WORKPLACE SOUNDSCAPE MAPPING: A TRIAL OF MACAULAY AND CRERAR'S METHOD

Iain McGregor, Alison Crerar, David Benyon and Gregory Leplatre

Napier University, School of Computing
10 Colinton Road, Edinburgh, EH10 5DT
{i.mcgregor, a.crerar, d.benyon, g.leplatre}@napier.ac.uk

ABSTRACT

This paper describes a trial of Macaulay and Crerar's method of mapping a workplace soundscape [1] to assess its fitness as a basis for an extended soundscape mapping method. Twelve participants took part within 14 separate environments, which included academic, commercial and domestic locations. Results were visualized and subsequently collapsed to produce typical responses to typical environments, as well as specialist responses to a shared workplace.

1. BACKGROUND

Macaulay and Crerar's method [1] was method was chosen for this study as it addresses the mapping of auditory environments from a human computer interaction perspective, rather than the more commonplace acoustic ecological perspective. The original authors identified a 'gap in the research agenda of the auditory display community' and attempted to 'utilize ethnographic techniques' rather than the traditional 'cognitive science model' in order to fill this 'gap'. The method takes the form of a 'context of use' through 'activity' in the form of an 'analytical tool' where each sound event is classified according to its sound type, information category and acoustical information providing a form of metadata (Table 1).

It also goes further than a traditional Gestalt figure ground approach by having a third contextual dimension which can be utilized as level of listening. This third layer provides an insight into auditory monitoring, that is a sound event which provides information about what is happening around a listener without actively engaging with it, such as is required when monitoring the output of a computer printer while actively listening to music.

Sound Type	**Example**
Music	Any type of identifiable music, radio/stereo
Abstract	Unusual sounds not normally experienced, video recorder chewing a tape
Speech	Conversation
Everyday	Identifiable recognised sounds.
Information Category	**Example**
Visible entities and events	The phone ringing
Hidden entities and events	The photocopier round the corner being used
Imagined entities and events	Something big is happening on the political desk (it has gone quiet).
Patterns of events/entities	Someone is batch copying a large document
The passing of time	It's nearly deadline time (because the shift change is happening)
Emotions	The sports desk sub-editor is unhappy (tapping)
Position in Euclidean/ acoustic space of entities/ events and of the listener	The editor is at the foreign desk behind me (can hear his voice)
Acoustical Information	**Example**
Foreground	Computer beep to attract your attention.
Contextual	Door opening (Help you orient to the nature of your environment.)
Background	Whine of disk drive providing reassurance or information about the state of the world.

Table 1: Macaulay and Crerar's Workplace Soundscape Mapping Tool Questionnaire

This tool was originally designed to be utilized by fieldworkers and designers, in order to preview the 'workplace context' creating a 'rich picture' prior to the introduction or development of an auditory interface or system. It was developed during a yearlong ethnographic study at *The Scotsman* offices in Edinburgh, and is based on the work of Ferrington [2] for the acoustical information as well as Truax [3] and Chion [4] for the sound types and information categories. The resultant map could then be used 'to add auditory aspects to ethnographic vignettes, as well as providing a 'shared language' which would facilitate 'comparative studies'.

One of the key elements not addressed by Macaulay and Crerar was the end user or inhabitant. Each individual experiences a unique soundscape, based on their previous experiences and interests, and as such will provide unique responses, what can also be termed as a Rashomon phenomenon [5]. Maps created by multiple inhabitants can provide a further insight into the typical versus the individual experience. The designer's perspective can be compared to that of individuals, or a typical response for a specific environment, or a typical response to a typical room. This would allow an anthropocentric approach to the design of auditory systems suitable for shared auditory environments. An evaluation of their method was required, as no examples had been included in the original paper, or subsequently published.

2. METHOD

This preliminary study took the form of fourteen different maps, using thirteen participants in twelve individual locations, these were divided between *regular*, *intermittent* and *new* inhabitants (Table 2).

Participant	Location	Environment	Type of inhabitant	Duration	No. Events	Notated by
SL	Library	University	Intermittent	15 mins	8	Participant
GC	Computer Room 1	University	Intermittent	15 mins	8	Participant
JN	Computer Room 2	University	Intermittent	15 mins	7	Participant
GA	Computer Room 2	University	Intermittent	15 mins	11	Participant
MB	Staff Canteen	University	New	40 mins	16	Participant
EM	Staff Common Room	University	New	30 mins	21	Participant
MP	Computer Room 3	University	Intermittent	15 mins	17	Participant
KM	Computer Room 3	University	Intermittent	15 mins	8	Participant
CR	B/W Darkroom	Photographic Lab	Regular	55 mins	25	Observer
RW	Colour Printing	Photographic Lab	Regular	85 mins	35	Observer
FD	Reception	Photographic Lab	Regular	20 mins	20	Observer
DK	50's style Diner	Diner	Regular	180 mins	59	Observer
IM	Kitchen	Domestic	Regular	180 mins	53	Participant
IM	Study	Domestic	Regular	60 mins	46	Participant

Table 2: Subjects and locations for trials of Macaulay and Crerar's method

Regular denoted that the participant regularly spent many hours in that environment, either due to it being their workspace or in the case of the domestic environments it formed part of their home and their familiarity was therefore high. *Intermittent* denoted that participants had visited the space periodically, but it did not constitute their main work or home area. *New* described a participant who was new to the environment, having never entered it before. This method of describing was not part of Macaulay and Crerar's method, it has been added in order to notate a participant's familiarity with an environment.

The first iteration involved eight individual participants in one of six different University environments. Each participant was asked to list in written form all the sounds they could hear in their local environment within a minimum of a 15 minute

period. All of the eight participants spontaneously closed their eyes and stopped what they were doing in order to consider their responses. They subsequently opened their eyes in order to confirm what they had heard in each instance, before re-closing them in order to continue. The majority of the participants missed at least a few of the sound sources or events. One major consequence was that participants stopped creating any noises themselves in order to listen more carefully, thereby omitting a major contribution to their personal soundscapes. Each sound event was later classified according to Macaulay and Crerar's method and then visualized as detailed in section 2.1.

To compensate for the problem of participants dimishing their usual auditory environment, by stopping what they were naturally doing and closing their eyes, a further four participants were observed while working. Notes were made of all the sound sources/events by a trained observer (sound engineer), with classification taking place after the prescribed period. During questioning, which took place in the same environment, if the participants were not aware of any of the sounds it was notated as 'not aware' and omitted from the subsequent map. Subjects were also asked if there were any sound events which they wished to add due to possible omission by the observer, no additions were suggested.

The final two studies, were conducted in a home environment, with the participant making a record of each new sound that was heard, over a specified period of time, as normal work was conducted. This enabled a longer time period to be studied without the need for an observer, which might potentially alter their behaviour or working practices, although this constant monitoring of the sounds was still an artificial condition.

2.1. Visualization

In figure 1 each concentric circle represents the acoustical information with *foreground* being located in the centre. The seven segments on the left represent the information category, as labeled.

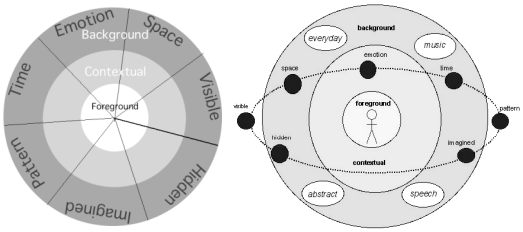

Figure 1. *Pictorial Representation of data, based on an original map by Macaulay and Crerar (unpublished)*

The sound type, was notated by the labeling of each 'bubble' with a symbol. *Music* was a couple of notes ♫, *abstract* a series of numbers **123**, *speech* a series of letters **abc** and *everyday* by an everyday item in this case an apple which appears to have had a bite taken out of it . Sound events were cross-referenced to letters within each 'bubble' to help prevent the image becoming too cluttered by confining the contents to a letter and a symbol, rather than a textual description of the source and event. The visualization did not use colour for individual maps, this was confined to maps with aggregated responses, allowing easy differentiation between the two different types. The individual colours in the latter case either represented different participants' responses, or the quantity of responses for each sound event. Each source or event was placed into the diagram according to the participant's responses.

2.2. Environments

Locations studied included a library, computing rooms, a commercial photographic processing laboratory as well as a 50's style diner, domestic kitchen and study. The intention of this diversity was to establish if the specified categories were appropriate for a variety of auditory environments, and to establish if any additional categories or layers would come to light through the questioning.

2.2.1. Academic Environments

The first type of academic environment chosen was one of Napier University's libraries. This environment was a mix of traditional library, and computer workstations. In order to test the method to create a typical response to a typical room, three separate 50 seat computing rooms, which were occupied for practical work, were mapped by five individuals, each contributing once. The last two academic environments studied a the staff canteen and common room, these was chosen to contrast with the previous work orientated environments, as well as to elicit responses from new inhabitants, as both participants had never previously visited either location before.

2.2.2. Photographic Processing Laboratory

A city centre photographic processing laboratory was chosen to represent a shared workplace environment in which a wide variety of tasks take place. The layout within the processing area was open plan, with small internal rooms leading off for specialist tasks such as printing and film development. The reception area adjoined the processing area through a communicating door, and the main entrance to the lab was via large glass doors, which opened onto a quiet lane.

2.2.3. 50's Style Diner

A busy 50's style diner located on a main road in Edinburgh's city centre was observed in order to represent a completely different type of workplace environment. A central kitchen and bar area was surrounded on three sides by customer seating for up to 80 diners. Typically up to five people worked in this environment, sharing all of the tasks from food preparation through to washing dishes, serving drinks and calculating bills.

2.2.4. Domestic Environments

Two domestic environments were included to evaluate the method's suitability for notating domestic as well as workplace environments. The first was a kitchen over a busy three hour time period preparing a Christmas dinner. The kitchen was located on the top floor at the rear of a five-storey building, with few exterior sounds penetrating the double-glazing. The second environment was a home office or study within the same building, this contained many of the elements found in the University computing rooms.

3. RESULTS

The sounds listed by the participants were classified during interviews with the first author using the Macaulay and Crerar model. All of the responses fitted easily into the categories supplied, all of which were applied. Within the sound types, *music* and *speech* were readily understood, and applied,

although sometimes if the speech was a *background* sound then if was classified as *everyday*. *Abstract* and *everyday* were not consistently applied, there were three main interpretations applied, the first was that of *abstract* representing a sound that the participant was familiar with, but thought was unusual, or differed from the norm for that particular object. With *everyday* being applied to a sound which was closer to what they had expected. The second was in terms of natural or artificial, *everyday* representing natural, and abstract representing artificial or man-made. The final interpretation was that of other known for *everyday* and other unknown for *abstract*.

Within the information category *visible* was applied to sound sources which could either be seen, or where a source could be identified, even if it were not immediately visible. *Hidden* was referenced for both sound sources which were not visible as well as when the specific sound source was visible but could not be isolated. *Imagined* was applied when an estimate was being made as to the source rather than trying to interpret the meaning of a sound, or lack of sound. *Pattern* denoted either a series of connected sounds over a short period, or an irregular long-term sound. The *passing of time* was applied only once, and referred as a reminder that it was time to go home. *Emotions* were routinely used when referring to speech when it was a *contextual* or *foreground* sound event. It was also applied to actions which informed others about someone's mood, when referring to impact sounds. *Position* was applied to moving objects rather than marking where a stationary object was located.

The application of acoustical information did not match with the original aims of the Macaulay and Crerar paper [1], in that this information was intended to illustrate the richness of the information being gathered, a foreground sound provides very little about what is going on in the world around you, such as a 'beep' whereas contextual informs you about what is going on contextually in your acoustic environment, and background provides reassurance about what else is occurring in the vicinity. The results more closely represent levels of listening as suggested by Amphoux [6], where foreground sounds were actively monitored and interpreted (sonic symbols), contextual sounds told the participants about the place they were inhabiting (sonic ambience) and background was applied to sounds that were not paid attention to.

3.1. Typical Environments and Responses

Within the twelve environments studied, three were computer labs which five respondents experienced. These were combined to form a typical soundscape map of a typical computer lab (Figure 2 and Table 3). A note was made of the number of responses and then the descriptions were collapsed and sequenced according to the amount of responses in descending order. Colours were added in order to give a precise figure of the number of responses per sound event, this started at the red end of the spectrum for the highest value of six, and worked through to yellow for single responses.

The combined results show a fairly even mix of acoustical information, which would be expected in a shared environment. There were similar amounts of *visible* and *hidden* sound events, although all of the *hidden* were *background*, and most of the *visible* were *foreground*. *Emotions* figured next highest, all of which were associated with people either talking, whispering, laughing, coughing or sighing. Only the laughter was *foreground*, everything else was classified as either *contextual* or *background*. Illustrating that the variety of forms of vocalization did not intrude upon the shared acoustic environment to any real extent. Within the sound type *everyday* formed the largest group, which showed familiarity with the environment. This was confirmed through only a couple of the events being *imagined* ('computer alert' and 'paper movement'). All of the sounds created by the computers were classified as *abstract*, compared to the sounds created by inhabitants, which were either *emotional* or *everyday*. This highlights the perception of the computers despite their familiarity as being artificial, whereas the constant road traffic was referred to as an *everyday* sound event.

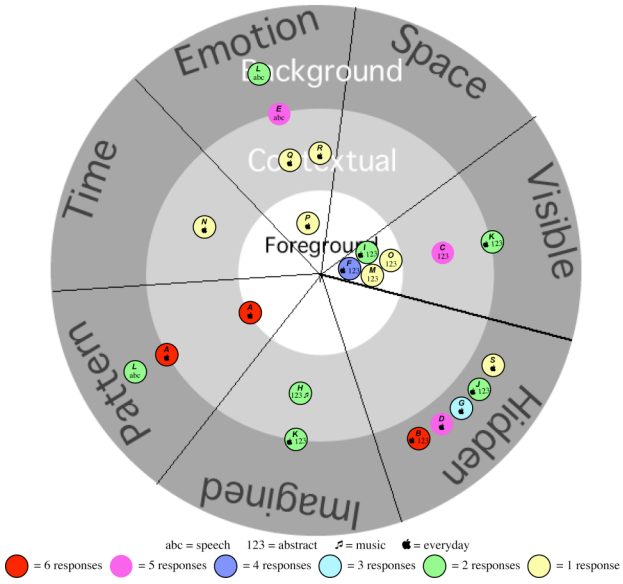

Figure 2. Visualization of a Typical Response to a Typical Computer lab.

A	Keyboard Typing	H	Computer alert	O	Purse velcro
B	Chair movement	I	Mouse movement	P	Laughter
C	Computer Fan	J	Air conditioning	Q	Coughing
D	Door opening & closing	K	Paper movement	R	Sighing
E	People talking	L	Whispering	S	Foot movement
F	Mouse clicking	M	Computer hard drive		
G	Traffic	N	Bag being zipped up		

Table 3: Key for Figure 2

With the photographic processing lab it was possible to create a map which represented the entire auditory environment, rather than just the specific areas studied in isolation. Responses from the individual regular inhabitants of the three main areas were combined, only a single sound event was shared, that of the telephone, which both FD and RW classified as *everyday*, *visible* and *foreground* (Figure 3 and Table 4). This approach reifies the soundscape from the perspective of an individual who inhabits the space most regularly.

The map clearly showed the varying control over the sound events within this environment. CR had most control over his environment, which resulted in 69% of the sounds being *foreground*, whereas FD has least control with only 30% being *foreground*, the inverse was true for *background* sounds with 7 and 35% respectively. *Visible* formed the largest group within the information category at 62%, followed by *patterns* (18%) and then *hidden* (12%). There were no instances of *imagined* as all of the sound events were easily identified, nor were there any references to *time*. *Everyday* formed the largest portion of sound type at 74%, illustrating the familiarity due to repeated exposure, compared to 19% *abstract*. There were surprisingly few instances of *speech* (5%) as most of the work was conducted in isolation, even, surprisingly, in the reception area.

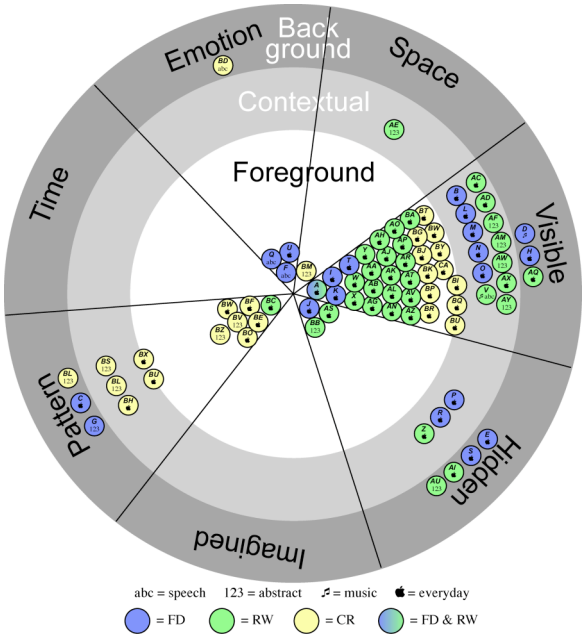

Figure 3. *Visualization of Combined Responses to the Photographic Processing lab.*

A Telephone ringing	AA Paper door closing	BA Slide Mount
B Doors opening to lab/computer room	AB RA4 paper processor Temperature beep	BB Hairdryer (drying prints)
C Coffee machine	AC Closing paper insert lid	BC Trays banging
D Radio	AD Paper bag rustling	BD Radio 4
E Customer's mobile phone	AE Footsteps on floor	BE Paper towel
F Customer	AF Trapped air in paper box	BF Mixing chemicals while measuring temperature
G Chair noises	AG Light switch	BG Ilford 2150 RC Processor fan
H Door alert	AH Bremson enlarger focusing switch	BH Processor warming up
I Telephone ringing	AI Bremson fan	BI Processor ready for next print
J Telephone hands free dialling	AJ Keypad buttons	BJ Brochure for timings
K Fax ringing	AK Keypad confirmation of settings	BK Tapping bottom of processing tank
L Modem dialling	AL Locking lens into position	BL Splash of fluid
M Keyboard tapping	AM Racking enlarger up and down	BM Throwing empty canisters into metal bucket
N Till beeping	AN Aperture selection	BN Handling plastic/paper bags
O Cash drawer	AO On/off switch (Bremson)	BO Air canister
P Switch receipt	AP Inserting film carrier	BP Easel adjustments and opening/closing
Q Conversation	AQ Enlarger easel	BQ Running water cleaning film
R Traffic	AR Air buster	BR Enlarger on/off switches
S Fan heater	AS Staff knocking to warn approach	BS Enlarger fan
T Printer	AT Revolving door	BT Timer confirmation
U Keys	AU Water pressure gauge (omnipro)	BU Timer countdown (seconds)
V Stereo	AV Print finished beep (omnipro)	BV Click of light switch
W Enlarger controller buttons (Buick)	AW Paper handling	BW Printing paper box opening/closing
X Exposure transport x 4	AX Door banging	BX Water run off from tanks
Y On/off switch (Buick)	AY Nitrogen Generator	BY Squeegee film
Z Fan (Buick)	AZ Trimming prints	BZ Cleaning squeegee
		CA Knocking excess water off reels

Table 4: Key for Figure 3

4. CONCLUSIONS

Macaulay and Crerar's mapping tool proved very easy to use, with the combination of categories covering every perceived sound event. Participants uniformly found it useful as a starting point for analyzing their auditory environment, but all of the individuals wished to contribute a greater amount of information than the classification requested.

A number of omissions became evident through the study, which could be split evenly between quantitative and qualitative. The first was quantity, a room could have twenty inhabitants, but speech was only detailed a single time. There was no indication if only one person was talking or everyone was talking. There was also no indication of how often these conversations took place, whether they were continuous or intermittent, as well as if they were concurrent with other sounds or were isolated events. In addition, location was omitted, in some cases the sound sources were equally spaced around the inhabitant, whereas on other occasions they were clustered. Directivity would have provided information about whom the sound was intended for, with speech it is common to direct the sound toward the intended recipient by facing them, whereas ubiquitous sounds, such as computer fans, tend to be omnidirectional. The last two quantifiable omissions were how loud and how high or low in terms of pitch or frequency sounds were. Whilst it would be extremely unrealistic to expect responses in terms of decibels and hertz, simple terms like loud/quiet and treble/bass might have provided consistent responses.

This method clearly showed the relative percentages of type, category and acoustical content, but was poor at representing the original sound event. Recording the time of each and every instance in real-time proved beyond the capabilities of the researcher, outwith unusual events. In order to produce an accurate record for data gathering, an auditory recording was required of the location and a highly skilled listener to decipher the recording; no software currently available would be capable of automating this task. It was also apparent that obvious sounds predominated, *foreground* and *contextual* sounds were notated first, whereas *background* sounds were notated last. This conforms with the way in which individuals interpret the world around them, but it does allow omissions due to perceptual masking, where a sound event is being established for notation and a quieter less intrusive sound is ignored, only to be notated if it is repeated after the predominant sounds have been detailed. This problem can be alleviated through recording the time period and notating the complete set of events from the recording.

In qualitative terms, the classification of *emotions* did not provide information about the mood which was being expressed, whether it was anger, frustration or relief. There was also no indication of what the type of interaction was, whether it was produced by air passing through an object or an impact, this was partially achieved through detailing the event, but not fully, as the requirement is to represent the sounds which were perceived rather than just a list of objects and actions. Bill Gaver's classification [7] of interacting materials by the focusing on the simple sonic events could easily be added, as the participants found it easy to recall the tonal qualities of the sound even an hour after the acoustic event. An understanding of the sound's information content would allow an insight as to how the sound event was interpreted by the participant, such as defined by Delage [8], whether it was an error alert or confirmatory sound or even unwanted. Further detail about a sound's perceived aesthetics would also prove useful in terms of communicating to designers the listener's preferences such as detailed in Gabrielsson and Sjorgen's method [9].

5. REFERENCES

[1] Macaulay, C. and A. Crerar, 'Observing' the Workplace Soundscape: Ethnography and Auditory Interface Design. ICAD 98, International Conference on Auditory Display, 1998.

[2] Ferrington, G., *Keep Your Ear-Lids Open.* The Journal of Visual Literacy, 1994. **14**(2): p. 51 - 61.

[3] Truax, B., Acoustic Communication. 2nd ed. 2001, Norwood: Ablex Publishing Corporation.

[4] Chion, M., Audio-Vision: Sound on Screen, ed. C. Gorbman. 1994, New York: Columbia University Press.

[5] Altman, R., The Material Heterogeneity of Recorded Sound, in Sound Theory/Sound Practice, R. Altman, Editor. 1992, Routledge: New York. p. 15-31.

[6] Amphoux, P., L'identite sonore des villes Europeennes. 1997, Cresson/IREC: Grenoble/Lausanne.

[7] Gaver, W.W., How Do We Hear in the World? Ecological Psychology, 1993. 5(4): p. 285 -313.

[8] Delage, B., On sound design, in Stockholm, Hey Listen!, H. Karlsson, Editor. 1998, The Royal Swedish Academy of Music: Stockholm. p. 67-73.

[9] Gabrielsson, A. and H. Sjogren, *Perceived sound quality of sound-reproducing systems.* Journal of the Acoustical Society of America, 1979. **65**(4): p. 1019-1033.