INVESTIGATING CONCURRENT AUDITORY ICON RECOGNITION

Eoin Brazil

Interaction Design Center
Department of Computer Science
and Information Systems
University of Limerick, Ireland
eoin.brazil@ul.ie

Mikael Fernström

Interaction Design Center
Department of Computer Science
and Information Systems
University of Limerick, Ireland
mikael.fernstrom@ul.ie

ABSTRACT

This paper an investigation of the identification of concurrently presented auditory icons[1]. The motivation for this work was to get a better understanding of the identification of an everyday sound scene. We collected a set of descriptions for a set of everyday sounds as classified by the participants, using their free text responses. Two different experiments were conducted. The first experiment used no sub-categorisation or classification information when choosing the auditory icons. The second experiment used object and action descriptors in the selection of auditory icons. Our hypotheses was that by ensuring auditory icons did not have the same object or action descriptors, the identification of auditory icons would improve. Both experiments used an onset-to-onset gap of 300 ms between auditory icons. The results show that when there was no overlap between the object and the action descriptors of the concurrent auditory icons, the identification of the auditory icons was significantly improved.

1. INTRODUCTION

Auditory scene analysis [1], auditory alarms [2] and Brewster's work on Earcons [3] have shown that truly artificial auditory displays with symbolic mappings can work successfully, however we lack the knowledge and empirical investigations of the everyday sound scene, in particular of auditory displays using everyday sounds. This led us to the two experiments presented in this paper. This work also links with our ongoing research [4] in mobile and ubiquitous computing.

Work in Auditory Display has often concentrated on situations where a single event or message, sometimes complex, is being conveyed but with increasingly more complex interoperating systems there is a growing need for the ability to convey multiple events or messages simultaneously. The lack of guidelines for designers wishing to convey multiple events or messages using auditory icons led us to our investigation of concurrently presented auditory icons. Researchers such as Papp [5] and McGookin [6] have investigated the concurrent audio presentation issue in a more formal manner. Their work has, however focused on Earcons. Our research uses methods based upon the work of previous researchers such as Papp and McGookin but is focused on auditory icons rather than Earcons.

Auditory icons are representations of everyday sounds (environmental sounds) designed to convey information from computer events[7, 8, 9]. These everyday sounds that have previously been learnt in our everyday activities have a meaning dependent on the context. Mynatt [10] discussed the recognition problem of how choosing the right sounds for the interface is an art with many hidden dangers and is dependent upon the skills of the designer. This can include sounds that are confusing when heard without context, or that create different imagery by simple rearrangement of the macro-temporal sequencing.

In an effort to reduce listener confusion, a subset of 40 sounds from a larger collection of 104 high-quality (44.1 Kilohertz 16-bit) everyday sounds (durations between 0.4 and 28.5) were used in our investigations [2]. The sounds had been classified by two categories, the object category of the sound and the action category of the sound. Another 32 sounds were selected from a collection of sound effect CDs [15] and classified in a similar manner to the previous 40 sounds. Fernström et al [16] have previously described this type of classification. An interesting result from Fernström's work was that actions of sounds are better identified than the objects involved in a sound. Research from Ballas and Howard [17] found that semantic context in sound interpretation is an important factor and that auditory perception is directed towards awareness of the sources of sounds i.e. the events producing sounds. They also stated that the function of auditory perception is to recognise events rather than processing acoustic patterns. As events consist of actions, objects and context, we have concentrated on these properties for our research and this paper is focused on the action and object properties of concurrent auditory icons. In order to further research two experiments were conducted where the identification of concurrently presented auditory icons sounds was investigated.

2. AUDITORY ICONS

Auditory icons were defined by Gaver [18] as "*everyday sounds mapped to computer events by analogy with everyday sound producing events*" and demonstrated by Gaver in systems such as the SonicFinder. Historically auditory icons have been used to repre-

[1]Concurrent auditory icons is the playing of several auditory icons together and simultaneously, to build more complex and compound auditory icons. This is inspired by how real sounds work in the world. For example, at the time I'm writing this, I happen to have a quiet news broadcast on the radio in the background; cars are passing on the road outside; the birds are singing outside my house and the wind is blowing through an open window.

[2]The sounds used were complex, dynamic and informational events with different temporal patterns [11, 12]. These sounds were edited to a duration which allowed for the "sound event" or "sound object" [13] to appear to occur naturally [14].

sent distinct events or items, where the relationship is a mapping from an event in the everyday sound to a computer event. The identification of an auditory icon directly relates to the analogy used between the everyday world and how this is mapped to a system. A bad mapping can result in a mismatch between what feedback the system provides and the feedback the user needs in the particular interaction. The metaphorical association of an auditory icon to the interacting object is built on our everyday listening skills. Gaver has described and investigated auditory icon comprehension in an ecological sense [7, 8]. Auditory icons can be parameterised as discussed by Gaver [19] to reflect the relevant dimensions of the interacting object. The identification of an auditory icon is related to the interpretation of its learnt meaning and this identification becomes more difficult when concurrently presenting auditory icons as the more icons presented the greater the possibility of confusing the interpretation. Modification may be made to help improve auditory icon identification; but the modification is constrained due to the metaphorical association that must be preserved to maintain the mapping between the events and the sounds.

3. AUDITORY SCENE ANALYSIS

Auditory scene analysis (ASA) is the model of perceptual grouping of auditory stimuli into auditory streams. Auditory streams are a way of putting sensory information together. The world is a rich sensory environment where people are exposed to hundreds of different sounds simultaneously and are still able to segregate and sort out the important parts of this environment or "*auditory scene*". A great variety of research relating to perceptual grouping of auditory stimuli into streams has been done, and described, by Bregman [1]. The perceptual grouping can be described by heuristics based upon the *Gestalt* principles of psychology [20]. The idea of temporal proximity has been used in this research as prior work by Darwin and Ciocca [21] and by McGookin [22] has shown a 300 ms onset-to-onset gap can improve steam segregation and identification. Factors that affect ASA seem to be dynamic and subject to change so there is no rigid rule set available that will take an auditory scene and determine how it will be perceived but there is a body of work on computational auditory scene analysis [23, 24, 25] attempting to address these issues. It is however inconclusive and does not offer any solutions that could be generalised to auditory icons at this point in time.

4. AUDITORY ICONS AND ASA

The issues discussed in the previous section show how the playing of concurrent auditory icons might have undesirable and unforeseen consequences but that can, to a degree, be explained by ASA. These problems are important where common everyday sounds are used as elements contained within a particular sound as they may fuse with a different sound and create a different representation for the auditory scene. In our experiments we aimed to determine how well identified the sounds are and what people describe when listening to several concurrent auditory icons.

5. CAUSAL UNCERTAINTY

Causal uncertainty is a measure of how a single sound may be produced by different causes. It was originally developed by Ballas et al. [26] and is derived from the uncertainty metric *H*, inspired by information theory. A *H* value for a sound is calculated based on the number and frequency of the different identifications of the sound. Ballas and colleagues found these values were stable for different examples of a particular sound [27], correlated with identification time for a sound [26], with subjective ratings of uncertainty [17] and with rating of the identifiability of a sound [28] These values were also found to be consistent across different listener groups with secondary school, college students and older listeners being examined [27, 29]. Ballas's research found that the *H* values had a greater correlation with identification time than with either percentage correct or the number of alternative identifications of the sound. This suggests that for everyday sound identification there is some type of parallel processing of alternatives or other information. The aim of this research continues from our earlier research [16] where only a single auditory icon was presented, while in this approach the concurrent presentation of three or more auditory icons was used. As part of this research, a listening test method was used, as used by several other researchers e.g. [30, 31, 32]. Our experiments also sought collect and analyse a set of descriptions for the auditory icons used. In analysing these descriptions Ballas method of causal uncertainty [30] was used. Ballas et al found that the identification time for everyday nonspeech sounds was a function of the logarithm of the number of alternative interpretations of a sound [26]. H_{CU} is a measure of causal uncertainty for sound i, where p_{ij} is the proportion of all responses for sound i sorted into event categories j and n is the number of categories for responses to sound i as shown below in equation 1.

$$H_{CU} = \sum_{i}^{n} p_{ij} \log_2 p_{ij}$$

Equation 1: Causal Uncertainty

6. EMBODIED COGNITION

Our research has been inspired by previous work by Gibson on *affordances* [33], by Lakoff and Johnson [34] on *embodied realism*, by Varela et al and their definitions of *embodied* and *action* [35], by Rosch and her concepts of two levels of category-specific organisation of concepts [36, 37, 38]. These concepts and recent research in neurophysiology focusing on the role of action in cognitive processes with specific regard to the perceptual recognition of objects, actions and their conceptual categorisation [39] have influenced our thinking. Garbarini discusses a viewpoint combining these concepts where the mind is seen "*as as a biological system rooted in bodily experience and interconnected with bodily action and interaction with other individuals. From this perspective, action and representation are no longer interpreted in terms of the classic physical-mental state dichotomy, but are closely interconnected. Acting in the world, interacting with objects and individuals in it, representing the world, perceiving it, categorizing it, and understanding its significance are perhaps simply different levels of the same relational link that exists between organisms and the local environments in which they operate, think, and live*" P105 [39]. Garbarini further refines the idea of *mental representation* as being "*intrinsically linked to the sphere of action and is expressible in the same terms that control it ...(the) mental representation in which the experience is "constructed" on the base of categories, which are no longer theoretical, but pragmatic, deriving from the dynamic interaction of the organism with its adaptive environment*" P106 [39]. The object-action dichotomy we use in our experiments has been influenced and informed by these researchers. Our studies aim to explore the object-action dichotomy derived from the participants responses.

7. EXPERIMENT RATIONALE

The motivation for these two studies was to gain a greater insight into the perception of everyday sound scenes. Previous work by others on Earcons or sinusoidal tones have proven to be valuable for design of auditory display, but the research into everyday sounds and use in complex auditory displays is lacking. The two experiments focus on the question of the identification of the everyday sounds within a complex scene where several everyday sounds are occurring concurrently. The concurrent presentation of auditory icons could cause them to interfere with one another and can be difficult to distinguish individually but as far as we have searched there are no studies that have investigated the extent to which identification of auditory icons is impaired by their concurrent presentation. McGookin et al [6] have conducted similar studies for Earcons and Brungart et al [40, 41] have investigated concurrent speech presentation. Both of these studies have shown that where the number of concurrent audio items is increased, the total proportion of identified audio is reduced. The auditory icons used in our experiments were monophonic sounds as we tried to reduce the number of variables in the problem domain.

The work by McGookin et al [22] on Earcons suggested using at least a 300 ms onset-to-onset gap as they found using a onset gap between the starts of concurrently presented Earcons improved identifiability. Our experiments tried to explore the impact on auditory icon identification by increasing the number of concurrently presented auditory icons. The two experiments we present in this paper deal with two sets of everyday sounds and their identification. The criteria of how we selected these sounds is important to distinguish, the first experiment used no prior classification of the sounds and the sounds were selected on a random basis for inclusion in the experiment. The second experiment used a prior classification of each sound in the collection with categorisation about its actions and the objects to exclude sounds being selected for a experiment condition where a sound already existed in that condition with either the same action or object classifications. Both experiments collected the set of descriptions for the sounds provided by the participants using their free text responses.

8. DESCRIPTION OF CONCURRENT AUDITORY ICONS WITH POSSIBLE OBJECT AND ACTION OVERLAPPING

The first experiment investigated concurrent presentation of auditory icons where the auditory icons within the conditions were selected randomly with no use of object or action descriptors. Example sets of auditory icons could, for example, contain several banging sounds or several sounds from the same type of object such as glass objects or motor vehicles.

8.1. Training

The focus of this training phase was to familiarise the users with the concurrent presentation of auditory icons. A training interface allowed participants to stop, start and loop up to seven auditory icons whose descriptions were provided on-screen. The stimuli used in the training phase were not used in the later tests. The participants spent 10 minutes using the interface after a short introduction on its operations. Users had headphones to listen to the sounds (in mono) while interacting with the system.

8.2. Participants

11 participants were recruited from the postgraduates at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the experiment from all participants.

8.3. Stimuli

The 72 everyday sounds as discussed in section 1 where used as the stimuli for the experiment. These sounds were not classified according to any object or action properties. In order to prevent the merging of streams if the stimuli were all presented simultaneously, a 300 ms onset delay was used between all auditory icons to stagger the onsets of the sounds.

8.4. Procedure

The participants listened to the recorded sounds (mono) in random order using headphones, responding in free-text format to what each sound was, using the interface shown in Figure 1. The conditions varied from three to six sounds being concurrently presented. The interface below used ten description boxes, as the complex scenes with multiple sounds may lead to richer descriptions than just one descriptor per recorded sound. Using a within-subjects design, the stimuli were presented in fixed order within the conditions and the task order was counter-balanced for the conditions (three, four, five or six concurrent auditory icons).

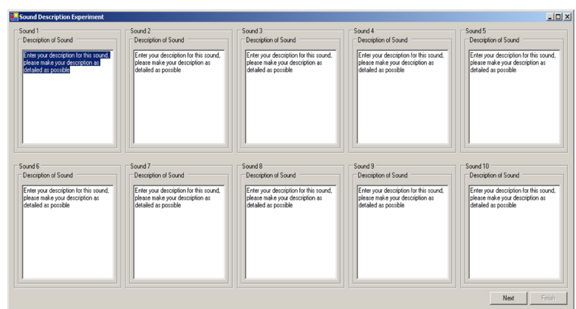

Figure 1: *A screen shot of the dialogue used by participants to fill in the descriptions of parallel sounds presented.*

8.5. Results

Three methods, application data logging, participant observation, and post experiment questionnaires gathered the results for this experiment. The application data logging captured the sound description entered by the participant as well as timing the duration each participant spent per condition. To determine the number of auditory icons correctly identified by participants, the following method was used. For each set of (three, four, five or six) concurrently presented auditory icons, the set of auditory icons presented (SAIP) and the set of participant responses (SPR) to those auditory icons were compared. If the description of an auditory icon in the SPR matched an auditory icon in the SAIP, and if that auditory icon has not already been identified and matched with a previous description, the number of correctly identified auditory icon was increased by one, and the auditory icon description was marked as allocated. An example mapping between the SAIP and SPR, from the four auditory icon condition is shown in figure 2. Due to the differing numbers of auditory icons in each condition a direct numerical comparison between the auditory icons and the correctly identified auditory icons would not be possible. An alternative comparison can be made using the average number of auditory icons identified per participant and converting this average into a percentage of the number of auditory icons that were concurrently presented. The average proportion of correctly identified auditory icons across all participants is presented graphically in Figure 3.

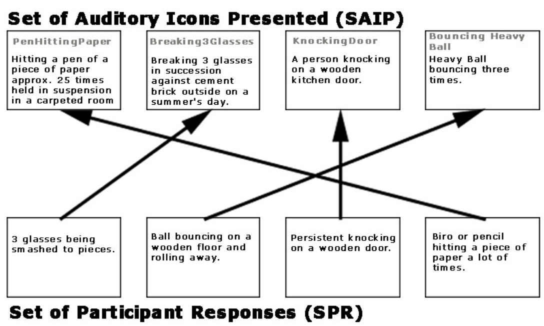

Figure 2: *An example from the four auditory icon condition of how the set of participant responses (SPR) were mapped to the set of auditory icons presented (SAIP) in order to determine the number of correctly identified auditory icons.*

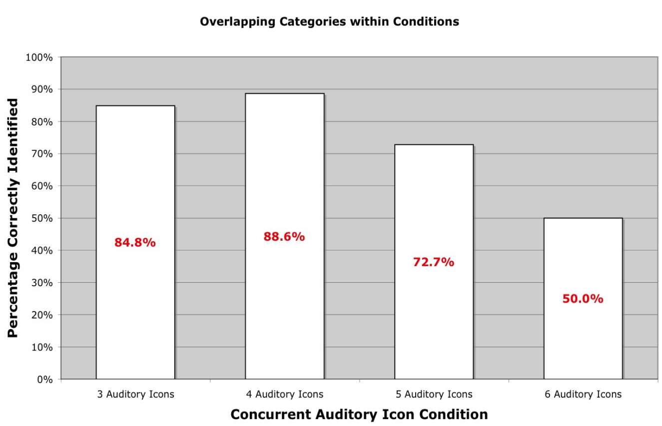

Figure 3: *Graph showing the average proportion of auditory icons identified for the three, four, five and six auditory icon conditions.*

To determine if any of the differences shown in Figure 3 were statistically significant, a one-way analysis of variance (ANOVA) test was carried out for the percentage of correctly identified auditory icons and was found to be significant (F (3,40) = 10.091, p < 0.0000044).). A post hoc Tukey Honesty Significance Difference (HSD) [42] test showed significant differences between the 6 concurrent auditory icon condition and both the three and four concurrent auditory icon conditions. These results show that auditory icons identification grows worse as the number presented is increased ($p < 0.05$).

A post experiment questionnaire was used to collect data about the participants gender, musical ability and qualitative differences in the various tasks between participants. The participant debriefing session raised several interesting points as indicated by comments from the participants like "*I found it difficult with all the glasses breaking and had to close my eyes to solely concentrate on all the sounds, it was difficult and I needed the additional plays of the mix*" and "*It was the sound of twigs breaking but could have been crackling in the fire*".

8.6. Conclusions of First Study

The results from this experiment show that trying to identify concurrently presented auditory icons is relatively easy in conditions with 3, 4, 5 or 6 concurrent auditory icons, with on average 74.05% of auditory icons being correctly identified. The results from the three auditory icon condition showed around 34.85% increase in the accuracy of auditory icon identification as opposed to the six auditory icon condition, meaning that the three auditory icon condition was better identified than the six auditory icon condition. This result supports the hypothesis that varying the number of concurrently presented auditory icons has a significant effect on a participants ability to identify those auditory icons. Performance levels decreasing in accuracy due to increasing complexity is consistent with previous research from human factors [43]. Examining the auditory icon conditions we can see the accuracy rates decrease from 84.85% and 88.64% to 72.72% to 50% respectively. This shows a difference of 34.85% in identification between the three and six auditory icon condition. Taking the auditory icon conditions we can say that the greater the number of auditory icons concurrently presented to participants, the lower the proportion of those auditory icons that can be successfully identified. In the six auditory icon condition the number of correctly identified auditory icons is severely reduced when concurrently presented auditory icons with the same object and / or action classification are used. The identification of the individual auditory icon objecthood or actionhood in the six sound condition is very robust with 0.663 for the object H_{CU} and 0.498 for the action H_{CU} as the compound results of the six concurrently presented auditory icons. These figures are deceptively low as participants heard on average only three of the six sounds present, which could be due to perceptual merging of the auditory streams. The high identification rates for the conditions presenting low numbers of auditory icons with the associated low causal uncertainty measures suggest that from 3 to 5 well chosen auditory icons with overlapping action and / or object categories would be robust in real world tasks, but given the results of the 6 concurrent auditory icons condition and the issues of streaming/ merging of similar sounds, further investigation is required.

9. DESCRIPTION OF CONCURRENT AUDITORY ICONS WITH NO OVERLAP BY ACTION OR OBJECT CATEGORIES

The second experiment investigated the concurrent presentation of auditory icons were the auditory icons within the conditions were selected with regard to their object or action descriptors. An example of this selection approach would be to prevent the choice of several banging sounds in a condition; it would equally prevent several sounds from the same type of object such as a glass object being selected for the particular condition.

9.1. Training

The focus of the training phase was similar to the previous experiment. The stimuli used in the training phase were not used in the later tests. The participants spent 10 minutes using the interface after a short introduction. Users had headphones to listen to the sounds (in mono) while interacting with the system.

9.2. Participants

11 participants were recruited from the postgraduates at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the experiment from all participants.

9.3. Stimuli

72 everyday sounds, as previously mentioned in section 1, with durations between 0.4 and 28.5 seconds were used in this experiment. In contrast to the previous experiment, sounds had been pre-classified according to objects and actions and selected to prevent overlapping of those properties within a condition. As with the previous experiment, the stimuli were presented using a 300 ms onset delay between each auditory icon in the condition to stagger the onsets of the sounds.

9.4. Procedure

The experiment used the same interface and sound collection as described for the previous experiment. As with the previous experiment, a within-subjects design was used where the stimuli were presented in fixed order within the conditions and the task order was counter-balanced for the conditions (three, four, five or six concurrent auditory icons).

9.5. Results

As with the previous study application data logging, participant observation, and post experiment questionnaires were used to gather the experimental results. For each set of (three, four, five or six) concurrently presented auditory icons, the set of auditory icons presented (SAIP) and the set of participant responses to those auditory icons (SPR) were compared. The average proportion of correctly identified auditory icons across all participants is presented graphically in Figure 4. To determine if any of the differences shown in Figure 4 were statistically significant, a one-way analysis of variance (ANOVA) tests was carried out for the percentage of correctly identified auditory icons and was found to be significant ($F(3, 40) = 4.1892$, $p < 0.05$). A post hoc Tukey Honesty Significance Difference (HSD) [42] test showed significant differences between the conditions with 4 and 5 concurrent auditory icons as well as the condition with 3 and 4 concurrent auditory icons. These differences may be due to the sound of Running up Stairs being masked in the 3 concurrent auditory icon condition and the Chain Rattle being masked in the 5 concurrent auditory icon condition which we will discuss further in the comparison of the experiments. There were no other significant differences between conditions, which indicate the high levels of similarity between the levels of identifications but also show the need for further investigations with larger numbers of auditory icons.

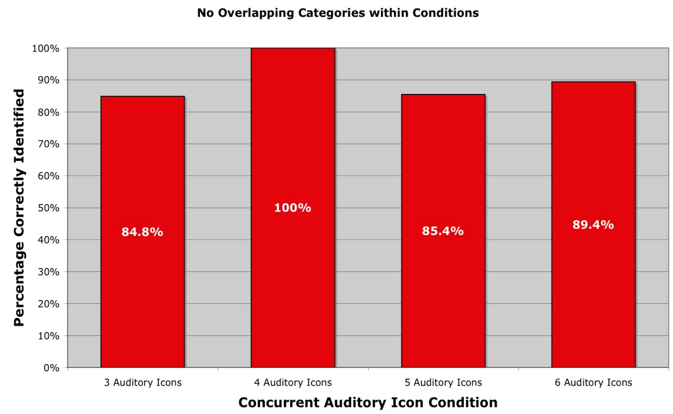

Figure 4: *Graph showing the average proportion of auditory icons identified for the three, four, five and six auditory icon conditions.*

In a similar manner and with similar topics to the previous study a post experiment questionnaire was used to collect data from the participants. The debriefing session raised several interesting points as indicated by comments from the participants like "*When you have the water going down a plug hole, so there was a sound like hitting a tap that started it and stopped it but I wasnt sure if it was part of the same action*" and "*Some sounds I heard once and even when concentrating on other sounds would drown out the other sounds I was trying to concentrate on, like the alarm clock or the stream bubbling sounds*".

9.6. Conclusions of Second Study

The results from this study shows that trying to identify concurrently presented auditory icons is relatively easy in conditions with 3, 4, 5 or 6 concurrent auditory icons, with on average 89.92% of auditory icons being correctly identified. The results from the three auditory icon condition showed around 4.55% decrease in accuracy of the auditory icon identification as opposed to the six auditory icon condition, that meaning that the six auditory icon condition was only slightly better identified than the three auditory icon condition. This result supports our hypothesis that preventing an overlap in the object and action properties of auditory icons in situations with concurrently presented auditory icons, can have an significant effect on a participants ability to identify those auditory icons. Examining the three, four, five, and six auditory icon conditions we can see the accuracy rates vary from 84.85% to 100% to 85.45% to 89.39% respectively. This shows a 4.55% increase in accuracy between the three and six auditory icon condition. Taking the three, four, five, and six auditory icon conditions we find even with a greater number of auditory icons concurrently presented to participants, the proportion of those auditory icons that can be successfully identified is not significantly different when ensuring there is no overlap in auditory icon object and action properties. Even in the six auditory icon condition the number of correctly identified auditory icons is only mildly reduced when the number of concurrently presented auditory icons is increased. The identification of the auditory icons objecthood or actionhood is very robust with 0.357 for the object H_{CU} and 0.383 for action H_{CU}

as the compound results for the case of six concurrently presented auditory icons. The high identification rates for the limited numbers of auditory icons used for this experiment and low compound H_{CU} suggest that from 3 to 6 well chosen auditory icons would be robust in real world tasks, but requires further investigation.

10. COMPARISON OF THE EXPERIMENTS

Examining the data from both experiments we can find an interesting trend when the results from experiment one (labeled as "No Categorisation Use") and the results from experiment two (labeled as "No Overlapping by Category") are shown below in Figure 5. The possibility of masking was found in two of the conditions in the second experiment after analysis of the results. Even considering the possible presence of masking, the overall results from experiment two still show identification rates well above 80%. The poor result of one sound "Running up Stairs" in experiment two in the three concurrent auditory icon condition is possibly due to it being masked by the "Bouncing Heavy Ball" sound in the condition. The sonograms[3] in Figures 6 and 7 show this to be the case. The confusion with the "Running up Stairs" is further explained taking the analysis of its action and object H_{CU} 's of 1.022 and 0.831 respectively, and also the low level of response to it by participants where only 6 of the 11 participants described the sound. A similar masking would account for the confusion with the "Metal Chain" sound in experiment two in the five concurrent auditory icon condition. Examining this graph we can see that until the six concurrent auditory icons condition that there is no significant difference between percentages of identification of the conditions. The six concurrent auditory icons condition shows a difference of 39.39% in accuracy between the two experiments. The sounds in experiment one were free to overlap in categories and the lower result of 50% could show that from six concurrent sounds onwards this overlap may cause difficulties in the identification of the sounds, while those sounds in experiment two that maintained a semantic difference by ensuring no category overlap had a better level of identification than even the first overlapping condition in experiment one with three sounds (89.39% versus 84.85%). This result shows that six well-chosen and semantically distinct concurrent auditory icons can be robust performers in real world tasks. Our future research will continue to investigate this trend using larger numbers of auditory icons to investigate the limits of identification for concurrent auditory icons.

Putting these results in context with work by McGookin et al [22] on Earcon identification is difficult given the Earcon research used both concurrent and compound Earcons with three data parameters encoded per Earcon. Examining the results for four concurrent Earcons and for four concurrent auditory icons the difference in correct identification is 42.5% for Earcons with a staggered onset, 100% for auditory icons with no overlapping action or object classifications and 88.6% for auditory icons with overlapping action or object classifications. In McGookin's study the ride types, intensities and costs were better identified with values of 75%, 70% and 65% respectively. The compound H_{CU} for the four auditory icons condition in the no overlapping experiment was 0 (action) and 0.17 (object). The compound H_{CU} for the four auditory icons condition in the overlapping experiment was 0.29 (action) and 0.34 (object). These compound H_{CU} measures

[3]The sonograms used an FFT size of 512 with 50 % overlap using a Hanning window, from 20hz to 4.7KHz.

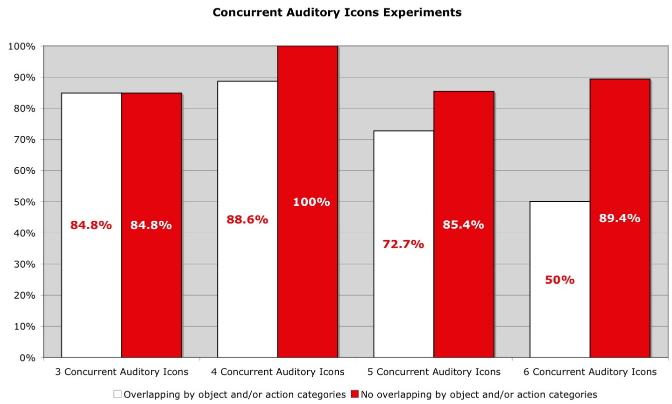

Figure 5: *The comparison of the results of the two investigations.*

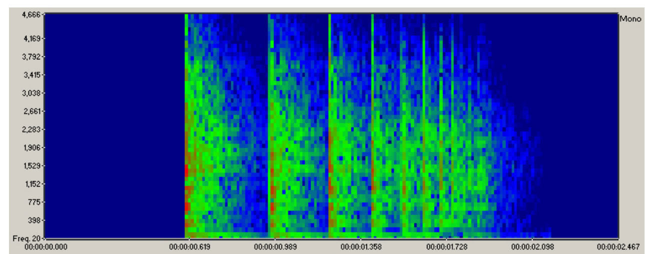

Figure 6: *Sonogram of the "Bouncing Heavy Ball" sound in experiment 2 (No Overlap By Action Or Object Categories) three auditory icon condition.*

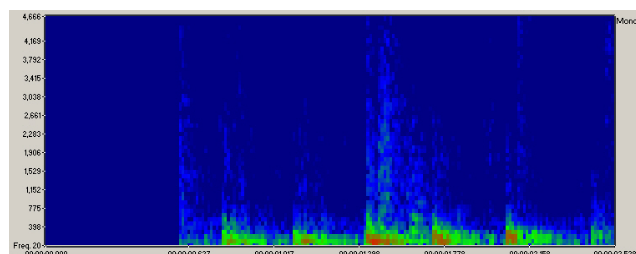

Figure 7: *Sonogram of the "Running Upstairs" sound in experiment 2 (No Overlap By Action Or Object Categories) three auditory icon condition.*

show a clarity with regard to identification by participants when identifying distinct or non overlapping classifications of the auditory icons. The compound Earcons use three bits of information encoded per Earcon in a particular directed task for browsing the rides in an amusement park, whereas this study used a free text response method as the task and as a result comparisons cannot be made directly between the two studies.

11. DISCUSSION

Listening test results can provide guidance in our understanding of how users interpret combinations of auditory icons and in creating designs based on the results from listening tests with possible metaphors being extracted from the participants descriptions of everyday sounds. More studies are needed on what people hear when they listen to everyday sounds to increase our understanding of the perceptual and cognitive processes involved. This research concentrated on the study of the effects of combinations of different concurrent auditory icons but more research is necessary in this area. All components in Auditory Displays also have aesthetic properties. In designing auditory displays it is possible to design sonifications and use auditory icons that are psychoacoustically correct and also quite efficient but that are unpleasant to listen to. Sound Design [44], Foley artists, theories from acousmatic music should all be drawn upon to further influence the work in auditory displays.

In this paper, everyday sounds were investigated to collect a set of descriptions for the concurrent auditory icons as classified by the participants in their free text responses. The results indicated that a factor that can increase the identifiability of auditory icons was to select auditory icons that had been classified using action and object descriptors and when choosing the auditory icons for concurrent presentation that these descriptors did not overlap. If there was no overlap between the objects and action descriptors of the concurrent auditory icons chosen, there can be a significant increase in the identifiability of the concurrently presented auditory icons. It should be noted a 300 ms onset-to-onset gap was used between each auditory icon in all of the experiments.

12. ONLINE INTERNET RESOURCES

The applications and non-commerical sounds used for this study are available for download at http://richie.idc.ul.ie/eoin/icad06/ along with several of the participant log files and the Microsoft Excel workbooks and scripts used for these experiments. Several of the sounds where taken from commercial CD's and information regarding these sounds is included but due to copyright issues, the actual sound files are not available. Composite sound files of each of the auditory icon conditions from both experiments are also available for download.

13. ACKNOWLEDGMENTS

Enterprise Ireland supported parts of this work under the Metamusic project and the Multimodal Browsing project. Science Foundation Ireland also supported aspects of this work under the Shared Worlds project. We would like to express our thanks to all the postgraduate students at the University of Limerick who participated in this experiment. We would also like to thank our colleagues, Annette Aboulafia and Liam Bannon for their comments on earlier versions of this paper.

14. REFERENCES

[1] A.S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, M.I.T. Press, Cambridge, MA, USA, 1990.

[2] N.A. Stanton and J. Edworthy, Eds., *Human Factors in Auditory Warnings*, Ashgate, Aldershot, Hants, UK, 1999.

[3] S.A. Brewster, *The Human Computer Interaction Handbook*, chapter Non-speech auditory output, pp. 220–239, Lawrence Erlbaum Associates, 2002.

[4] L. Bannon, *Future Interaction Design*, chapter A human-centred perspective on interaction design, Springer-Verlag, 2005.

[5] A.L. Papp III, *Presentation of Dynamically Overlapping Auditory Messages in User Interfaces*, Phd thesis, University of California, 1997.

[6] D. McGookin, *Understanding and Improving the Identification of Concurrently Presented Earcons*, Ph.D. thesis, University of Glasgow, 2004.

[7] W. W. Gaver, "What in the world do we hear? an ecological approach to auditory source perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[8] W. W. Gaver, "How do we hear in the world? explorations of ecological acoustics," *Ecological Psychology*, vol. 5, no. 4, pp. 285 – 313, 1993.

[9] W. W. Gaver, "Using and creating auditory icons," in *Auditory Display: Sonification, Audification and Auditory interfaces*, G. Kramer, Ed., pp. 417–446. Addison-Wesley Publishing Company, Reading, MA, USA, 1994.

[10] E.D. Mynatt, "Designing with auditory icons," in *Second International Conference on Auditory Display (ICAD '94)*, G. Kramer and S. Smith, Eds., Santa Fe, New Mexico, 1994, pp. 109–119, Santa Fe Institute.

[11] J. J. Jenkins, "Acoustic information for objects, places, and events," in *Persistence and Change - Proceedings of the First International Conference on Event Perception*, W. H. Jr. Warren and R. E. Shaw, Eds., Hillsdale, 1985, Erlbaum.

[12] S. McAdams, *Thinking in sound: The cognitive psychology of human audition*, chapter Recognition of sound sources and events, pp. 146–198, Oxford, Clarendon, 1993.

[13] R. F. Port, F. Cummins, and J. D. McAuley, *Mind as Motion: Exploration in the dynamics of cognition*, chapter Naive time, temporal patterns, and human audition, M.I.T. Press, Cambridge, 1995.

[14] F. L. Wightman and R. Jenison, *Perception of Space and Motion*, chapter Auditory spatial layout, Academic Press, San Diego, 1995.

[15] Premiere Edition, "The hollywood edge [sound recording] : sound effects library," 1990.

[16] J.M. Fernström, E. Brazil, and L. Bannon, "Hci design and interactive sonification for fingers and ears," *IEEE Multimedia*, vol. 12, no. 2, pp. 36–44, 2005.

[17] J. A. Ballas and J. H. Howard, "Interpreting the language of environmental sounds," *Environ. Behav.*, vol. 1, no. 9(1), pp. 91–114, 1987.

[18] W.W. Gaver, "Auditory interfaces," in *Handbook of Human-Computer Interaction*, M.G. Helander, T.K. Landauer, and P. Prabhu, Eds. Elsevier Science, Amsterdam, The Netherlands, 2 edition, 1997.

[19] W. W. Gaver, "Synthesizing auditory icons," in *Interchi'93*. 1993, pp. 228–235, ACM Press.

[20] K. Koffka, *Principles of Gestalt Psychology*, Harcourt-Brace, New York, 1935.

[21] C.J. Darwin and V. Ciocca, "Grouping in pitch perception: Effects of onset asynchrony and ear of presentation of mistuned component," *J. Acoustic Society of America*, vol. 91, no. 1, pp. 2281–3390, 1992.

[22] D. McGookin and S. A. Brewster, "Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition," *ACM Transactions on Applied Perceptions*, vol. 1, no. 2, pp. 130–155, 2004.

[23] D. Ellis, *Prediction-driven computational auditory scene analysis*, Phd thesis, Massachusetts Institute of Technology, 1996.

[24] A. J. W. van der Kouwe, D. Wang, and G.J. Brown, "A comparison of auditory and blind separation techniques for speech segregation," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 3, pp. 189–195, 2001.

[25] S.N. Wrigley and G.J. Brown, "A computational model of auditory selective attention," *IEEE Transactions on Neural Networks*, vol. 15, no. 5, pp. 1151–1163, 2004.

[26] J. A. Ballas, M.J. Sliwinsky, and J.P. Harding, "Uncertainty and response time in identifying non-speech sounds," *Acoustical Society of America*, vol. 79, 1986.

[27] J. A. Ballas, K. N. Dick, and M. R. Groshek, "Failure to identify "identifiable" sounds," in *Proceedings of the 31st Annual Meeting of the Human Factors Society*, Santa Monica, CA, 1987, Human Factors Society, pp. 144–146.

[28] J. A. Ballas, "Acoustic and perceptual-cognitive factors in the identification of 41 environmental sounds," Tech Report ONR-89-2, Center for Behavioral and Cognitive Studies, George Mason University, Fairfax, VA, 1989.

[29] J. A. Ballas and M. E. Barnes, "Everyday sound perception and aging," in *Proceedings of the 32nd Annual Meeting of the Human Factors Society*, Santa Monica, CA, 1988, Human Factors Society, pp. 194–197.

[30] J. A. Ballas, "Common factors in the identification of an assortment of brief everyday sounds," *J. of Experimental Psychology*, vol. 19, no. 2, pp. 250–267, 1993.

[31] W. W. Gaver, *Everyday Listening and Auditory Icons*, Ph.d., Cognitive Science and Psychology, San Diego, University of California, 1988.

[32] N.J. Vanderveer, *Ecological Acoustics: Human Perception of Environmental Sounds*, Phd, University Of Cornell, 1979.

[33] J.J. Gibson, *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates Inc. Publishers, Hillsdale, NJ, USA, 10 (1986) edition, 1979.

[34] G. Lakhoff and M. Johnson, *Philosophy in the flesh: The embodied mind and its challenge to western thought*, Basic Books, New York, 1999.

[35] F. J. Varela, E. Thompson, and E. Rosch, *The embodied mind: Cognitive science and human experience*, MIT Press, Boston, 1991.

[36] E. Rosch, *Study in cross-cultural psychology*, chapter Human categorisation, pp. 1–49, Academic Press, London, 1977.

[37] E. Rosch, "Family resemblance: Studies in the internal structure of categories," *Cognitive Psychology*, , no. 7, pp. 573–605, 1975.

[38] E. Rosch, *Cognition and Categorisation*, chapter Principles of Categorisation, pp. 27–48, Lawrence Erlbaum Associates, Hillsdale, NY, 1978.

[39] F. Garbarini and M. Adenzato, "At the root of embodied cognition: Cognitive science meets neurophysiology," *Brain and Cognition*, , no. 56, pp. 100–106, 2004.

[40] D.S. Brungart, M.A. Ericson, and B.D. Simpson, "Design considerations for improving the effectiveness of multitalker speech displays," in *International Conference on Auditory Display ICAD-02*, Kyoto, Japan, 2002, pp. 424–430.

[41] D.S. Brungart and B.D. Simpson, "The effects of spatial separation in distance on the information and energetic masking of a nearby speech signal," *J. Acoustic Society of America*, vol. 112, no. 2, pp. 664–676, 2002.

[42] R.B. McCall, *Fundamental Statistics for Behavioral Sciences*, Brooks/Cole Publishing, Pacific Grove, 7th edition, 1998.

[43] C.D. Wickens and J.G. Hollands, *Engineering Psychology and Human Performance*, Prentice Hall, New Jersey, USA, third edition edition, 2000.

[44] D. Sonnenschein, *Sound Design: The Expressive Power of Music, Voice, and Sound Effects in Cinema*, Michael Wiese Productions, Studio City, CA, USA, 1st edition, 2001.