Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, July 6-9, 2003

PERCEPTIVE STUDY AND RECOMMENDATION FOR SONIFICATION CATEGORIES

Patrick Susini, Piotr Gaudibert, Emmanuel Deruty, Louis Dandrel

Institut de Recherche et Coordination Acousique/Musique (IRCAM-CNRS)
1 place Igor Stravinsky, F-75004 Paris, France
susini@ircam.fr

ABSTRACT

In the field of audio signaletics, most sound designers have their own recipes to make samples that convey a certain meaning, which we could call auditory function. The aim of the present article is to compare the perceptive representation and the functional representation with the usual sound categories designed to fulfill specific actions of user's interface. The article finally proposes recommendations for the designers according to perceptive results.

1. INTRODUCTION

This study, part of a project to develop techniques for the creation of on-line radio programs based on an individual's thematic choice, focuses on the ergonomics of sonification in the context of web radio. The sonifications, superimposed on the audio stream, reflect the functions of this medium: hyperlinks, alerts, "assistants", etc. … on the one hand, additional information about the program being listened to or about a similar program proposed in a real time channel, and on the other, unrelated events, such as the arrival of mail, a message from a background process, or a reminder of an appointment. The goal is to inform the listener simultaneously about supplementary information :
- specific to the content of the program : "sonified hyperlink" [1] [2], "RadioLink"
- not specific to the content of the program but specific to the user's interface : "mail in / out", "alarm", "feedback yes / no".
The aim of this work was to study how well a given sound fulfills its function according to the ususal categories and to compare the functional representation with the perceptive representation. From a practical point of view, it's interesting to check whether the auditory function of a sound is indeed understood by the actual user, and if not, what the user exactly understand from the samples he's listening to. This knowledge could help to create unambiguous sounds used in audio assistive interfaces for visually impaired people or enforced vigilance performance in an audio-visual context [3] [4] [5].
The final purpose of the present article is to propose acoustical recommendations for sounds used to fulfill different functions.
The sounds used in this study were taken from common existing software, including OS and games and were grouped in six defined categories (§2.1). Two categorization experiments were performed with the same sound samples and two different groups of subjects (§2.2). For each experiment, the instructions given to the subjects were different in order to study respectively the perceptive representation and the functional representation (§3.1 & §3.2). In the last step, subjects were asked to select unpleasant or offensive sounds from each the six categories (§3.3). Finally, some recommandations are proposed (§4).

2. EXPERIMENTAL PROTOCOL

2.1. Material

The first step in the study was to survey sonificated programs such as instant messengers, mail clients, operating systems, and games. A few sonificated web pages were also examined : a description and sound samples can be found at this address : http://www.ircam.fr/equipes/design/activites/sonification/VI_Studies/Study01/index.html
Each program is described using the following method :
- short program description
- which actions are sonificated
- the audio files : location on the computer, format, how to modify them (in case of archive files)
- official audio samples (designed by/for the software society)
- some unofficial audio samples if possible (designed by anonymous users)
Then, a classification of the different functions using a sound and an average morphological portrait were made. From this classification, 48 sounds were choose to convey the six functional categories defined in table 1 :
- In mail - Out mail - Radiolink - Feedback yes - Feedback no - Alarm. The definitions of those categories correspond to the auditory functions that were included in the global project of on-line radio programs.

The **RadioLink** is a specific functional category created for vocal navigation utilities in a streaming radio context ; it can be explained by : - A user is listening to a radio broadcast, and other "audio articles" about the current subject are available on another stream : a RadioLink auditory signal warns the user that other streams are available about this subject.
Or : - A user is listening to a radio broadcast, and something important, that isn't necessarily related to the current subject, is suddenly available from another stream - a goal in a soccer match, a public safety alert, or anything else : a RadioLink auditory signal warns the user.

The *In_msg* category : is an undefined category, something close to the *in mail* category.

Functional categories	Definitions	Nbr
1 - **Mail-In**	"new mail"	6
2 - **Mail-Out**	"successfully sent"	6
3 - **Feedback-Yes**	"success of an action"	6
4 - **Feedback-No**	"wrong or unvailable action"	6
5 - **Alarm**	"a schedule reminder"	6
6 - **RadioLink**	*see below*	12
7. In_msg	*undefined function*	6

Table 1: Definitions of the auditory functions and the number of sounds for each category.

2.2. Procedure

Two experiments were done. The stimuli were amplified by a Yamaha P2075 stereo amplifier and presented diotically over a Sennheiser HD420 headset. The listeners were seated in a double-walled IAC sound booth. The experimental session was run using a Matlab interface[1] running on an Apple computer.

Experiment 1 : timbre representation
Subjects : 18 listeners were recruited for the experiment. Listeners were informed of the purpose of the project (on-line radio) and were computer users. None of the listeners reported having hearing problems.

Procedure : three successive steps have been done :
- Step 1 : Using the computer interface, subjects were asked to group sound samples into as many categories as they wanted, according to their similarities.
- Step 2 : Subjects were asked to describe verbally each category formed according to their meaning and particularly in the context of the on-line radio project.
- Step 3 : Afterwards, the six definitions of the functional categories were presented; the subjects were asked to try to name their perceptive categories using the definitions.

Experiment 2 : functional representation
Subjects : 18 new listeners were recruited for the experiment. Listeners were informed of the purpose of the project (on-line radio) and were computer users. None of the listeners reported having hearing problems.

Procedure : three successive steps were done :
- Step 1 : Using the computer interface, subjects were asked to group sound samples into six categories defined by table 1.
- Step 2 : Subjects were asked to choose in each category the sample that seemed be the most representative of its groups (*prototype*)
- Step 3 : Then, the subjects were asked to select and remove one sample per category that was perceived unpleasant or offensive.

3. RESULTS

For each experiment data were compiled into a similarity matrix where the values indicate the number of subjects sorting each pair of stimuli within the same class. The similarity matrix was then processed through a hierarchical cluster analysis. The results are represented by a tree indicating the similarity

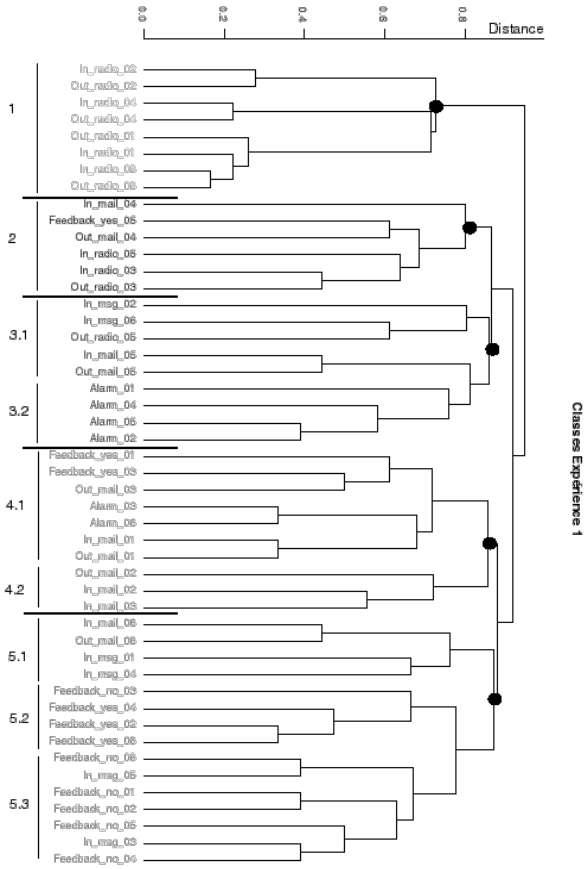

Figure 1 : Perceptive representation. Labels of the sound samples correspond to the functional category they should convey.

distances between the sound samples : the more nodes between two samples, the less often these samples were associated together [6].
Figures 1 and 2 show the results of the analysis respectively for experiments 1 and 2.

3.1. Experiment 1

From the cluster analysis, it's possible to distinguish five major classes. The classes 3, 4 and 5 can be split up respectively into 2, 2 and 3 sub-classes.
A global examination of the classes shows that :
- Sound samples from class 1 can be caracterized by a melodic profile, such as "Jingle".
- Sound samples from classes 2 & 3 correspond to identified sound sources such as musical instruments, "metallic" sounds, domestic devices or warning signals (alarm clock, car horn, …).
- Sound samples from classes 4 & 5 correspond to abstract sounds such as sounds that are used in cartoons. In addition, samples from class 5 sound like wood percussion (marimba, woodblock...).

[1] The experimental interface was developped by Vincent Rioux at Ircam (rioux@ircam.fr)

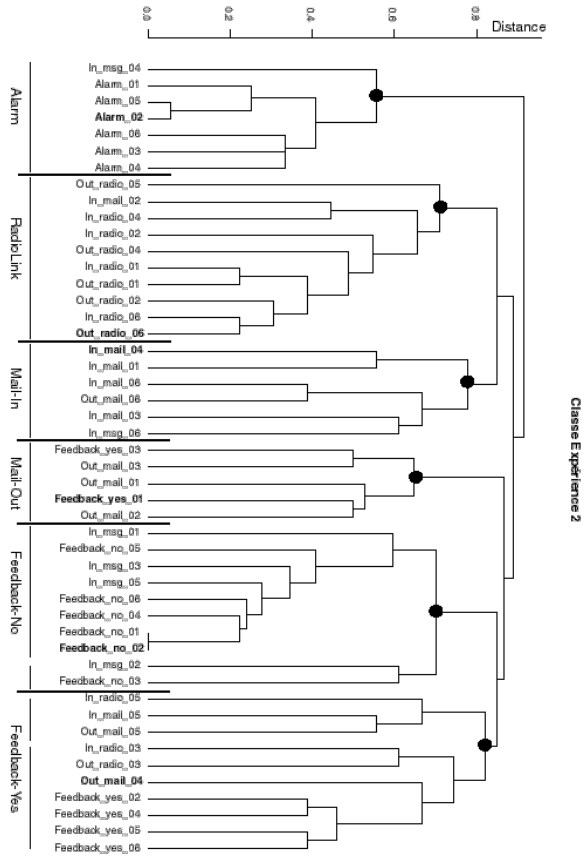

Figure 2 : Functional representation. Labels of the sound samples correspond to the functional category they should convey. Labels of the six functional categories are indicated on the left and the prototype sound for each category is bold.

The results from step 3 reveal that only classes 1, 3.2 and 5.3 are associated with a functional category, respectively RadioLink, Alarm and Feedback-No. In addition, the sound samples belonging to these three classes were choosen to convey this meaning. That means that the functional representation fits well with the perceptive representation for the categories RadioLink, Alarm and Feedback-No. In return, Feedback-Yes and Mail-In are dispersed into most of the perceptive categories which means that there are no common perceptive factors to represent those categories.

In summary, classes of experiment 1 can be described in terms of (1) perceptive and functional association and (2).important / prominent aspect :

Class 1
1. [jingle=long=musical=+/- complex] with ["welcome"]
2. Musical aspect

Class 2
1. [short, high-spectrum/pitched samples] with ["acknowledgment of a usual action"]
2. "Hi medium only"

Class 3
1. [repetition] with ["alarm"]
2. Highly realistic vs abstract

Class 4
1. None
2. Belongs to a certain culture ; in this case, cartoon like.

Class 5.1
1. [very short samples] with ["just a notice"]

Class 5.3
1. [low spectrum ~ (contant melodic profile) & (double repetition)] with ["error"]

More details can be found with sound samples at http://www.ircam.fr/equipes/design/activites/sonification/VI_Studies/Study01/index.html. The web page indicates three kinds of descriptions, extracted from steps 2 & 3, for each class : - 1 - identified sources - 2 - subjective appreciation - 3 - eventual auditory function.

3.2. Experiment 2

As was expected the cluster analysis reveals six categories (figure 2). A first overview shows that the classifications are different between experiment 1 and 2.
- Classes 1, 3.2 and 5.3 from experiment 1 are entirely included respectively in the classes RadioLink, Alarm and Feedback-No of the second experiment. This result confirms the first experiment's conclusion. In other words, those three groups of sounds are perceptively unambiguously related to those three functions.
- In addition, it's interesting to notice that some sound samples selected to convey a function are separated from the others in experiment 1, for example Alarm_03 and Alarm_06, and are now grouped together. This observation can be made especially for the In_mail sound samples that are dispersed into different perceptive categories in experiment 1 and are mostly grouped into the function category Mail-In in experiment 2.
- On the orther hand, the function category Feedback-Yes is particularly unhomogeneous but contains 4 vs 6 sound samples selected to convey this function.

In the second step of the experiment, subjects were asked to select one sound from each category that is the most representative of the associated function (prototype).
- It appears again that the sound samples selected to convey the function categories Alarm, RadioLink and Feedback-No are the most representative of those categories with the respective percentages 84, 89 and 79. The prototype sounds are Alarm_02, Out_Radio_06 and Feedback_No_02. Table 2 shows the summary of the results for RadioLink category.
- In the other hand, the sound samples selected to represent the In-Mail category are quite different. For the Feedback-Yes category the sounds which are the most representative belong to the group of sounds selected to convey the Out-Mail function (47%, see table 3). In the same way, the most representative sound of the Feedback-Yes belongs to the Out-Mail category.

RadioLink		Family
Proportion	Name	
42%	*Out_Radio_06*	**Radio_In/Out**
16%	Out_Radio_02	
11%	Out_Radio_05	
11%	In_Radio_05	**89%**
5%	In_Radio_06	
5%	Out_Radio_01	
5%	Alarm_01	
5%	In_Mail_04	

Table 2 : Selected prototypes for the RadioLink category

Feedback Yes		Family
Proportion	Name	
11%	Feedback_Yes_05	**Feedback_Yes**
5%	Feedback_Yes_02	
5%	Feedback_Yes_03	**26%**
5%	Feedback_Yes_06	
5%	Feedback_No_05	
5%	In_msg_06	
5%	In_Radio_06	
26%	***Out_Mail_04***	**Out_Mail**
5%	Out_Mail_02	
5%	Out_Mail_03	**47%**
5%	Out_Mail_05	
5%	Out_Mail_06	
5%	Out_Radio_02	
5%	Out_Radio_05	

Table 3 : Selected prototypes for the Feedback-Yes category

In the last step, subjects removed from each category the most unpleasant sound. Five samples were found unpleasant by more than 50% of the listeners. The conclusion that can be drawn from listening to these samples is that one there must not be too much energy in the hi-medium frequencies ("be smooth !").

4. Conclusion

The results of experiments 1 and 2 show that some perceptive and functional categories are quite well related, such as Alarm, RadioLink and Feedback-No. There's a strong link between the perceptive and the functional representations. In addition, sounds usually used in different softwares to convey those functions seems to be well designed.
In the other hand, when a specific instruction is given about the meaning, sounds are more easily grouped to fulfill the functional category they are supposed to convey. It's the case for example for the In-Mail category.
Finally, it seems that the sounds usually used to inform about an Out-Mail action or a Feedback-Yes acknowledgment don't suit very well.

Recommendations :
- **Alarms** that sound like "real life" alarms with a repeated pattern are considered as the most representative of their category.
- High spectrum, simplicity and lightness can produce recognizable **feedback yes** samples.
- Sounds lasting ~ 0.25 sec, with stable melodic profiles and low spectra, repeated twice, fit very well with a **feedback no** function.
- Clear downhill profiles are good for **mail in** samples, although stable or slightly uphill alarms are not bad.
- An image of something being thrown in the air is excellent for **mail out** samples. Clearly uphill samples are also very good.
- A "technological sounding", quite complex and long sample is appreciated as a **Radio Link**. "Technological sounding" seems the most important criterium.

5. Acknoledgment

This study is part of an online thematic radio project, named "RadioThem" (RNRT) and was made in association with the software firm Hyptique, Radio France and France Telecom.

6. Reference

[1] P. Susini, S. Vieillard, E. Deruty, B. Smith, C. Marin. "Sound Navigation : Sonified Hyperlinks". ICAD, Kyoto, Japan, 2002.
[2] Ph. Truillet and al., " Présentation multimodale de documents électroniques structurés". Colloque sur la multimodalité, Grenoble, France, 2000.
[3] S. Maury, S. Athènes and S. Chatty. "Vers une interaction basée sur le rythme : les menus rythmiques". IHM 99: 11ème Conférence Francophone sur l'Interaction Homme-Machine, 1999.
[4] W. C. Osborn, R. W. Sheldon, and R. A. Baker."Vigilance performance under conditions of redundant and nonredundant signal presentation," J. Appl. Psychol., 47, pp. 130-134, 1963.
[5] O. Le Bohec, and E. Jamet. "Effet de redondance et effet de modalité réunis." Journées d'étude en Psychologie ergonomique, Nantes, France, 2001.
[6] E. Rosch and B. Lloyd, "Cognition and categorization" L. Erlbaum, Hillslade (N.-J.), London, 1978.