'Observing' the Workplace Soundscape: Ethnography and Auditory Interface Design

Catriona Macaulay and Alison Crerar

Department of Computing
Napier University
219 Colinton Road
Edinburgh EH14 UK
Email: {c.macaulay, a.crerar}@dcs.napier.ac.uk

Abstract

This paper identifies a gap in the research agenda of the auditory display community – the study of work practice and the uses (current and potential) of the workplace 'soundscape'. The paper presents a case study derived from a one year activity theory-oriented ethnographic study of information gathering work at a UK daily newspaper. We consider the soundscape aspects of mediating collaborative activity in the newsroom, and conclude with a discussion of the issues arising from this attempt to utilise ethnographic techniques within the auditory display design domain.

1. Introduction

The auditory display community has made significant steps forward in identifying many of the design, usability and technical issues associated with the use of sound in the interface, as can be seen from the proceedings of past ICAD conferences. Concurrently, within the wider interface design community recent years have seen increasing interest in the need to understand the *context of use* [1] of both existing and future computer systems. A range of studies of *work practice* [2], utilising ethnographic and sociological techniques and concepts, have been undertaken, (for example [3, 4]). Methods like Contextual Design [5] and Literate Development [6] and tools such as the Activity Checklist [6], Scenarios [8] and Rich Pictures [7] have been developed to integrate such approaches into design lifecycles. This paper proposes that work practice studies have much to offer the designers of auditory interfaces and that the auditory aspects of work have been largely over-looked. This lack of attention can be in part attributed to a paucity of analytical tools with which to undertake such research, and in part to the visual bias of Western culture [8]. This paper reports an activity theory-oriented field study of information gathering work at a UK national daily newspaper. The study has been conducted within the context of a project concerned with opening up a 'space of design possibilities' for auditory interfaces in the information gathering domain. As such it is an explicit attempt to investigate the problems and benefits of adopting contextualised approaches to the design of auditory interfaces. The rest of the paper is structured as follows. The next section outlines the methods used in the work presented. Section 3 provides a necessarily brief introduction to the fieldwork setting, and in Section 4 the issue of how to think about the auditory aspects of real-world activities is addressed. Section 5 presents an illustrative case study from our fieldwork, and the case study is developed in Section 6 where a space of design possibilities for supporting one particular aspect of information gathering in the newsroom is discussed. The paper concludes by considering some of the questions this work has prompted.

2. Methods

2.1 The Case for 'Contextualising' Interface Design

Since the late 1980's there has been a growing discontent with HCI's reliance on the cognitive science model of human behaviour [9]. It would be foolish to completely discard all that the cognitivist approach has to offer [10], but few would deny the validity of the argument that we can no longer ignore the cultural, sociological, political and historical influences on human behaviour. That is to say, it has been recognised that effective design requires an understanding of the context of use of computer systems since all computer use is 'situated' within a social, cultural, etc., context. One of the obvious methods Human Computer Interaction (HCI), and more frequently Computer Supported Cooperative Work (CSCW) researchers have turned to for insight into context and situated practice is ethnography. Ethnography literally means 'writing culture' and is based on the idea of the fieldworker going into some field (traditionally exotic locations, nowadays often closer to 'home') and attempting to produce a rich picture of what it is to be part of that culture [13].

2.2 Methods Used

A one year ethnographic study of information gathering at The Scotsman is being conducted using a mixture of participant observation (approximately 42 days over 10 months), interviews (31 so far, typically lasting around 30-40 minutes), documentary materials gathering and historical/desk research. Observations and interviews have been carried out in all of the principle editorial departments (news, foreign news, political news, business, features, sports, picture and production desks, and specialist writers) and in The Scotsman library, the IT unit, and with management. The applied nature of the ethnographic study would tend to point towards the usefulness of conceptual or analytical frameworks as a way of orienting and focussing fieldwork. Whilst all fieldworkers will take some theoretical or conceptual framework into the field with them, these may be more or less explicit and more or less formal [11]. Although some ethnographers have a dislike of what is often called Grand Theory, it is ultimately theory which distinguishes ethnography from a journalistic descriptive account [12]. Clearly any ethnographer may appeal to a mixture of formal and informal theories. For this particular project it was decided draw explicitly on ideas from activity theory in particular (see Section 2.3 below) to explicitly orient both the fieldwork and the analysis, since activity theory provides a way of quickly focussing in on aspects that might be particularly relevant for design [13]. Edited field notes and interview transcripts were coded for themes, after multiple re-readings, using Ethnograph software.

The use of ethnography in systems design is not without its critics of course [14]. This is in part due to a traditional reliance on Positivistic methods and in part to the clash between the largely abstraction driven world of engineering and the traditional lengthy narratives of ethnography. A particular criticism seems to be the generalisability and reliability/validity of ethnographic 'findings'. Whilst for a Positivist, reliability is usually equated with repeatability, it would be an unusual ethnographer who would seriously entertain the notion that their fieldwork could be repeated by another ethnographer with exactly the same results [15]. Validity, on the other hand, is a question which many ethnographers have addressed [11, 16, 17]. The validity of an ethnographer's findings is something that has to be demonstrated, through making explicit the choices made in the fieldwork (for example theoretical choices), and the conduct of the fieldwork [15].

2.3 Activity Theory

It is outwith the scope of this paper to present a detailed introduction to Activity Theory (for excellent overviews see, [18, 19]), however a few key points will useful inform the rest of the paper. Briefly, activity theorists think of human behaviour in terms of an *activity hierarchy*, wherein activities are composed of actions, which are in turn composed of operations. An activity is usually undertaken by collaborating subjects, and is oriented towards some objectified motive. For example, a production team work together to create a film. Actions are consciously carried out and have a goal. So for example the cameraperson will pan the camera slowly along a wall in order to film a scene in which an actor runs alongside the wall. Actions are comprised of operations, which are unconscious and are oriented to the conditions under which they are made. So an experienced cameraperson will not need to consciously think about using their arm to pull the camera around on its tripod, one of the operations required to perform a pan. However if conditions change unexpectedly, for example because the tripod head has stuck a little, the operation will become a conscious action as the cameraperson attempts to compensate and keep the pan going. Activity theory is rooted in the idea that all human behaviour is mediated by artefacts and that all artefacts have both material and conceptual aspects [20]. Whilst in common-sense terms we might regard a pencil as a physical tool, its physical aspects are not its entire nature. A lump of wood with some lead through the middle is only a pencil if certain ideal properties, or significance's [21] are attached to it. In other words we cannot understand how a pencil mediates the activity of writing through considering its physical properties in isolation. Furthermore, *"sociocultural context shapes the selection of cultural tools"* [22] p26. Therefore in order to understand *how* an artefact mediates action we must understand both its physical and its conceptual properties, which in turn reflect a certain historical development. And in order to understand *why* a particular artefact is used to mediate action rather than another we have to understand the context of the actions in question. Activity theory in some ways overcomes the problem of deciding what 'context' is by identifying activity as context [20], and this use of activity as the 'unit of analysis' guides our observations [23] and allows us to bridge the divide between the socio-cultural and the cognitive aspects of human behaviour [20]. Thinking of information gathering in activity-theoretical terms encouraged us to focus on the different artefacts available in the newsroom and the ways in which they mediated information gathering, as well as on the socio-cultural and historical context of such mediated activity. As we shall see later, considering the problems existing artefacts for mediating information gathering is one way of identifying design opportunities and developing design ideas.

3. The Scotsman Newspaper

The Scotsman is a UK national daily newspaper in a group of three major and several local newspapers owned by The Scotsman Publications Limited, and based in Edinburgh, Scotland. The parent company had a change of ownership in 1995, which brought with it an increase in investment in the technological infrastructure of the group. Consequently the Scotsman became one of the first UK newspapers to provide desktop Internet access to all its editorial staff. The majority of Scotsman staff work at its offices in the group headquarters in Edinburgh.

The main newsroom is home to the business, news, foreign and specialist teams, as well as the picture desk. Large and old-fashioned (they are currently based in an imposing old building in the city centre), the main newsroom contains workstations for around fifty people. At the centre of the room is the ‘control’ area (referred to by some informants as ‘the Bridge’) where section editors[1] and sub-editors[2] sit and direct their teams and discuss developing stories amongst themselves and, crucially, with the editor. Before moving on to consider the specific aspects of the study that relate to the design of auditory displays, we present a brief ethnographic vignette by way of introduction to the field setting. This vignette condenses a number of key themes from observations and interviews conducted in a number of departments of The Scotsman. It ‘sets the scene’ for our later consideration of one particular aspect of newsroom life.

3.1 A Tale from the Field: Pulling it All Together

The Scotsman newsroom is never still, people are constantly moving around. Piles of photographs, newspapers and documents are carried around the huge room in a constant stream. Well trodden paths from area to area are smoothly traversed by staff seemingly well used to the minefield of cables, papers, books, chairs and sharp edges. Impromptu conversations occasioned by a meeting on one of these paths turn into heated debates as people sitting nearby chip in. Questions, thoughts, jokes and asides fly about. As the deadline approaches, the noise and movement levels increase. Then, suddenly, a lull. The first edition has left the building. From a flurry of talking, moving around, searching text and picture databases, formal and informal meetings, visits to the library, shouting, phoning, typing, editing of photographs, and playing with layouts a newspaper has been born.

James[3] is head of the business news, and as such he spends his working life ‘pulling together’ the daily section. Even before he gets into the office he will have been listening to the radio, reading papers or watching television. His ‘official’ work morning starts with a flurry of reading other newspapers, checking electronic information sources such as the ‘wires’ (online, real-time electronic news feeds which most major newspapers rely on as a major source for story ideas), news from the stock-markets and talking to colleagues and regular . From this he will generate a story schedule, a list of potential stories which he takes into the daily morning editorial conference. At conference he will learn how many of his stories will win the editor’s go-ahead, and how much space he has to fill that day. There will be several more such meetings, formal and informal, throughout the day, during which any one of ‘his’ stories may be dropped, or promoted to a more significant position. Some business news stories will even turn out to be interesting to the main news section, and he may lose them altogether. After morning conference, James will assign the stories on his schedule to one of his team of staff and freelance journalists. The business section has a large number of staff permanently based at The Scotsman’s offices in Glasgow and London, as well as several in Edinburgh, so for James a lot of this process will involve lengthy stints on the telephone. Like most section editors, James tries to provide his writers with as much information as he can when briefing them. Often he will have clipped stories from other papers, or have some suggestions for possible sources (human, and non-human) the journalist could try using. At this stage it can be hard to predict how the story will work out, so he cannot be sure how long the story ought to be. But journalists are only too aware of the vulnerability of all stories to sudden changes in length or emphasis during the course of writing, and of course many of the stories assigned in the morning will, for a number of reasons, be abandoned (referred to as ‘spiked’). It would be difficult to think of the story schedule at this stage as anything but the most tentative of plans.

Until the late afternoon James will now be, as he thinks of it, “pulling everything together”. On one side of James’s desk is a computer which is displaying information from a specialist stock-market information service, identical to the one stock-market traders use. On the other side, James uses his other computer to flick between the wire services, an application which tracks the production status of submitted stories and a text editor he uses when checking stories before passing them on to his chief sub-editor. A steady stream of press releases and office memos are brought to the bulging paper trays on his desk, and the floor around his desk is littered with discarded newspapers. The phone rings constantly and he regularly gets up from his desk to wander around the newsroom chatting with other section editors, to grab a quick word with the editor, or to chat to the rest of the business team. Contrary to our stereotypical images of the newspaper business, journalism is as much a performance as literary art. Newspapers are not so much written as they are ‘talked up’. Throughout the rest of the day James will be monitoring the various information sources he has available for new stories, checking on the progress of already assigned stories, and on events in other sections of the paper. He will use all this information to make decisions about which stories are failing to ‘stand up’[4] and ought therefore to be ‘spiked’, and which are turning out to be better than anticipated and ought therefore to be moved to a more prominent position in his schedule. He will read every story at some point in the afternoon or early evening, checking them for general slant, accuracy, etc. By around 6 p.m. James will have told most of his team the fate of the work they

---

[1] Section editors generate and assign stories, check that they are meeting expectations, negotiate with other section editors and the editors for the 'rights' to a story, space in the paper, etc.

[2] Sub-editors check copy for accuracy, house style, spelling and grammar mistakes, write headlines and lay-up pages.

[3] Names and some details have been changed to protect the anonymity of our informants.

[4] Stand up – refers to whether or not a story once progressed a little still appears worthy of publication.

have produced. Some will have been disappointed to hear their story has been spiked, others will have been asked to do a rewrite. For the rest of his day James and his team of sub-editors can now concentrate on editing the stories that have survived, adding headlines, making final decisions about size and page layout, selecting pictures or ordering graphics. Of course the world does not stop during this period, and so James will not escape the round of constant information source monitoring that is the most dominant feature of his day. He will still frequently flick between the stock-market and wires news services since events in the world outside can invalidate any one of the stories he has nurtured to this fairly late stage (hence he always keeps a couple of stories in reserve ‘just in case’). It can be an extremely stressful time, but by about 8pm or so he can be fairly certain that his section is shaping up for the first edition, which will be sent to the printers at around 10pm. With a bit of luck he might even get home fairly soon.

4. ‘Observing’ the Soundscape

4.1 ‘Voices in the Forest’

This paper proposes that understanding the ways in which sound does, or could, support work activities is a neglected but potentially important part of the auditory display community’s research agenda. But how are we to go about ‘observing’ the soundscape, and how do we use our findings in design? How can we think about the aspects of work practice that might be relevant to design and how can we design auditory displays that compliment and enhance he existing workplace soundscape rather than disrupt it? The remainder of this paper considers these questions through focussing on one particular aspect of information gathering at The Scotsman. The authors were brought up, and work, in Western societies. It has been argued that Western society is visually biased [46] – so much so that the idea that any other sense but vision could principally mediate our experience in and of the world is almost unthinkable. As with many occasions when one wishes to ‘think the unthinkable’, it is often necessary to look outside one’s own culture in order to get a fresh perspective on the matter. Therefore prior to going out into the ‘field’ a literature review was conducted that focussed on conceptions of sound and audition from not only psycho-acoustic perspectives, but from culturalist perspectives as well.

Whilst sound is undoubtedly under-attended to within anthropology and sociological studies of work, there are some role models which helped orient the subsequent study. Within anthropology a small sub-field has developed which devotes itself to exposing the visual bias of Western culture, for example [24, 25] . For example, studying the importance of birds to the Kaluli of Papua New Guinea, it was observed that *“When asked direct questions that include the name of a bird, the response “It sounds like X” is universally presented by Kaluli before any sort of “It looks like X” statement.”* [26] p.72. To the Kaluli the world is not primarily conceptualised and related to through vision, but through audition. These insights were not easily available to Feld, it was some time before he began to realise the importance of the auditory world in Kaluli society. He recounts spending a frustrating day with a member of the Kaluli, trying to get him to describe some of the many species of birds that lived in the rainforest that Feld was attempting to classify. In traditional Western style, Feld was attempting to develop a classification based primarily on describing the visual appearance of the birds. His informant, frustrated, finally turned to Feld and announced that *“…to you they are birds, to me they are voices in the forest”* [26] p.44. The Kaluli are, or rather were since the Kaluli as a distinct community have since disappeared, not unique. A study of the Kalapalo of north-eastern Brazil concluded that *“Through sound symbols, ideas about relationships, activities, causalities, processes, goals, consequences and states of mind are conceived, represented and rendered apparent to the world.”* [27] p.311. Writing about the Kalapalo’s narrative performance art, Basso comments on how striking the rich weaving of large numbers of sounds with the more straightforward narrative text of the story is; the sounds help the audience follow the story. For example, birds can appear either as birds or as representative of magical beings. The sound type used when the bird appears lets the audience know in which capacity the entity should be read as appearing. This concern with the cultural and contextual aspects of audition is mirrored in the ecological psychology [29] and acoustic communication [30] approaches, which highlight the importance of considering the real acoustic environment we live in; that complicated balance of multi-layered sounds that makes up the *soundscape.* For the Kaluli, living in a dense rain forest, the benefits of relying heavily on audition as they went about their daily lives were obvious. As both systems designers and users struggle to keep on top of the increasingly crowded information ‘forests’ many of us must now navigate, voices in those forests may well become a more accepted part of our working lives.

4.2 ‘Mapping’ the Auditory Environment

One of the problems with ‘observing’ the workplace soundscape during ethnographic fieldwork is finding ways to think about and represent it. For applied ethnographic work such as this it is particularly important that a way is developed of attending to the soundscape, since it is necessary to both consider the current soundscape and imagine future soundscapes with newly designed elements. Musical notations were, of course, developed to give humans a way to think and communicate about music, similarly film theorists have also had to develop ways of thinking and talking about the filmic soundscape. In order to address the auditory aspects of the field setting work by Chion, Schwartz [31, 32] and others was drawn upon to develop a simple ‘map’ of the soundscape. For example, in his theory of acoustic communication [28] gives centre stage to the informational aspects of sound, situating the listener’s subjective interpretations of the soundscape more clearly in the middle of the acoustic experience. [29] has suggested that real life soundscapes are composed of three layers of acoustical information: foreground sounds, contextual sounds which support the foreground sound, and background or ambient sounds. Listen to the sounds around you now. Some, such as the whine of a disc drive will be background sounds, providing reassurance or information about the state of the world. Others provide context, helping you orient to the nature of the environment. Foreground sounds such as the beep of the computer attract your attention. Isolated, intermittent sounds (as often found in interface designs) only provide a single layer of acoustical information, that of the foreground.

From these various approaches we have synthesised a ‘map’ of the soundscape therefore characterises the soundscape along three dimensions: sound type (speech, music, non-speech/everyday, non-speech/abstract), acoustical information level (background, foreground and contextual) and information category (visible/hidden/imagined entities/events, passing of time, position in space, patterns in entities/events and emotions.) In Figure 1 a 3D rendering of the ‘map’ is provided to try and give the impression of these various interacting soundscape dimensions – although as is obvious representing the soundscape visually is not an easy task! Whilst the parts of the soundscape are described at a high level, they are sufficiently detailed to allow us to develop an idea of the main elements of the soundscape. When approaching the design of individual elements in an artificial soundscape a framework such as Gaver’s [30] may be turned to for an appropriately lower level of detail. However, for our purposes this model provides a useful framework for investigating the way people hear and interpret the soundscapes, and not just the sounds, around them. This ‘map’ represents a first attempt at developing a practical analytical tool which could be used by any fieldworker attempting to describe a workplace soundscape, or by designers seeking to understand the ways in which their future system might fit into the workplace context. It can be used to add auditory aspects to ethnographic vignettes.

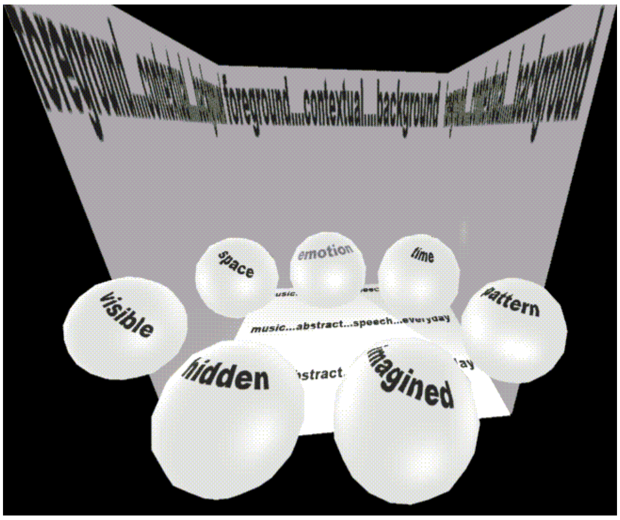

Figure 1: Map of the Soundscape

Let us return to the newsroom. It is nearly five o’clock, a busy time in the newsroom as section editors struggle to get all the necessary stories in from their writers so they can start the long process of editing. James has told one of his journalists to call him at five, when he will have time to chat a little about the story they have been having problems with. The hall is full of people as a shift change takes place. Everyone is busy, except for the sports sub-editor who has been held up at a crucial point in his day as he waits for a print to be located by the picture desk, and is tapping a ruler on his desk. The editor is wandering around asking people what they are doing, while on the political desk there has been an unexpected breakthrough in a story which the team are trying to digest. The soundscape might supply James with information about:

- Visible entities and events: e.g. The phone ringing on his desk.
- Hidden entities and events: e.g. The photocopier round the corner being used.
- Imagined entities and events: e.g. Something big is happening on the political desk (it has gone very quiet).
- Patterns of events/entities: e.g. Someone is batch copying a large document.
- The passing of time: e.g. It’s nearly deadline time (because the shift change is happening).
- Emotions: e.g. The sports desk sub-editor is unhappy (ruler tapping rapidly).
- Position in Euclidean space of entities/events and of the listener: e.g. The editor is behind me.

5. Case Study

5.1 Co-ordinating Newsroom Collaboration

As we have seen, section editors must balance numerous complex and interrelated factors when trying to pull their section together. Although James is identified (by himself and others) as the one who 'does' this pulling together, it is in fact a highly co-operative activity. A large number of people must complete their 'individual tasks' (writing stories, monitoring current and breaking stories, etc.), but such individual tasks cannot be completed without an awareness of what everyone else is doing. James is the hub of this co-operative web of activities, and as such rarely gets an uninterrupted period of time to focus on one task. On top of this, everyone involved must deal with the tension resulting from the pressure of deadlines. One of the ways in which James can be seen to cope with the conflicts between the need to attend to so many important activities at once, the need to get everything in on time and the interruptions he must inevitably endure is through the exercise of control, wherever possible, over how he spends his time. The exercising of this control is supported through the spatial layout of his immediate environment and in the everyday activities of the business section. James' desk is at the hub of the business section - between the writer's and the sub-editor's areas. Information flows within the section (phone calls, mail and press releases, conversations, newspapers, the wires and Reuters online information service) all revolve around and are easily available to him. He is ideally positioned to hear everything his colleagues say, is geographically at the heart of the section's space and within easy reach of all other business section staff. Of course, we do not wish to imply that there is some simple, and explicit, determining of human behaviour through spatial organisation of the physical space or through any formalising of the section's everyday activities. People can, and do, break with these practices. Rather we wish to suggest that considering the 'produced' spatial organisation [31] and lived practice of the business reveals much about the co-ordination of newsroom collaboration.

5.2 Outlouds

'Outlouds' were discussed in a study of stock-market dealers by Heath et al. [32], and have also been discussed in the context of air traffic control rooms [33]. An 'outloud' happens when someone shouts a piece of work-related information. They are a way of communicating information to a number of others, of maintaining peripheral awareness of goings-on in the workplace, and of initiating collaboration cf. [32]. Outlouds are a common feature of life in the Scotsman newsroom. As the following fieldnote extracts[5] demonstrate, they are used for a number of reasons:

> **Extract 1:** Sarah had been asked to hurry with the Stonewall story by a voice from the central area. Sarah is shouting out whilst typing 'what's the send command for the news desk?', someone shouts F9, then she shouts 'Stonewall'. The first voice asks for the name again, and she shouts Stonewall.

> **Extract 2:** The editor sits in 'the Bridge' (central area of the newsroom occupied by subs, section editors and the editor himself) and asks "So what's the story on X" (he's not looking at the person who answers). He listens to the response and then says "So the line is" and starts discussing whether that is the best line to take. Two other people in 'the Bridge' join in and they take about five minutes to agree. Editor: "Right then, so we'll just (give it the line)."

In the first example we can see two separate uses of the outloud in one. First is 'broadcasting a request for information' - in this case a request for help with a computer application command. Next is 'alerting a colleague to an action that is relevant to them' - in this case the completion and 'sending' of a story that has been previously chased up. In the second example we can see how outlouds are also useful as peripheral monitoring devices for others who may have a collaborative interest in an activity. The editor did not know who was responsible for the story so he simply announced his request to the people in the immediate vicinity. The response could be overhead by others, two of whom used this 'peripheral information' to join in the conversation. Such scenes are often observable in the newsroom, and as the day progresses *seemingly* impromptu 'outloud-initiated' conversations arise with ever greater frequency. One use of outlouding is to let the section editor know you have finished one of your assignments and it is available for editing:

> **Extract 3:** One of the reporters behind Mark shouts without turning round "Right you've got Widows" and Mark, without looking up/round, shouts "OK" .

Using outlouds to inform James of the availability of a story is highly efficient for both James and the sender. Neither need interrupt their current activity in order to achieve it. James can choose, as in the case, to

[5] Fieldnote extracts are edited versions of notes taken in the field. Text in quotes represents reported speech transcribed as it was spoken. Text in brackets are fieldworker comments, summaries.

offer some kind of verbal response. But since both sender and receiver are co-present this is not always necessary. The sender is aware of the intended recipient’s presence and activities, and can assume that no further action is necessary. In addition to choosing whether or not to respond with confirmation that he has received the message, James can choose whether or not to act. In this case he did not. Control over when to act on an outloud is very important. James must juggle many activities throughout the day and is constantly adjusting shared work plans. Outlouds are here functioning as both an efficient way to communicate information (without interrupting either the sender's or the receiver's current activity) and a mechanism for achieving control over the activity space. Thinking in terms of the soundscape map, this kind of outloud provides those who hear it with information about both visible and hidden entities and events, patterns, the passing of time and positions in spaces. This particular type of outloud is so efficient because the sender and recipient(s) are co-present, but can they be emulated in situations where the parties are physically remote? As we shall see, staff at the Scotsman attempt to do just this, using the telephone.

5.3 Remote Outlouds

The one aspect of James' daily life that he cannot control so easily is the telephone. The telephone, unlike some of his colleagues, is always responded to immediately. Telephone calls represent a significant challenge to James’ ability to control his activities. If we consider the activity ‘pulling a section together’, we find that James, as the subject, has an objectified motive of gathering and checking stories. Staff in the other offices can be regarded as in some ways trying to *emulate an outloud* when they phone James to tell him a story is ready. However, in this situation there is a problem with the way the activity of outlouding is mediated, since the communication requires that James stop what he is doing and attend solely to it. Whilst everyday practice at the Edinburgh office discourages journalists from asking for such feedback at inappropriate times, a telephone conversation opens up greater opportunities to elicit feedback, and since the journalist is not in the same office as James it is harder to know when might be the most appropriate time to ask for feedback. James rarely responds immediately to notification via co-present outloud, since he is often engaged in some other task at the time. However with what we may term 'telephone emulated-outlouds' he is forced to attend more *directly* to the communication, and is now vulnerable to being induced into dealing with the story at a time that may not be of his choosing.

Using the Activity Theory framework to think about the activity 'outlouding' we find two principle subjects - the sender and the receiver of the outloud. The sender's aim is to update the (co-present or remote) recipient whereas the recipient's aim is to gather the information necessary to let them update the mental and physical representations of the story list. Whilst outlouds function well when the initiator and recipient are co-present and therefore spoken language is the only mediating tool used, remote (i.e. telephonic) outlouds suffer from the introduction of the telephone as a support mediating tool. With the emulated outloud James can no longer orient his response effectively since he is now constrained by the nature of the telephone as a mediating artefact. Attending to the outloud now becomes an action rather than an operation, it has shifted up a level in his consciousness. He loses control over <u>whether</u> to attend to an outloud in the background or foreground - the telephone cannot easily be ignored until such time as James is ready to respond. Furthermore, he loses control over <u>how</u> to attend to an outloud. With a co-present outloud James may or may not immediately attend to the event (i.e. the sending of a ‘finished’ story to his workspace) announced. With a telephonic outloud he is more likely to be drawn into conversation about it, if not into actually reading it and providing feedback whilst still on the phone. There are alternatives to the emulated outloud, such as frequently checking the computer to see if any new stories have arrived or calling a reporter to ask if the story has been sent. However they are far less likely to meet the recipient's goals since they require that they *act* rather than simply engage in background monitoring of their auditory environment. Furthermore, the telephone-emulated outloud loses another potentially important side-effect of the co-present outloud, informing others in the area at the time of the current state-of-play. Journalists often develop a sense of whether 'their' stories will get into the paper by monitoring events around them, such as who is working on what stories. Outlouds are an important (and obvious) feature of the auditory environment in the newsroom and are a useful way of keeping in touch with all sorts of situations throughout the day.

6. The ‘Space of Design Possibilities’

Supporting virtual outlouds presents a challenge for the designer. Outlouds are not co-operative in the simple sense of joint attention to a task, rather they are a subtle communicative aspect of a wider group activity whose motive is the pulling together of the next day’s business section. This wider activity is in part achieved individually, in part mutually, but in all cases requires more or less explicit, and more or less on-going, communication between colleagues. “*Indeed, it may be the case that in attempting to support collaborative activity, we consider mutually focussed collaborative activity on a single task, as perhaps one of the rarer forms of real world co-operation”* [33], pp166-167. Within the media space literature it has been remarked that the telephone is perhaps the pre-eminent example of an audio-only media space [34]. However, as we have seen, the telephone does not always mediate collaborative activity appropriately. We can now turn our attention to media space (and related) research for guidance as to ways in which virtual outlouds might be more appropriately supported. One solution to the problem of emulating outlouds might simply be to provide a pop-up message on James’ screen. He would ‘get the message’ without necessarily having to act upon it immediately and without the danger of being drawn into inappropriate conversations. However James’ virtual and real desktop is a very cluttered place. The problems of living with the huge amounts of information now available to us have been the subject of much debate recently, with the phrase ‘information sickness’ entering the vernacular. But rather than thinking of the problem in terms of ‘information overload’, we can think of the problem in terms of ‘word overload’. The problem might not be too much information, but crucially, *the way that the information is presented* (typically as text). Within the auditory display research community the domain of ‘background monitoring’ of events and people has attracted much interest e.g. [35], [36]. And of course co-present outlouds are a well established feature of James’ auditory environment, hence it would seem appropriate to investigate the auditory presentation of virtual outlouds.

First on the list of ‘obvious solutions using sound' might be simply to have an ‘open mic’ at the remote offices allowing the people there to 'transmit' their outloud into the Edinburgh office. However there are a number of disadvantages with this approach. How will they know if James is actually at his desk? If he is not he will miss the message. Some kind of buffer to save messages could be added but this defeats the purpose of not requiring James to consciously attend to receiving (or in this case retrieving) the outlouds. Furthermore, if he is at his desk it may not be a good moment to outloud, for example he may be engaged in a lengthy conversation with someone else. Co-present colleagues can monitor (aurally and/or visually) the environment for cues as to when to issue an outloud, remote colleagues cannot. Using a two way audio channel might offer an opportunity to provide some means of signalling when James is and is not at his desk. Similarly, streaming the audio[6] at James’ end would overcome the problem of what happens if outlouds from more than one person/remote office are sent at once. We are now moving into the arena of the media space, and more particularly of the audio only media space. A good example of such a system is Thunderwire, a system which allows both speech and ambient workplace sounds to be shared between remote locations [34]. Thunderwire was found to have been successful in that it allowed users to engage in social conversations in an apparently 'natural manner'. However a particular set of problems were identified around how to deal with the ambient background noise the system provided. We would like to suggest that some of these problems might be attributable to the problems of reproducing real world sounds electronically, since of course it is not 'real' ambient sounds that are being transmitted across Thunderwire but electronic reproductions of them. As filmmakers found out many years ago, simply attempting to copy the way sounds appear in real life would result in a cacophony [38].

In some ways the telephone actually meets quite a lot of the requirements for a mediating artefact for virtual outlouds. A great many of the usual cues about whether the intended recipient is available, what else they are doing, and who the sender is are potentially lost when the actors are remote. The telephone in part overcomes these by ‘grabbing’ the recipient’s attention. An alternative system design must address these issues, and it is our contention that whilst the auditory aspects of co-present outlouding are speech-only, a speech-only audio-media space would be insufficient for a virtual outloud mediating artefact since there are numerous visual cues available to co-present outlouders that would need to be communicated in a non-attention grabbing way. Speech would simply be too cumbersome. Finally, any alternative design solution must address some important environmental considerations. Primary amongst these is that the auditory environment in the newsroom generally is important. As we have seen, a number of different kinds of outloud are issued in the newsroom and taken together are an important way for those working there to stay aware of peripheral events around them. In the heat of the newsroom what seems to be a peripheral event can all to quickly become highly relevant. Therefore an alternative mediating artefact that used headphones would be unsuitable. Additionally, James and his colleagues do spend time away from their desks, visiting other departments, chatting in the hallways, etc. A desktop system would severely limit their mobility.

Recent work on *nomadic auditory environments* [39] points the way towards the concept of a context/environment sensitive wearable audio media space device. Using a Nortel patented research prototype called the *Soundbeam Neckset,* Sawhney and Schmandt have come up with the metaphor of nomadic radio for systems which allow users to send and receive speech and non-speech audio. Messages of a particular type or from a particular source are streamed to provide a useful spatial metaphor, as well as to aid comprehension of simultaneous messages. The speakers and microphone built into the *Neckset* overcome the problems for the virtual outlouder of tying users to their desk and of headphones preventing the monitoring of the real-world auditory environment [41]. Whilst the technology is at a very early research stage and is still prohibitively bulky, the stated aim of those involved is to move to the point where the technology can be embedded in clothing or jewellery. Of course such solutions throw up some ethical and political considerations. Firstly, as we have seen whilst remote outlouds mediated by the telephone do not necessarily support James’ aims, they may support those of the sender. There is a conflict in goals here, and the mediating technology used to support outlouding will affect who ‘wins’. Secondly, the idea of being tracked around the office may well not be too everyone’s taste and there are potentially serious civil liberties questions if employers decided to use such technology for reasons other than supporting outlouding.

---

[6] Research into auditory stream segregation indicates that splitting simultaneous sounds into separate spatial auditory streams increases their comprehensibility [37].

Whilst a concern with the potential of sound in CSCW applications is not new [35], we do offer a clear example of a situation in which speech or the simple 'transmission' of ambient noise is not enough. It might be argued that the case presented here is highly specific to the Scotsman and that our observations regarding the need for appropriate technological solutions for the problem of virtual outlouding are therefore not generalisable. However, as was noted above, outlouds are a feature of collaborative working that have been observed in other settings. Perhaps the common feature amongst these settings is that they are all environments in which work is highly time constrained. This would suggest that our findings may be applicable in other highly collaborative, time constrained work settings where colleagues are physically distributed.

7. Conclusions and Future Work

This paper’s central claim is that work practice studies have a role to play in the design of auditory displays, just as they do in the wider HCI and CSCW communities. For entirely understandable reasons, auditory display research hitherto has largely been driven by technology-inspired design work and experimental projects addressing specific cognitive issues. It is proposed that the time has now come to widen the scope of the research agenda to address the issue of *context* in auditory display design. The methods and case study presented here represent an early attempt at explicitly dealing with the auditory aspects of the workplace, and the potential for auditory displays. A number of areas for further research have of course been identified, including developing a shared language with which to talk about the auditory aspects of work and the workplace soundscape – something necessary if comparative studies of future research in this area are to be undertaken. As far as we know, our work represents a unique attempt to tackle some of the issues associated with undertaking contextualised design projects for largely or solely auditory displays (although Ackerman et al, [34], have addressed contextual issues in the evaluation of an audio-only system). However because of the nature of our project we only go as far as developing a ‘space of design possibilities’ for information gathering systems which use (to a greater or lesser extent) auditory elements. Since our work has been exploratory in nature we have had the luxury of a relatively long time-frame within which to conduct our fieldwork, and no particular technical or design constraints. A very different set of issues may arise when attempting to adopt this approach in a specific design project. For example, participatory design [40] has often been the design method of choice for the ethnographically-inclined. However, participatory design techniques such as paper prototyping and storyboarding are obviously much more suited to GUI design than auditory display design and further work is needed to identify appropriate tools and methods for the participatory design of auditory displays.

8. Acknowledgements

The authors would like to thank all of our informants at The Scotsman for their time, co-operation and patience. We are especially grateful to Dr. Alan Munro of Napier University for his comments on outlouds.

9. References

[1] L. Suchman, “Making Work Visible,” Communications of the ACM, vol. 38, pp. 56-65, 1995.

[2] G. Button and R. Harper, “The Relevance of 'Work-Practice' for Design,” Computer Supported Cooperative Work, vol. 4, pp. 263-280, 1996.

[3] R. H. R. Harper, “Gatherers of Information: The Mission Process at the International Monetary Fund,” presented at Fifth European Conference on Computer Supported Cooperative Work, Lancaster, UK, 1997.

[4] J. Hughes, V. King, T. Rodden, et al., “Moving Out from the Control Room: Ethnography in System Design.,” interactions, vol. II, pp. 57-65, 1995.

[5] K. Holzblatt and H. Beyer, “Contextual Design: Principles and Practice,” in Field Methods Casebook for Software Design, D. Wixon and J. Ramey, Eds. New York: John Wiley and Sons, Inc., 1996, pp. 301-334.

[6] V. Kaptelinin and B. Nardi, “The Activity Checklist: A Tool for Representing the "Space" of Context,” Umeä University 1997.

[7] A. Monk and S. Howard, “The Rich Picture: A Tool for Reasoning about Work Context,” interactions, pp. 21-30, 1998.

[8] W. Ong, Orality and Literacy. The Technologizing of the Word. London: Routledge., 1982.

[9] J. M. Carroll, “Human-computer interaction: psychology as a science of design,” International Journal of Human-Computer Studies, vol. 46, pp. 501-522, 1997.

[10] T. R. G. Green, S. P. Davies, and D. J. Gilmore, “Delivering cognitive psychology to HCI: the problems of common language and of knowledge transfer,” Interacting with Computers, vol. 8, pp. 89-111, 1996.

[11] M. Hammersley and P. Atkinson, Ethnography: Principles in Practice., 2nd ed. London: Routledge, 1995.

[12] H. F. Wolcott, The Art of Fieldwork. Walnut Creek: Altamira Press, 1995.

[13] B. A. Nardi, “Activity Theory and Human-Computer Interaction,” in Context and Consciousness: Activity Theory and Human-Computer Interaction, B. A. Nardi, Ed. Cambridge, MA: The MIT Press, 1996, pp. 7-16.

[14] G. Mantovani, “Social Context in HCI: A New Framework for Mental Models, Co-operation and Communication,” Cognitive Science, vol. 20, pp. 237-270, 1996.

[15] R. Sanjek, “On Ethnographic Validity,” in Fieldnotes: The Makings of Anthropology, R. Sanjek, Ed. Ithica: Cornell University Press, 1990, pp. 385-418.

[16] K. Hastrup, A Passage to Anthropology. Between Experience and Theory. London: Routledge, 1995.

[17] D. Silverman, Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction. London: Sage Publications, 1993.

[18] S. Bødker, Through the Interface: A Human Activity Approach to User Interface Design. Hillsdale, NJ: Lawrence Erlbaum Associates, 1991.

[19] B. A. Nardi, “Context and Consciousness: Activity Theory and Human-Computer Interaction,” , B. A. Nardi, Ed. Cambridge, MA: The MIT Press, 1996, pp. 7-16.

[20] M. Cole, Cultural Psychology. A Once and Future Discipline. Cambridge, MA: The Belknap Press of Harvard University Press, 1996.

[21] D. Bakhurst, “Activity, consciousness and communication,” in Mind, Culture and Activity. Seminal Papers from the Laboratory of Comparative Human Cognition., M. Cole, Y. Engestrom, and O. Vasquez, Eds. Cambridge: Cambridge University Press, 1997, pp. 147-163.

[22] J. W. Wertsch, P. del Rio, and A. Alvarez, “Sociocultural studies: history, action and mediation.,” in Sociocultural Studies of Mind, J. W. Wertsch, P. del Rio, and A. Alvarez, Eds. Cambridge: Cambridge University Press, 1995, pp. 1-34.

[23] B. A. Nardi, “Some Reflections on the Application of Activity Theory,” in Context and Consciousness: Activity Theory and Human-Computer Interaction, B. A. Nardi, Ed. Cambridge, MA: The MIT Press, 1996, pp. 235-246.

[24] C. Classen, Worlds of Sense: Exploring the Senses in History and Across Cultures. London: Routledge, 1993.

[25] P. Stoller, The Taste of Ethnographic Things. Philadelphia: University of Pennsylvania Press, 1989.

[26] S. Feld, Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression, 2 ed. Philadelphia: University of Pennsylvania Press, 1990.

[27] E. B. Basso, A Musical View of the Universe: Kalapalo Myth and Ritual Performances. Philadelphia: University of Pennsylvania Press, 1985.

[28] B. Truax, Acoustic Communication. Norwood, NJ: Ablex Publishing Corporation, 1984.

[29] G. Ferrington, “Audio Design: Creating Multi-Sensory Images for the Mind,” , 1995.

[30] W. W. Gaver, “What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception,” Ecological Psychology, vol. 5, pp. 1-29, 1993.

[31] H. L. Moore, Space, text and gender: an anthropological study of the Marakwet of Kenya., 2nd ed. London: Guilford Press, 1996.

[32] C. Heath, M. Jirotka, P. Luff, et al., “Unpacking Collaboration: the Interactional Organisation of trading in a City Dealing Room,” presented at ECSCW '93, 1993.

[33] P. Luff, C. Heath, and D. Greatbatch, “Tasks-in-interaction: paper and screen based documentation in collaborative activity,” presented at CSCW 92, 1992.

[34] M. S. Ackerman, D. Hindus, S. D. Mainwaring, et al., “Hanging on the 'Wire: A Field Study of an Audio-Only Media Space,” ACM Transactions on Computer-Human Interaction, vol. 4, pp. 39-66, 1997.

[35] J. Cohen, “Monitoring Background Activities,” in Auditory Display; Sonification, Audification and Auditory Interfaces (Proceedings of ICAD 92, Santa Fe), G. Kramer, Ed. Cambridge, MA: Addison-Wesley, 1994.

[36] J. A. Ballas, “Delivery of Information Through Sound,” in Auditory Display; Sonification, Audification and Auditory Interfaces (Proceedings of ICAD 92, Santa Fe), G. Kramer, Ed. Cambridge, MA: Addison-Wesley, 1994.

[37] A. S. Bregman, Auditory Scene Analysis. The Perceptual Organization of Sound., 1st paperback ed. Cambridge, MA: The MIT Press, 1994.

[38] K. Reisz and G. Millar, The Technique of Film Editing. London, UK: Focal Press, 1968.

[39] N. Sawhney and C. Schmandt, “Design of Spatialized Audio in Nomadic Environments,” presented at ICAD 97, Palo Alto, 1997.

[40] A. H. Namioka and C. Rao, “Introduction to Participatory Design,” in Field Methods Casebook for Software Design, D. Wixon and J. Ramey, Eds. New York: John Wiley and Sons, Inc., 1996, pp. 283-300.