also be amongst the most disliked. In such instances, children were invariably being badly behaved. For example, one of Household 2’s recordings contained the sound of their children ‘whining’. On rehearsing this recording, Mark stated that it is a sound which made him ‘instantly tense’ and which he would never like to listen to again. However, Angie thought this sound might gain value over time. On a number of occasions, participants stated that they did not like the sound of their own voice, and found listening back to it to be a slightly uncomfortable experience. Also, recordings which overstepped notions of privacy were also disliked, for example, when one half of a telephone call was recorded by accident.

Also, the length of the recording, at times, had a bearing upon the degree to which recordings were liked and valued.

Sometimes longer recordings were liked, in that they provided an ongoing temporal capture of an event in which participants could immerse themselves in. However, longer recordings were more often a source of frustration to participants. Most often this stemmed from the fact that participants were aware of interesting sections in the recording, but had no way of easily or accurately ‘skipping’ to those moments. Mark (H2) voiced his frustrations and aspirations in the following:

“There’s an element of boredom. It starts creeping in when you’ve got a long recording and you’ve got moments here which are good and moment here which are bad. [...] you need to explore this, how to make it simple, to get bits out of your sound”.

At certain stages, all participants expressed similar feelings, stating that they would like to have short but significant ‘snippets’, ‘snapshots’ or ‘chunks’ of sound rather than longer recordings. Such a finding was further born out by participant responses to the sound clips presented in the sonic gem bowl. In advance of the second home visit, we prepared and edited a series of short recordings and associated them with objects in the bowl. Participants were often delighted with these short recordings. For example, we had strategically taken short fragment of Household 2’s daughter speaking about characters from Star Wars from a recording which Mark had otherwise described as ‘just a very matter of fact conversation’. On listening to this Mark was clearly moved, stating:

“There was just a level of excitement and enjoyment in her voice and, you know, that whole innocent childhood thing was going on and everything was in that, a lot’s encapsulated in those two words.”

This is typical of the short recordings which participants most valued, i.e. fragments of sound which encapsulated something important to the listener.

3.5 The experience of listening back

After each recording, participants were asked to describe their experience of listening back. A very wide range of reactions to the sounds were registered, ranging from boredom, frustration, to delight and, on occasion, tearful recollection. In the following we unpick these experiences. Through this we highlight the unique affordances of audio as a form of sentimental memory capture, whilst also providing further insight into the sounds which participants valued and why they did so.

3.5.1 Sound and Memory

One of the strongest themes to emerge from the data was that participants’ recordings often provided them with very strong memory cues. On listening back, individuals were often transported ‘back into the moment’ from which the recordings had originally been taken. A large element of this appeared to rest upon the fact that listening often gave rise to vivid mental images of the event in question. For example, on listening back to a recording of an everyday conversation with friends and family, Angie (H2) stated that *“I’ve got the whole [...] visual [...] impact of where we were and what we were doing as well”*. Such mental images often unfolded at the same pace as the recording, enabling the individual to mentally recreate the original event in real time. This aspect of listening was mentioned at least once by each participant, and often on a number of occasions.

While recordings were often found to inspire images of their original contexts, strong visual impression of an event could still be derived by people who were *not* present at the time of recording. Mark (H2), describing his experience of listening

back to his children at play, provides the only such example of this from the study:

"I could just see Emma, I could just see her doing stuff and I could see her expression while she's doing it and stuff... I can see Chris running up and down as well... You do get more visually out of it than what you would expect to because most of your memories tend to be more visual things and kind of jogs that visual memory... I know what Chris does and I know how he is and I can hear that and so I can say that's what it would have looked like."

The nature of the visual images which this recording inspired differed significantly from those described in the previous example. Whereas Angie's mental images of the event were recalled from memory, here, Mark is in fact *constructing* the event in his imagination, using the recording as a prompt and combining memories from other events.

The link between sound and strong visual and emotional recreation of time and space was also revealed through one household's experience of listening to 'canned', or pre-recorded sounds. Amongst the sounds recorded by H1 were sounds of the ocean and the sound of a nightingale singing, downloaded from a free sound effect website. Their experiences of listening back to these recordings revealed another order of association between sound and memory. When listening back to the ocean, Phil was taken back to bittersweet memories of peaceful yet lonely times of his life in which the sound of the ocean had figured. Similarly, on listening back to the sound of a pre-recorded nightingale, Sue (H1) was transported back to a very specific moment in her history:

"Well, it's, it's emotional for me 'cause I heard it once with a dear friend who I don't see anymore. We both stood there going, wouldn't it be nice if we was in love. It was one of those good drinking pals that you don't love one and other but you spend a lot of time together. It's a nice little memory".

These examples are notable because they show that generic and 'canned' sounds, such as birdsong and the ocean, can in fact trigger the recollection of utterly unique events, moods and feelings from the past. It may not matter that these sounds were not taken from the specific event or time.

These examples reveal the relationship between sound and memory operating in three separate ways. In the first, sound prompted the (often visual) recollection of events; the second, sound prompted the visual recreation of events through the imagination; and, in the third, sound triggered associated feelings and memories. In each of these, sound is operating as a scaffolding mechanism for memory and imagination. The ability of sound to stimulate such often unbidden images and memories seemed to lie, in some key aspects at least, in its lack of visuality, as we shall now explore.

3.5.2 Sound and Vision

Most established forms of memory capture (video, photograph) have significant visual elements, and so we were particularly interested to understand how participants responded to unaccompanied sounds. A range of opinions were recorded regarding this issue, and while participants sometimes felt that sounds would be improved with visuals, more often participants felt that the unaccompanied sounds had a particular and unique value, and were sufficient in themselves.

One of the key factors determining whether visuals were necessary or not was the issue of context. This is well illustrated by Mark's and Angie's (H2) differing reactions to a recording of Angie as a child. While this recording had not been made for

the study, both participants had listened to it in the past, and their reactions are noteworthy. For Angie, this recording recreated nostalgic feelings of childhood, and while she felt visuals may have provided additional details, they were not necessary for her to create the feeling, and neither did she feel that they would in anyway strengthen this feeling. Mark's reactions were in contrast. While he felt that the recording had value, and give him an insight into his wife's history, he felt he needed visuals in order to fully understand the clip:

"Because I wasn't there or anything, it didn't give me a sense of place or anything, so where Angie was saying from the perspective that she was there, so the video wouldn't add anything, but for me not being there, the video would have added a lot, to have seen other things going on around probably.."

In the previous section, we saw how Mark was able to construct vividly an event from a recording, even though he had not been present at the original occurrence. He was able to do this because he possessed a good deal of knowledge about the participants and their context. In the example above, Mark had no such knowledge – he did not have a direct memory of his wife as a child, nor of the house where she had lived – and consequently had little frame of reference in which to place the recording. Therefore, the recording did not provide enough information of itself to be fully satisfying.

While participants sometimes felt unaccompanied audio to be problematic – either for themselves as listeners or for others – in the majority of cases participants found audio alone to be a desirable and evocative form of capture. Often, this seemed to spring precisely from its lack of visual input, a circumstance which opened the space for a particular form of engagement with memory which was not possible with more visual forms of capture. The following three quotations, each referring to the experience of listening back, begin to illustrate this:

Sarah (H3): *"You've got your own imagination as opposed to what's there in front of you."*

Phil (H1): *"You concentrate on the sound much more I suppose."*

Sue (H1): *"I think sometimes just sound alone makes you go into your, your own memory more."*

These quotations further illuminate the experience of listening back to audio only and, in doing so, draw attention to the particular relationship between sound, vision and memory that this form of listening draws out. Because there are no visual elements, participants often found themselves having to concentrate in a focused way on the sounds in order to understand them. As illustrated in the previous section, missing visual information was often supplied by the listener, in the form of vivid mental images of the event in question. The dual act of focused listening and imaginative visualisation appeared to take listeners further back into the memory. This may be because, rather than being presented with a definitive image, visual memory is being stimulated and the image is recreated, and in some ways, re-experienced, anew. However, because of the mental reconstruction process, unaccompanied audio relies heavily on the listener and, crucially, the listener's memory, to provide missing context and sense. We would therefore expect audio only recordings to have less value and meaning to listeners who were not at the original event and had limited background knowledge of the event and its participants. This is borne out by the examples considered above.

3.6 The experience of capture

Although recording audio has similarities with taking a photo or making a video, it nevertheless differs in some major respects.

Here, we investigate participants' experiences of sound recording, how this compared to other forms of capture and the impact this had on the nature of participants' recordings

Participants strongly felt that audio capture was more natural and less intrusive than photo or video capture. The form factor of the device, along with the one-click record, afforded immediate spontaneous capture, and meant that the device could easily be kept in a pocket or held discreetly in the hand. In addition, unlike video or photography, the recording device required very low levels of attention both before and during capture, allowing the individual to participate more fully in the event. This issue seemed to be of key importance to the families, particularly Household 2 who described how their experiences of audio recording different from video capture, which they had tried but not found to be useful for them:

Mark: *"Yeah, you just stick that in your pocket, you haven't got to focus it or anything."*

Angie: *"...and also you're not putting up a barrier between you and the world in the way that we found with the video camera, which is why we didn't use it, because we thought you weren't experiencing the event because you had to be the other side of the video. We thought we would use it all the time but we didn't actually want to sit there the other side of a video camera recording, we wanted to be there with them. Whereas that, you can just set it to record, you're still there, you're still in the moment."*

Audio recording appeared to bring down and blur the boundaries between observer and participant. Video and photography are largely unidirectional forms of capture, in that they only visually capture what is directly in front of them. Because of this, the role of capturer is essentially one of observer, standing *behind* the camera. Audio, however, is omnidirectional, meaning that the information is received from all directions, despite the orientation of the recording device (to an extent at least and depending upon the type of microphone used). As the above illustrates, this enabled the observer to become more participant in the event.

The spontaneous and discreet nature of audio capture might lead to captured sounds having a different value to that of video or photos as they can be more natural and less staged. This was noted by Sarah (H3), who also stated that children might be less likely to show off or get shy if an audio recorder was used rather than a video.

"I think it would be useful for parents for when kids do say quirky things but you haven't got time to set up a video recorder, and you just pick it up, push a button and record, um, without it being forced and, um, staged."

Nonetheless, it should be noted that many of the captured sounds in the study were staged sounds and, thus, it should be assumed that such a device may be used for both staged and un-staged recordings.

One concern with the easy and discreet nature of recording was that of privacy, as it was possible to record without other people noticing, or participants sometimes forgot the recorder was running. Household 2 discussed the fact that they hadn't been worried about privacy during the experiment, since it was an experiment, but that outside of the experiment situation it would be more of a concern. They discussed the fact that it is important to ask peoples' permission before recording them. Household 1 was also slightly concerned with privacy. They tended to tell people after they finished recording to make sure they realised they had been recorded. In addition, at the end of one recording they were heard to say "I'll turn this off now because this is too private". In contrast, Sarah (H3) said she had

not asked permission to record because all her recordings were of relatives and she felt that they would not mind. In one recording, Amy (H4) had forgotten the recorder was running, and became embarrassed when listening back to herself:

"I'm interested in the noises Joshua makes, but, because when you talk to him, you don't care really...well, if I know someone's listening, I wouldn't do that. So I think that's why it's embarrassing, because I either didn't know it was being recorded, or had forgotten, and when I hear it back, I think, oh, I sound so stupid."

Although recorders being left on or activated by accident can cause embarrassment, they can also lead to recordings which can later become special. In Section 3.2, we reported a number of accidental recordings from the days of cassette recorders which ultimately were very precious to the families. In one recording for the study, Household 2 started, then forgot about the recorder and, thus, captured a regular day, which otherwise would not have been captured. Although the participants did not report this as particularly special to them now it is possible that, over the years, like the accidental recordings mentioned earlier, this snapshot of everyday life might take on some sentimental value, as may the recording mentioned above, even though it is currently perceived to be embarrassing.

Participants generally found the Sony sound recorders easy to use and did not report any major difficulties.

Participants' experience of capturing audio revealed a great deal about the dynamics of capture and the processes which come into play before, during and after the 'record' button is pressed. Whereas taking video or photographs is an established practice, all participants had to negotiate the new ground of discreet audio capture, figuring how and when to ask for permission and dealing with the difficult issue of privacy. However, for all the problems, participants perceived the benefit of a less intrusive, more naturalistic recording.

3.7 Playback, archiving and sharing

While participants clearly valued their recordings, suggesting that sound does indeed have potential as a form of sentimental memory capture, a number of research questions remain. For example, what do participants want to *do* with audio recordings, how do they wish to store and display them, do they wish to share them? In the following, we explore each of these areas and report on participants' reactions to the prototype sonic gem bowl.

3.7.1 Storage and playback

We asked participants to tell us what they imagined doing with the sounds which they had captured, and what they would like to do with any sounds they might capture in the future. Each household stated that keeping sounds filed away on a computer would be one very likely course. This was perceived as a safe and convenient form of storage, with the benefit of quick potential access to sounds. However, such a form of storage would require the effort of labelling files, which was flagged up as a negative. Others suggested making themed CDs of particular times (H4) and using recorded sounds as ring-tones and alarms (H1, H3).

A wide range of views were expressed regarding how often participants felt they would like to listen back to sounds. Even though the idea of listening to unaccompanied sounds was a somewhat unfamiliar concept, most participants instinctively felt they would handle sounds in much the same way as their photograph collections, i.e. to keep them safely stored, and view them on an 'every now and then' basis. For example, Angie (H2) made the following comment:

"I think I'd like to keep them, have them, and occasionally play them back. We so rarely sit down and look at photographs even these days. We have them all digitally, so we don't print them out, so we don't look at them. And I have a sense if we did the same with sounds it would be the same".

In this example, sounds would be listened to infrequently, and safely stored until such time as they were needed. Sue (H1) stated that there were certain sounds she would listen to rarely, because of the emotional impact these sounds have upon her. These were of family members who had passed on.

However, other forms of playback were suggested which would involve sounds being used in a far more regular way. As noted above, participants also suggested using sounds as alarms and ring tones in everyday items, as a form of sentimental personalisation. Sue (H1) exemplifies this:

"I think it would be nice if you could personalise the sounds that the computer makes anyway. Like, you know, a thing that starts off 'ba, ba' and windows comes on actually have like a member of the family saying 'hi' or something like that."

All participants stated that they would wish to play the sounds they had recorded of their children back to them at certain points in the future, for example at weddings, birthdays or other significant events.

As noted in the previous section, participants often stated that they wished to have short recordings. Mark (H2) noted that the challenge with displaying or listening back to audio recordings will be to extract the interesting short fragments of sound, the "sonic gems", from the long recordings. He notes that current audio editing tools are too complicated, and it would take too much time. He notes that they find it difficult to find time to even label their photos, and this would be even more time consuming.

To prompt people to think about sound playback, we used the Sonic Gems bowl to present their sounds to them. We selected short sound clips from within their recordings and put these in the bowl. One of the main elements of the sonic gems bowl which participants found appealing was the short sound fragments. We had to take a guess at what would be special, but in some cases we had managed to select particularly emotive fragments and the participants responded extremely well to the sounds. This was mentioned earlier, in Section 3.4, where Mark (H2) was entranced by the short sound clip of his child talking about Star Wars, whereas he had not noticed the same clip in the context of the longer recording. Mark also noted that he liked the fact that he could easily find that clip to play again by picking up the same object:

"The thing that was cool about that was just going, 'oh, that's that sound I heard, I want to hear it again, I want to hear it again, and the good thing about that was once you located that sound it was very easy to play it again and so that was the key thing about that."

Another comment about the Sonic Gems bowl was that it would be important to be able to mark the objects so you knew which sound was associated with which object. Participants also responded well to the idea of associating the sounds with relevant objects (rather than just the identical plastic eggs used in this prototype). Angie (H2) commented that the experience of interacting with sounds in the bowl was different to listening back to them on a PC and said that "it makes the sound totally tangible".

Several participants said that if they had realised that their sounds would be put into a device like this, they would have recorded more sounds. This indicates that by showing the bowl to participants they gained new insights into ways in which sound could be displayed, which then inspired them to think of the kind of sounds they might want to display.

3.7.2 Sharing special sounds

The data gave us insight into the types of sounds which participants would like to share, how they would like to do this, and also an idea of those which they would not like to share.

The sounds of children were often highlighted as sounds that participants would like to share with family and close friends. However, for Mark (H2), the method through which he would share sounds of his children arose as a significant issue. Mark had great reservations about sharing these sounds over the internet as he felt that in digitizing the data, he would effectively lose the power to control what happened to it. Because of this, he would only send sounds to a very small and trusted group of friends of family.

The sounds which people would like to share often consisted of their children saying, or sometimes singing, something short and appealing. For example, Mark (H2) stated that he would like to share the sound of his daughter saying 'Luke Skywalker' and singing her own version of the song Barbie Girl. Apart from children, Sue (H1) stated she would like to share the sounds of birdsong recorded in their back garden with her sister in Mexico (a sound she felt her sister would miss). The recordings which participants expressed an interest in sharing were all short clips.

There sounds which participants felt they would not like to share, or would not be comfortable in sharing with other people. Sarah (H3) stated that she would like to share the sounds of a drunken evening from holiday with the people who were there, but not with those who were outside the group. This is largely because she felt the recording portrayed her in a negative light, and therefore would not like those who were not involved in the night to hear it.

3.8 DISCUSSION OF FINDINGS

At the beginning of this study we set out to investigate the potential of audio as a form of sentimental memory capture. Our findings, gathered both from interview data and participants' recordings, indicate much scope for sound in this area. Indeed, sound recording already appears to be an established, but largely invisible, practice, with participants recording and sharing small fragments of sound via mobile phones, or recording them as ring tones. If sound recording has a stake in the present, it appears to have had a larger presence in the past. Certain recordings once made on tape recorders still survive, and have become very precious to their owners, whether or not they have devices to play them on. By offering participants the opportunity once more to be able to record sounds we uncovered a set of sounds which sample members desired to keep and share.

Audio only was found to have two major sets of affordances marking it apart from video and photography. One of these related to the capture of sound, the other to its playback. To begin with the former, audio recording proved to be a very low impact form of capture. In some senses, this aspect of capture served to recast the role of the recorder from observer, who usually stands *behind* the camera (and thus *outside* the event) to participant, whose involvement in the event is less encumbered by the duties of framing, focusing and following the action. As a result, sample members often found audio to be a more naturalistic form of capture, requiring less attention than video

and photography and thus more able to capture authentic and spontaneous moments of everyday life. The second major affordance of audio revolves around participants' experiences of re-listening to recordings. Listening back could often be a highly immersive experience, giving rise to vivid mental images of events and recapturing lost moments. The data suggests that one major feature in audio is its lack of visual stimulus, which focuses attention more directly on the sound, thus taking the listener further into memory and imagination as he or she mentally recreates the scene.

The findings also revealed difficulties which participants encountered whilst using, or thinking about using, audio in this way. Issues such as privacy, and the best way of archiving, editing, sharing and displaying sounds were all flagged up by participants as important. Such issues present significant challenges and opportunities for the design of audio recording and related technologies. These are discussed now.

4. DESIGN CHALLENGES

One motivation for this study was to understand the design implications for the capture, archive and playback of sonic gems. While the creative acts of everyday sentimental photography and video recording are well established and supported, the equivalent area of audio recording is far less so. In the field of digital photography, new technologies, e.g. digital photo frames, are becoming available to support the display of personal photos, and editing technologies for home use are widely available. Audio display technologies, in contrast, are almost exclusively for music use, and current audio editing software tends to be designed for music or radio professionals and neglects the home market, particular for non-music related audio. From our data we see that personal audio recordings can have high sentimental value, and be very precious to users. We, therefore, believe that it is important to consider how we can support users in capturing, editing and playing back their "sonic gems". The following sections outline some of the challenges that our study has highlighted which should be considered when designing such technologies.

4.1.1 Capturing Sonic Gems

The digital voice recorders used in our study proved to be appropriate for the capture of Sonic Gems. The participants noted that the one-click record should be present in any such recording device so that it can be activated easily without interrupting the activity. The device should also be small and discreet enough to be carried in a pocket. Two families had previously recorded on their mobile phones and integrating the recorder into a phone or MP3 player is an option since these are generally carried with people at all times. However, the current quality of many mobile phone audio recorders is quite poor and this would need to be improved before it could be reliably used for capturing sentimental sounds. In addition, the option to have one-click recording even on a phone would be important.

Participants particularly valued the short sound snippets extracted from recordings. The temporal nature of sounds means that they are difficult to "frame". One solution to this might be a device that allows sounds to be recorded retrospectively (e.g. [5, 10]). Such a system could record audio all the time but only store when initiated by the user. Pressing the "store" button could store the last 30 seconds recorded and thus allow participants to capture and retrieve special moments that might otherwise be missed

4.1.2 Archiving Sonic Gems

One challenge when storing a large number of audio recordings will be the ability to locate the desired recording. Therefore, it will be necessary to find a way to enable people to label or

annotate their recordings to allow for simple browsing or searching. However, several families commented on the fact that they find it hard to find the time to label their digital photo collections. Labelling audio recordings will be even more important since there is currently no audio equivalent of a thumbnail to help people to identify recordings quickly. This poses interesting challenges to the design community in considering how one might provide automatic labeling or categorization sounds, and even how one might think about designing auditory "thumbnails" for sentimental audio. Such concepts exist for providing thumbnails for music [1, 2], but may take some adaptation to work for non-musical sounds, much as photo editing tools are currently used.

Another challenge in the archiving of sentimental audio recordings is to find a way to allow users to easily edit sounds and grab the "sonic gems" from long recordings. Current audio editors are generally designed for musicians, composers and sound designers and are too complicated for most typical users wanting to extract a sound clip from a recording. The important aspect here will be to design a simple sound editor that allows people to grab the start and end of an important section and extract it. Such technology exists but needs to be repackaged for this purpose.

4.1.3 Playback and Display of Sonic Gems

A final challenge will be to explore how sounds might be played back or displayed. The participants in this study responded well to the Sonic Gems bowl, as it gave them a new way to explore short clips from their sounds. This tangible kind of interaction is just one way in which people might want to playback their sounds. Other options which could be explored would be audio equivalents of digital photo albums, the ability to make sentimental audio "mix tapes" with ease to send to family and friends, or some kind of small portable audio playback devices that could be carried in a wallet, as photos are at present. Participants in this study also expressed a desire to be able to attach sounds to existing devices, e.g. as an alarm clock sound or phone ringtones. Of course, some of these are already possible, but it could be interesting to explore how to make this process simpler.

One other concern regarding playback was that of privacy when sharing clips. One participant was concerned about sharing sound electronically as he would then lose control of what was done with them. It may be interesting to explore whether it would be appropriate to apply some type of copy protection, like that used with digital music, to such recordings to overcome these concerns.

5. CONCLUSIONS AND FUTURE WORK

Every moment of every day, we are surrounded by sound. These sounds can be memorable or forgettable, annoying or beautiful, yet for all their power, we rarely record them. And when we do, it tends to be for the most functional of reasons ('this call may be recorded for training purposes'). Our study begins to reveal the potential of sound as a form of sentimental memory capture, whilst also illuminating the opportunities and challenges that this area offers and presents. Our paper opened with a discussion of the cassette recorder, and our findings reveal that sample members still owned, and sometimes still listened to, old recordings made with these devices. Anachronistic as they are, these old audio tapes, and the impact they have on participants, speak of the significant potential of sound as a form of sentimental memory capture. Our other findings support this. By mining the audio memory, we uncovered sets of sounds which participants remembered warmly and wished to hear again. Through the recording exercises, we revealed the different sounds which our

participants wished to keep, either to personalise objects, share with friends and family, play in the future or safely archive. In playing sounds back to participants we highlighted some of the unique affordances of audio in this area, illuminating the specific ways in which recorded sound provided scaffolding for memory and imagination, and how and why this differed to other more traditional forms of capture. The experience of capture, too, proved to be significant, often offering a more immersive and naturalistic method of recording than photography or video.

This study was quite small in scale and as such only begins to provide insights into what sounds people want to capture, and what people want to do with the sentimental audio that they record. From such a small scale study, it is not possible to generalise our findings with confidence. However, future work will conduct a longer term study with multiple and diverse users, over a number of months. During this time participants will be asked to capture and archive their sounds, and regular visits will be carried out to explore how often sounds are captured, whether these sounds are listened back to, and, if so, how often. This will provide greater insights into how audio capture might be used alongside other media over a longer time period. We shall also begin to address some of the design challenges already posed in this work.

ACKNOWLEDGMENTS

The authors would like to thank the participants who took part in this study. Thanks are also due to Abigail Sellen, David Frohlich and all in the Socio-digital Systems Group at Microsoft Research for their valuable input to this research. This paper is based upon work completed while Gerard Oleksik was employed at Microsoft Research, Cambridge.

REFERENCES

- [1] Bartsch, M.A. and Wakefield, G.H., "Audio Thumbnailing of Popular Music Using Chroma-Based Representations". *IEEE Transactions on Multimedia*, 7(1), 2005.
- [2] Burges, C.J.C., et al., *Using audio fingerprinting for duplicate detection and thumbnail generation* in *Proc. Acoustics, Speech, and Signal Processing* 2005, IEEE.
- [3] Frohlich, D.M., "Audiophotography: Bringing photos to life with sounds". 2004, Kluwer Academic Publishers.
- [4] Frohlich, D.M. & Murphy R. (2000) The memory box. *Personal Technologies* 4: 238-240.
- [5] Hayes, G.R., et al., *The Personal Audio Loop: Designing a Ubiquitous Audio-Bases Memory Aid*. in *Proc. Mobile HCI* 2004, ACM Press pp. 168-179
- [6] Hindus et al., *Ubiquitous audio: capturing spontaneous collaboration*, in *Proc Conference on Computer-supported cooperative work* 1992, p. 210-217
- [7] Kasesniemi, E. and Rautiainen, P., "Mobile Culture of Children and Teenagers in Finland", in *Perpetual Contact: Mobile Communication, Private Talk, Public Performance* Katz, J.E. and Aakhus, M.A. (Editors), 2002, Cambridge University Press. p. 170-192.
- [8] Kirk, D.S., et al., *Understanding Videowork*, in *Proc. CHI* 2007, ACM Press, p. 61-70.
- [9] Kirk, D.S., et al., *Understanding Photowork*, in *Proc. CHI* 2006, ACM Press, p. 761-770.
- [10] Oleksik, G., et al., *Sonic Interventions: Understanding and Extending the Domestic Soundscape*, in *Proc. CHI* 2008, ACM Press.
- [11] West, D. et al., *MEMENTO: A Digital-Physical Scrapbook for Memory Sharing*, in *Personal and Ubiquitous Computing*, 2007, Vol 11, No 4, p. 313-328
- [12] Wilcox, L.D., et al., *Dynomite: A Dynamically Organized Ink and Audio Notebook*, in *Proc. CHI* 1997, ACM Press, p. 186-193