FP6-NEST-PATH project no: 29085
Report Version: 1.0
Report Preparation Date: June 1, 2007
Classication: Pub.
Deliverable no. 4.1 part 2

Closing the Loop of Sound Evaluation and Design (CLOSED)

Deliverable 4.1

Everyday sound classification

Part 2
Experimental classification of everyday sounds

Olivier Houix, Guillaume Lemaitre,
Nicolas Misdariis, Patrick Susini (IRCAM)

Contents

1 Introduction

In the first part of this deliverable, we have provided a review of the literature on everyday sound perception and classification. This review has identified two kinds of listening: musical listening and everyday listening. While in the former case listeners focus on the properties of the sound signal, abstracted from its cause, everyday listening is directed toward the event causing the sound: the sound event. Identifying the sound source is the most spontaneous reaction when listening to everyday sound events. Therefore, the first part of this deliverable has provided as well a review of publications studying how listeners identify sound events and organize this knowledge. Particularly, we have focused on how listeners categorize perceived sounds, since this review has reported several experimental and theoretical data suggesting that the organization of perception and cognition is categorical.
In the first part of the deliverable, there was also a review of the ideas founding the domains of sound quality and design. And among these ideas has emerged the concept of basic interaction design. In short, this concept relies on the idea of providing to the designers the basic bricks, or categories, of interaction designs, to let them imagine new functions, new interactions with products through sound. Finally, we have provided a taxonomy of the sound syndissertation algorithms developed by the partners of the CLOSED project. This taxonomy is based on a classification of sound events imagined by researchers in the field of ecological psychology.
Thus, perceptual classification of everyday sound events appears clearly to be a key issue for the CLOSED project. Studying how listeners classify sound events has several interests: from a theoretical point of view, it allows to understand the listeners' ability to identify the causes of sounds, and how listeners organize and recover knowledge about the sounds. From a design perspective, it will allow us to define what are the basic categories of sounds. Theses categories of sounds will have to be associated with the basic interactions that users may experience with a product through the sounds, identified elsewhere in the CLOSED project (see deliverable 3.1). Moreover, these perceived categories of sounds will indicate what categories of algorithms are required or need refinements. Therefore, the classification model provided by this document is considered as recommendations for the development of sound syndissertation algorithms.
There are many ways of judging, appraising, and classifying sounds (as it will be reviewed in the first part of the document). But we are specifically interested in the perception of sound events: how listeners perceive and categorize physical events causing sounds. Indeed, the perception of the musical qualities of sounds has already been studied extensively, and a lot of results are already available and useful from a design perspective. Sound event perception is less documented, and we need particularly to determine what are the basic categories of sound events. Thus, this document will spend much effort on setting up a methodology to specifically study the perception, and the classification, of sound events.
This document is made of two main parts. First, we will summarize in section 2 the results of perceptual classification of everyday sounds already available in the literature. From this summary, we will define some hypotheses on how listeners classify everyday sounds. These hypotheses will be experimentally tested in the second main part of the document. These experiments will also allow us to form the first bricks of a classification model for everyday sound events (section 3).

2 Experimental classification of everyday sounds: a discussion

2.1 Results from available perceptual classifications of everyday sounds

Only few results of experiments in which listeners have to classify environmental sounds have been reported. Among these publications, we choose to focus in this section on the results of four studies: Nancy Vanderveer's Ph. D. dissertation [12], Frédérique Guyot's Ph. D. dissertation [2], a paper by Michael Marcell and al. [7], and Yannick Gérard's Ph. D. dissertation [3]. Each of these studies provides data which illustrates the variety of results that such experiments may provide. The outlines of Vanderveer's and Guyot's work were summarized in the first part of the deliverable 4.1 and we will focus here in more details on the classification experiments. Besides, it must be noted that Guyot's dissertation provides the individual results of the classification experiments, while the other works provide only "averaged" results. Guyot's, and Gérard's results (particularly the verbal descriptions of the categories) were originally written in French. Hence the descriptions of the categories provided by the experiment participants have been translated from the French, which might have missed the subtleties of these descriptions.

2.1.1 Summarizing the classification results

Vanderveer's experiments

In chapter 6 of her 1979 Ph. D. dissertation, Nancy Vanderveer reports the results of two free classification experiments. In each experiment, 20 participants listened to 20 sounds recorded on a tape, wrote down descriptions of the sounds on gummed labels, and then had to "sort the items based on the similarity of sounds" (p. 205). The experimenter provided to the participants some examples of "obvious" groupings (e.g. "filing" and "sawing" sounds). The two experiments were made with different sets of sounds and different participants.
For both experiments, she computed from the sorting data the co-occurence matrix and focused on the pairs of sounds grouped together by a number of participants significantly higher than chance.
Figure 2.1 and 2.2 reproduce the results of the two experiments. The clusters of sound descriptions link the sounds paired by a significant number of participants. From this rough analysis, she concluded that the basis for sorting the sounds was "twofold: acoustical similarity (or temporal patterning in particular) and relatedness of source events (or meanings)" (p.214). Indeed, the participants appeared to have sorted the sounds because they were caused by similar events (e.g. "drop pen", "drop can", "drop wood"), or because they shared acoustical similarities (e.g. "pin box", "sawing", filing"). Therefore, she concluded that the most important determiners of perception seemed to be the following (the labels *van...* have been added by the authors and will be used further in the document):
"A. Temporal patterning variables:
1. percussive vs. continuous events (*van1*)
2. rhythmic patterning (*van2*)
3. attack or decay time (*van3*)
B. Resonance or other characteristics of particular objects, surfaces or substances. For example:
1. metal (*van4*)
2. paper (*van5*)
3. rough texture (*van6*)

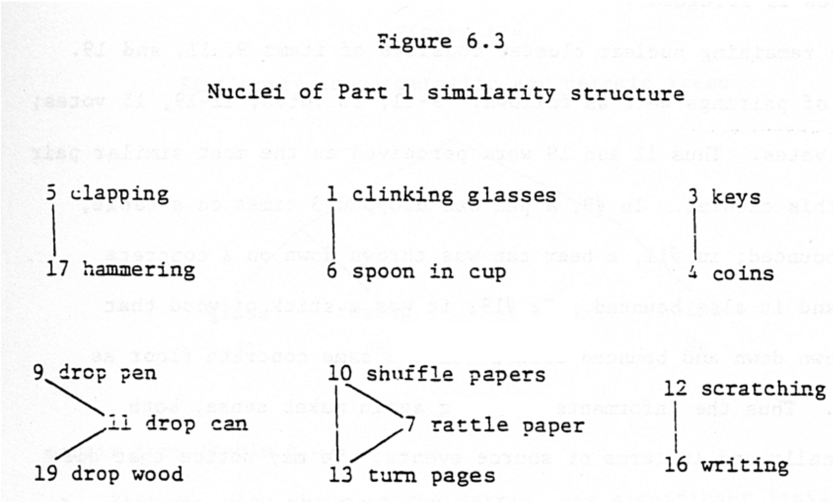

Figure 2.1: Vanderveer's analysis of the results of her first classification experiment. Reprinted from [12].

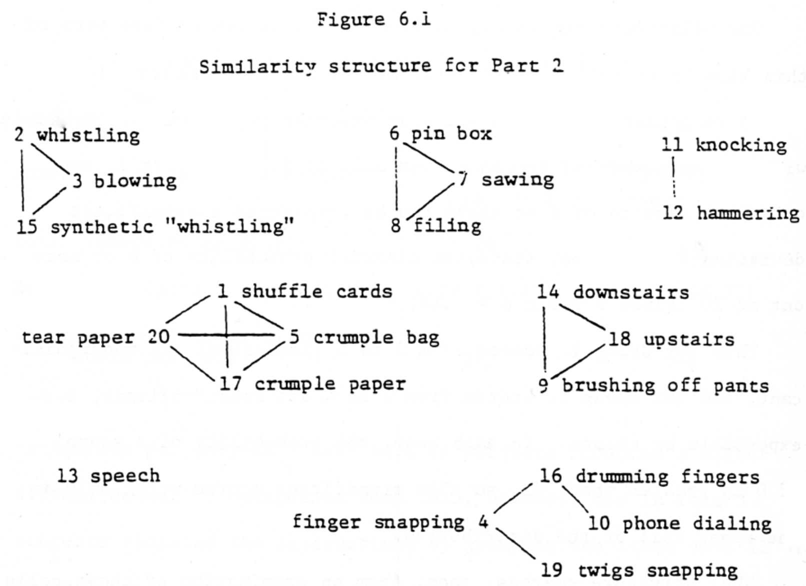

Figure 2.2: Vanderveer's analysis of the results of her second classification experiment.Reprinted from [12].

4. etc." (p. 225)

Guyot's Ph. D. dissertation

Frédérique Guyot reports in the chapter 5 of her dissertation (1996) a classification experiment. She studied 25 "domestic" noises (roughly 3 seconds long, see Table 2.1). Participants had first to group together noises, "according to their perceptual similarities" (p. 114). Afterwards, they had to characterize verbally the categories.

Two participants made categories on the basis of the acoustical properties of the sounds and explained the categories by describing the signal properties. One made three categories on the basis of the type of excitation (mechanical, electrical, electronic). The other ones made categories based on the type of

1. Cupboard door creaking	10. Telephone ringing	19. Window closing
2. Pen sharpening	11. Aluminium crumpling	20. Spray
3. Hair dryer	12. Door opening	21. Bottle opening
4. Tearing a sheet of paper	13. Nails cutting	22. Alarm clock tick-tock
5. Alarm clock	14. Vacuum cleaner	23. Velcro
6. Teeth brushing	15. Type writer typing	24. Telephone tonality
7. Match striking	16. Glasses clinging	25. Plates scrapping
8. Jar closing	17. Bread cutting	
9. Keys	18. Aluminium sheet tearing	

Table 2.1: Sounds used in Guyot's classification experiments (translated from the French).

movement creating the sound, on the identified source, on the properties of the signal, etc.
She represented the sorting data in a additive tree (see Figure 2.3). She interpreted the tree as the result of two modes (levels) of categorization: either the sounds are grouped together because they are made (or belong to) by the same source, or they are grouped together because they are made by the same movement/interaction (scratching, rubbing, etc.). Thus, she proposed a hierarchical orga-

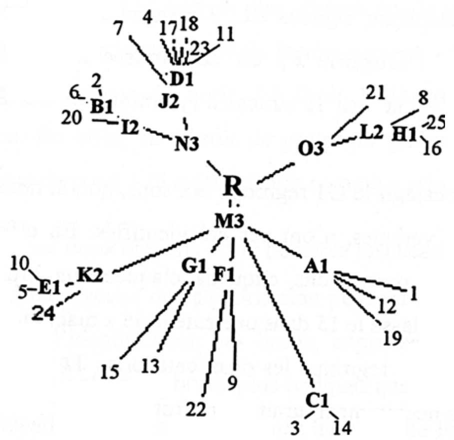

Figure 2.3: Guyot's additive tree obtained from sorting data. The sounds are listed in Tabel 2.1.

nization of sound events, made of three levels: superordinate level, base level, and subordinate level (following Rosch's prototypical categorization [8]): see Figure 2.4.

It must noticed that this organization is only based on descriptions of the physical events causing the sounds. However carefully examining the individual classification provided by each participants show that some of them also made categories of sounds grouped together because of acoustical similarities ("continuous", "musical", "short", "high in pitch", etc.) of because of a similar appraisal ("very unpleasant", "funny", "aggressive", etc.)

Marcell et al. study

In their 2000 paper [7], Michael Marcell and his colleagues aimed at building a normalized corpus of unambiguously identified and named sounds. Their third experiment consisted in asking 37 listeners to freely classify a set of 120 environmental sounds, and to describe the categories they made. "Participants were told that categorization involves placing something with other objects that have similar characteristics and are members of the same group" (p. 853). Then, two independent judges reviewed the descriptions and grouped those judged as equivalent in meaning. They found 23 categories listed in

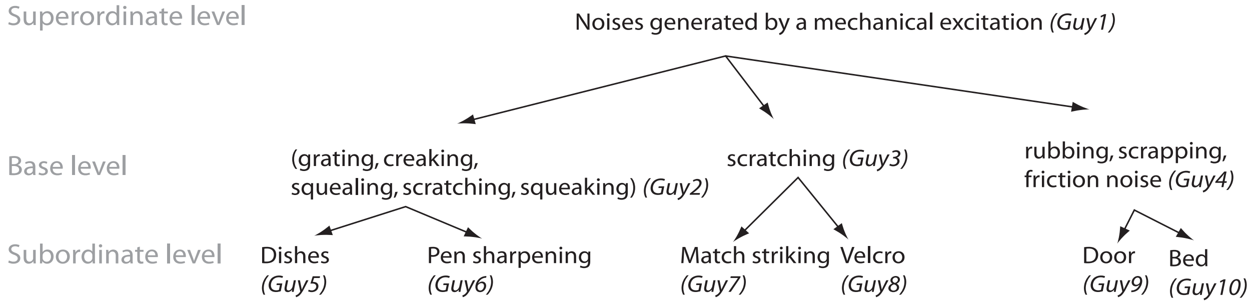

Figure 2.4: Guyot's proposal of a hierarchical organization of domestic noises (translated from the French). The labels *Guy...* have been added by the authors and will be used further in the document.

Table 2.2. The sounds studied by Marcell et al. were made by a large variety of different sources, from

4-legged animals (*Mar1*)	Game/recreation (*Mar7*)	Musical instrument (*Mar13*)	Sleep (*Mar19*)
Accident (*Mar2*)	Ground transp. (*Mar8*)	Nature (*Mar14*)	Tool (*Mar20*)
Air transp. (*Mar3*)	Human (*Mar9*)	Pet (*Mar15*)	Water (*Mar21*)
Bathroom (*Mar4*)	Hygiene (*Mar10*)	Reptile/amphibian (*Mar16*)	Weapon (*Mar22*)
Bird (*Mar5*)	Insect (*Mar11*)	Sickness (*Mar17*)	Weather (*Mar23*)
Farm animal (*Mar6*)	Kitchen (*Mar12*)	Signal (*Mar18*)	

Table 2.2: The 23 categories obtained by Marcell and al. from the free classification of environmental sounds. The labels *Mar...* have been added by the authors and will be used further in the document.

animals to machines. The categories that appeared in their experiment were very general, describing large categories of sources ("air transportation", "tool"), locations ("kitchen", "bathroom", "nature") or abstract ideas ("hygiene", "sickness", "sleep").

Gérard's dissertation

Yannick Gérard's 2004 Ph.D. dissertation studied the "semantic memory". In chapter five, he reports two classification experiments of everyday sounds. In the first one, 30 participants had to group together 24 sounds "which they may hear together in the environment". In the second one, 30 other participants had to group together sounds "on the basis of their acoustical characteristics, independently of their meaning".
He analyzed the results of the classification with cluster analyses. Figure 2.5 reports the results of the first experiment. The dendrogram obtained from the cluster analysis separates the sounds of inanimate objects and animate beings. These categories are divided into thematic subcategories: sounds which happen inside a house, transportation noises, sounds made by animals evoking holydays , sounds made by farm animals. Figure 2.6 reports the dendrogram obtained from the second experiment. In this case, the categories built from the sorting data group together sounds sharing acoustical similarities: same rhythmic, pitch, amplitude patterns.

2.1.2 Discussing these results

Comparing the results of these classifications clearly shows that listeners use different strategies, when they have to group together the sounds. They group together sounds, because they:

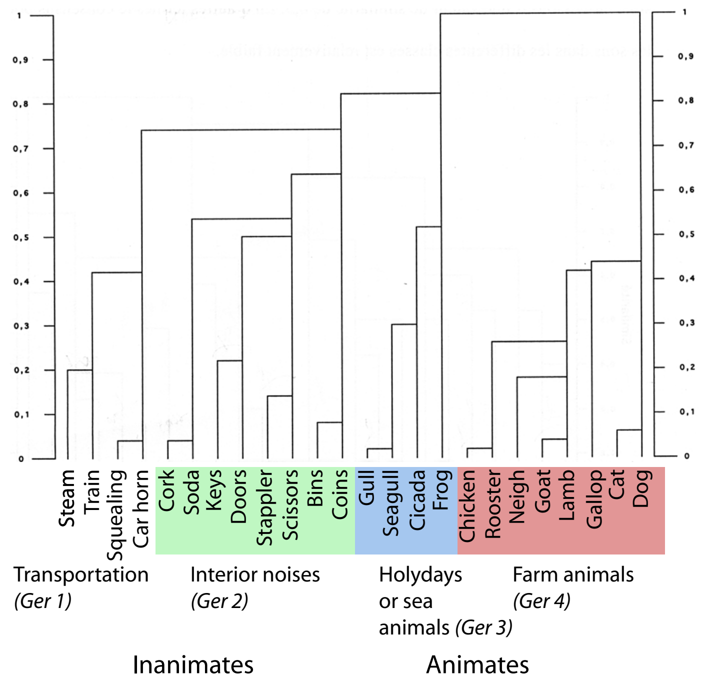

Figure 2.5: Gérard's dendrogram obtained by asking listeners to group together sounds "which they may hear together in the environment" (adapted and translated from the French). The labels *Ger...* have been added by the authors and will be used further in the document.

- share some acoustical similarities (same timbre, same duration, same rhythmic patterns): Vanderveer, Gérard,
- are made by the same kind of action/interaction/movement: Vanderveer, Guyot,
- are made by the same type of excitation (electrical, electronical, mechanical): Guyot,
- are produced by the same object (the same source): Guyot, Marcell et al.,
- are produced by objects fulfilling the same (abstracted) function: Gérard, Marcell et al.,
- occur in the same place or at the same occasion: Marcell, Gérard.

Some strategies to group together the sounds are based on the signal properties, or on the properties of the physical cause of the sound, and require little interpretation. Some others imply to identify precisely what has made the sound, to infer the situation in which the sounds have occurred, who was responsible for the event, for what reason we attribute a meaning to the sound, etc. Attributing this meaning thus relies on the knowledge of the listeners, requires their interpretation and is influenced by the context. By "knowledge", we mean any kind of knowledge, including "everyday knowledge", or experience, as well as learned skills.

Since grouping together sounds actually involves assessing the similarities between the sounds, we can therefore define three *types of similarities* used by the participants of these studies to group together the sounds:

- The similarity of acoustical properties: *acoustical similarity*
- The similarity of the physical event causing the sound: *event similarity*
- The similarity of some kind of knowledge, or meaning, associated by the listeners to the identified object or event causing the sound: *semantic similarity*

In Figures 2.7, 2.8 and 2.9 we have drawn the categories from the aforementioned studies, sorted into the three types of similarity, and tried to figure out a common organization. In each of these figures, the different categories of these different studies are represented with a different color, and indicated by a label referring to the name of the study (*van* for Vanderveer, *guy* for Guyot, *mar* for Marcell and al., and *ger* for Gérard). The proposed organization is represented by black and white boxes and

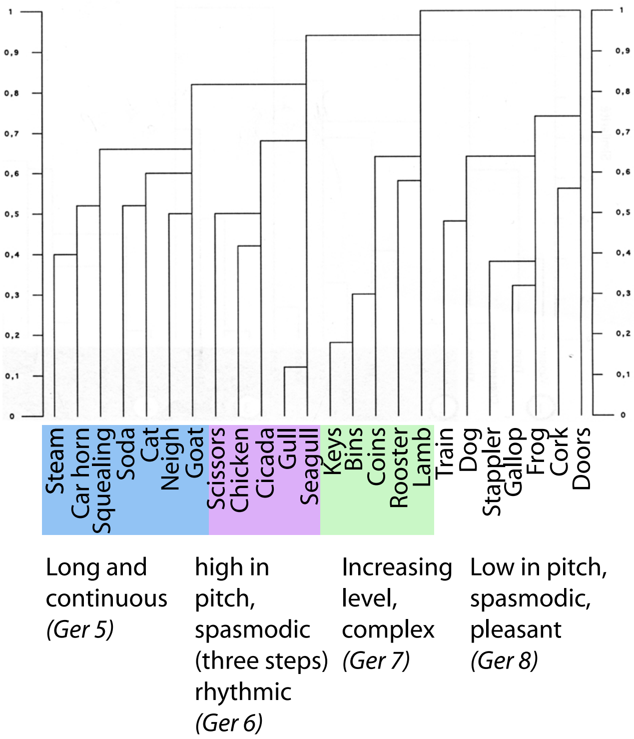

Figure 2.6: Gérard's dendrogram obtained by asking listeners to group together sounds "on the basis of their acoustic characteristics, independently of their meaning"(adapted and translated from the French). The labels *Ger...* will be used further in the document.

arrows.
Figure 2.7 draws the categories of sounds grouped together because they belong to general thematics (e.g. "hygiene", "sleep"), because the events causing the sounds occur on the same location (e.g. "bathroom", "kitchen"), or because the objects identified as the causes of the sound belong to the same category(e.g. "tools"). Except for these thematics, we do not have made a specific organization.
Figure 2.8 reports the categories of sounds grouped together because of acoustical similarities. We have separated the sequential, timbral and pitch properties. In each of these classes of properties, we have have tried to draw a hierarchical organization of the properties.
Figure 2.9 reports the categories of sounds grouped together because of similarities of the events causing the sound. In each of these classes of properties, we have tried to draw a hierarchical organization of the properties, generalizing the organization proposed by Guyot.

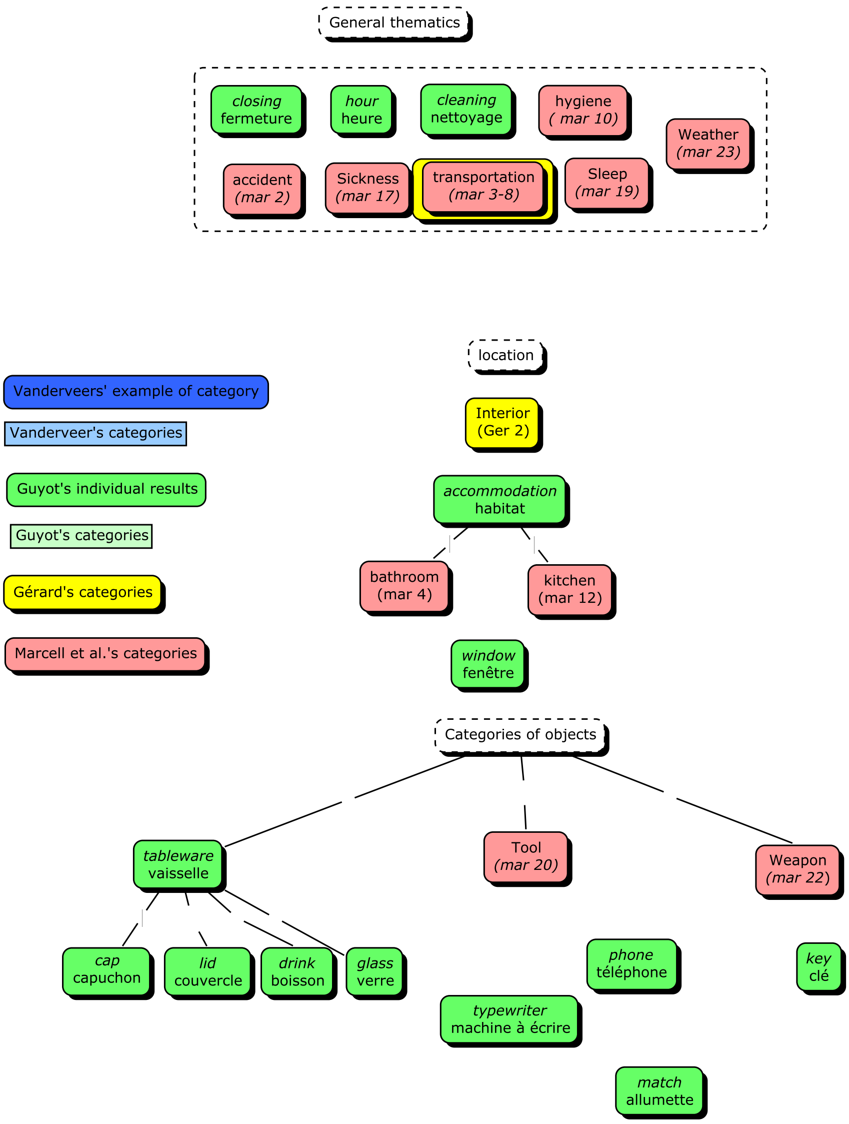

Figure 2.7: Categories from Vanderveer's, Guyot's, Marcell et al.'s and Gérad's work, corresponding to items grouped together because of their semantic similarity. The black and line boxes and arrows are an assumed organization of these categories.

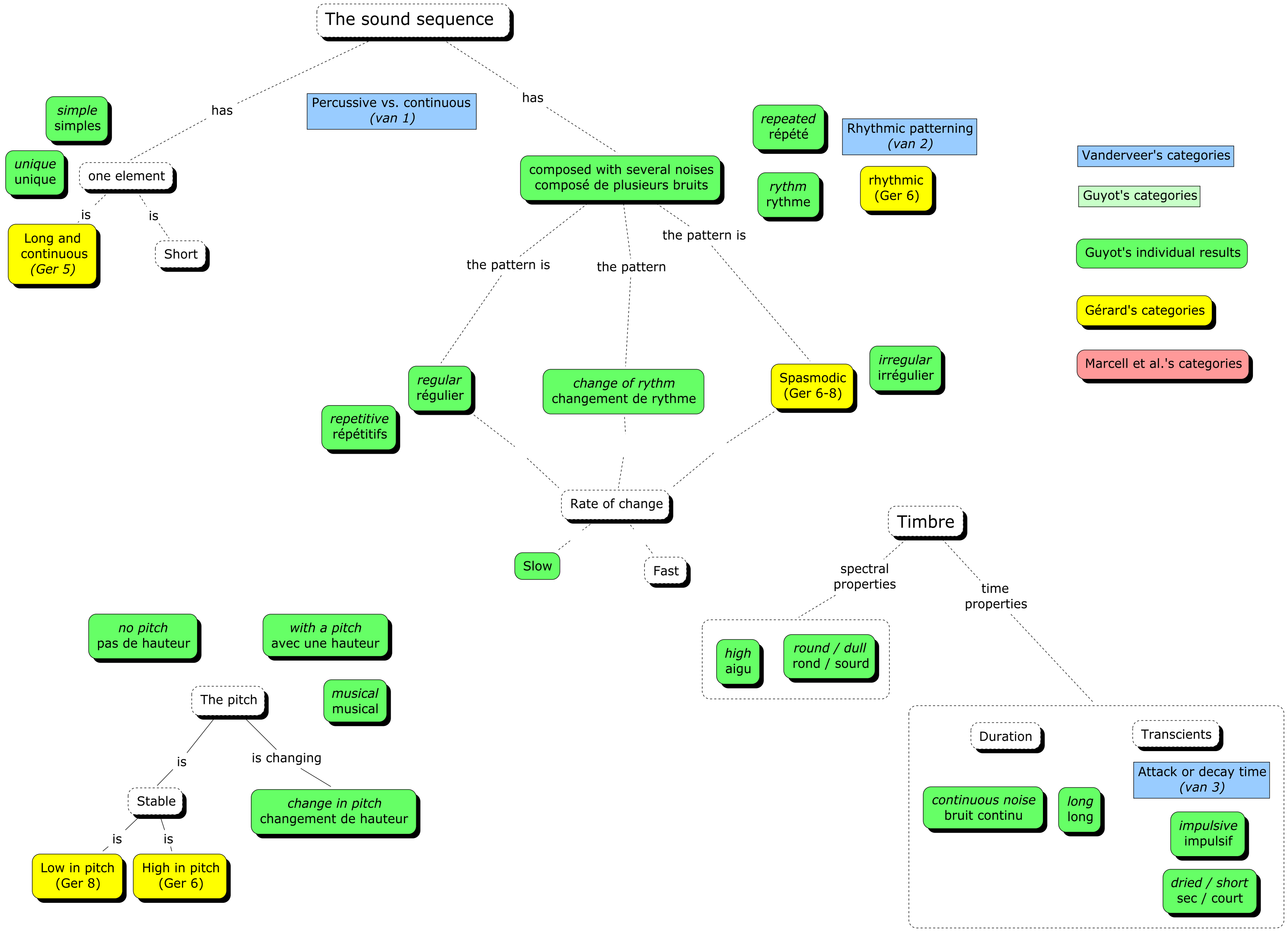

Figure 2.8: Categories from Vanderveer's, Guyot's, Marcell et al.'s and Gérad's work, corresponding to items grouped together because of their acoustical similarity. The black and line boxes and arrows are an assumed organization of these categories.

All these results show that classifying sounds is strongly related to the degree of sound identification. Indeed, except in the case where participants group together sounds because of the acoustical similarities, grouping strategies require to recognize at least the event (action and object) that has caused the sounds.
The reason why participants have used different strategies is not really explained by these studies. First of all, the instructions given by the experimenter to the subjects were not exactly the same. For instance, the instructions provided by Vanderveer gave exemples of sounds grouped together both because of the similarity of the physical event and of the sounds themselves (”sawing” and ”filing”). And subjects made classes of sounds mainly on the basis of these two criteria. By changing the instruction, Gérard changed the results of classifications as well.
These different strategies might have been also influenced by the listeners ability to analyze sounds. The participants in Vanderveer's experiment were university students. Those in Guyot's experiment were members of her laboratory. Marcell et al. worked with psychology college students, and Gérard did not provide any biographical data from the participants. None of these authors seems to have recorded the expertise of the participants with sounds, neither tried to assess the influence of the expertise on the individual results.
These results show another issue: as categories are described by words, a question is whether the categories labeled by words in a language actually correspond to perceptive categories. An interesting example is found in Guyot's dissertation [2]. Results are different when participants have to classify freely sounds, and then to name the sounds, from when other participants have to associate sounds with the category names, although provided by the participants themselves. Perceptual categories and linguistic categories dot not seem to fit together.
The question is then to understand why listeners may use these different strategies.

2.2 Hypotheses

From the previous discussion, we can formulate five hypotheses:

- H1. When listeners are required to group together sounds ”freely”, they may group together the sounds according to different kinds of similarities (acoustical, event, semantic),
- H2. When the sounds are not identifiable, listeners can only group together sounds which are acoustically similar,
- H3. When the sounds are identified, listeners can choose to group together the similar sound events, or to group together sounds because of property related to their knowledge of the source
- H4. The categories of sound events are hierarchically organized

We also suspect that skilled listeners will have more ability to assess acoustical properties of sounds than listeners without any musical or sound education (like professional musicians and sound engineers). Therefore we introduce another hypodissertation.

- H5. Listeners with an expertise in analysing sounds will tend to group together sounds using acoustical similarities more often than naive listeners.

The goal of the following experiments is to test these hypotheses.

2.3 Outline of the experiments

According to our hypotheses, two factors influence the strategies used by the listeners to classify a set of sounds without any other instruction. The degree of identification of the source of the sounds and the expertise of the participants. Assessing the expertise of the participants may be easily done by having them filling a questionnaire. Measuring the degree of identification of a sound requires however specific methods.
Back to the work of James Ballas (see the state of the art in the first part of this deliverable), we choose to focus on the *causal uncertainty* Hcu (see next section for more details). Causal uncertainty measures how many different sources listeners identify, for a given sound. It has been shown in [1]

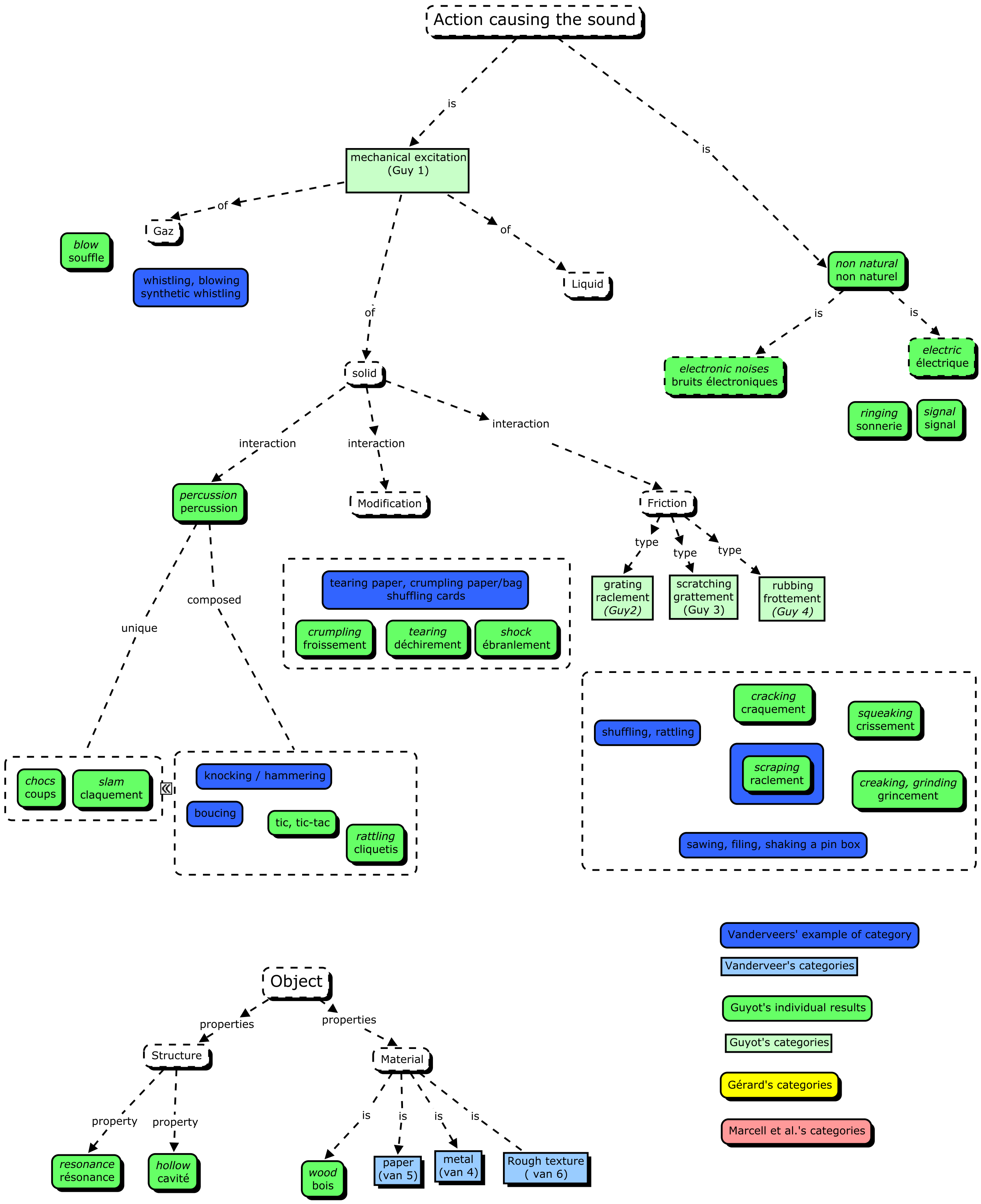

Figure 2.9: Categories from Vanderveer's, Guyot's, Marcell et al.'s and Gérad's work, corresponding to items grouped together because of the similarity of the physical event causing the sound. The black and line boxes and arrows are an assumed organization of these categories.

that causal uncertainty is one of the factor explaining the identification of everyday sounds.
We will focus on sounds recorded in a kitchen. There are two reasons for using kitchen sounds: first, we need to have sound sources from everyday life familiar for all the participants, and a single context that they can easily interpret. Second, there is a large variety of sounds occurring in a kitchen (machines, solid, liquid interactions, gaz, electronic alarms, etc.).
There will be two experiments. In the first experiment, we will measure the causal uncertainty for a set of 96 sounds. The aim of this first experiment (1.1) will be to assess the reliability of the Hcu measurement methodology, and to select a subset of sounds, providing an homogeneous sample of Hcu values, from sounds with a very low causal uncertainty, to sounds with a very high uncertainty. The second experiment (1.2) will have three steps: first we will make a free classification experiment, based on a sample of 60 sounds. Second, participants will also have to verbally describe their categories, and, thirdly, to decide which type of similarity they have used to make the category.
Examining the raw results of the classification will allow us to test H1 (participants use different strategies). Then we will analyse how the causal uncertainty of the sounds, and the expertise of the participants have influenced the choice of the strategy used to group the sounds. This will allow us to test hypotheses H2, H3 and H5. Finally, cluster analyses of subsamples of the data will allow us to test hypodissertation H4.

3 Experiments

3.1 Experiment 1.1: measuring the causal uncertainty

3.1.1 Introduction

The goal of experiment 1.1 is to "measure" the "degree of identification" of a set of sounds. One of the major work studying identification of environmental sounds is the series of papers published by James Ballas until 1993 (see section 1.1.3 of the first part of the deliverable for an overview). Particularly, he showed in [1] that the performance of identification were related to different variables, including *acoustic variables*, *ecological frequency* (the frequency with which a listener encounters a specific sound event in his everyday life), *causal uncertainty* (measured as the amount of reported alternative causes for a sound) and *sound typicality*. Results suggested that sound identifiability is related to the ease with which a mental picture of the sound is formed, context independence (when the sound can be identified easily without context), the familiarity with the sound, the similarity of the sound to a mental stereotype (the author indeed showed sounds that were more typical than others), the ease using words to describe the sound, and the clarity of the sound.
Since we are working in these experiments with sounds belonging to an identified and familiar context (the kitchen), and because we can safely assume that the degree of familiarity with the sounds occurring in a kitchen is shared by a random set of listeners (except for cooks), causal uncertainty appears to be potentially a good indicator of the degree of identification of our set of sounds.
To measure causal uncertainty, Ballas made an experiment in which listeners had to listen to each sound of a set and to write down one description of the possible cause of the sounds. After having sorted the descriptions into categories of identical descriptions, he defined the measure of causal uncertainty with the following expression :

$$Hcu_i = \sum_{j}^{n} p_{ij} log_2 p_{ij} \qquad (3.1)$$

where Hcu_i is the measure of causal uncertainty for a sound i, j are the different alternative causes provided by the participants, p_{ij} is the proportion of responses for a sound i as being caused by cause j, and n is the number of categories of identified causes for a sound i.

3.1.2 Method

participants

29 participants (14 women and 15 men) volunteered as listeners and were paid for their participation. They were aged from 20 to 47 years old (median: 35 years old). All reported having normal hearing. None of the participants reported being a professional cook or having a professional activity related to cookery. The participants were all French native speakers or demonstrated high skills in French.

Stimuli

All the sounds were recordings of activities occurring in a kitchen. The choice of the kitchen environment reflects two aspects of the project :

- Kitchen is a framework well adapted to produce scenarios related to new sound design products. The kitchen scenario is therefore a starting point of our different studies (see Deliverable 3.1 for examples)

- A large variety of sounds sources can occur in a kitchen environment (for example : motor, liquid or mechanical sounds, ...), belonging to different categories of objects and sharing different semantic relations, as illustrated in Figures 2.7 and 2.9.

To select the sounds, we asked members of our laboratory to answer a questionnaire describing which sounds they usually hear in their kitchen. This questionnaire helped us to define the usual sources of sounds occurring in a regular kitchen.
Then, the sounds were chosen among different commercial sound libraries : Hollywood Edge Premiere Edition I, II and III, Sound Ideas General Series 6000 and Blue Box Audio Wav. We selected 101 sounds representative of the usual sources identified by our questionnaires and our own personal experience. These sounds were stereophonic and, mainly, monophonic sounds with 16-bit resolution and with a sampling rate of 44.1kHz. We converted stereophonic sounds into monophonic sounds by using only the left channel in order to have a monophonic corpus of sounds. Some sounds were then edited to reduce noise at the beginning or at the end, or to isolate a piece of sound in a sequence. The description of the sounds and of the sources are detailed in Table 4.2 p.40.

Apparatus

The sounds used were played by a Macintosh Mac Pro (Mac OS X v10.4 Tiger) workstation with a MOTU firewire 828 sound card. The stimuli were amplified diotically over a pair of YAMAHA MSP5 loudspeakers. Participants were seated in a double-walled IAC sound-isolation booth. Levels were calibrated using a Brüel & Kjær 2238 Mediator sound-level meter. The equivalent average sound level measured for the sound reference (sound number # 99 Table 4.2 p.40) was 70 dB. The experiment was run using the PsiExp v3.4 experimentation environment including stimulus control, data recording, and graphical user interface [10]. The sounds were played with Cycling'74's Max/MSP version 4.6. For each experiment, the order of the sounds was randomized.

Ecological adjustment of sound levels

The sounds, selected across different commercial libraries have been initially recorded with different techniques. For example some recordings were near field recordings (i.e. with the microphone close to the source), while some others were far field recordings (with the microphone far from the source). This may cause problems when playing the sounds to the listeners: for instance, some sources usually sounding a very low level (for example an "ice cube"), played at a too high level may be perceived louder than normally louder sources (for example "water flow") recorded with different techniques, or simply may become unidentifiable, because of this "acoustical zoom" effect.
That is why we adjusted the level of the sounds in order to reproduce the "usual" or "ecological" level of the sounds in a kitchen in a preliminary experiment. We call "ecological level" the natural level of the different sound sources in a kitchen environment.
Six people from our laboratory were presented with pairs of sounds. Each pair was made of the same reference sound followed by the sound to be adjusted. We asked them to adjust the level of the sounds as they would have heard them in their kitchen (see the verbatim of the instructions in Appendix A. p.36). They had to move a cursor changing the level of the second sound. The interface was implemented under PsiExp v3.4 [10]. # the reference sound (# 99 Table 4.2 p. 40) corresponded to "filling a sink with water". This sound was selected because of his high physical level in a kitchen in real situation. For each sound we provided a description of the sound (see Table 4.2 p.40) to ensure that participants had identified the sound to be adjusted.
The result of this experiment are, for each participant, the ratio applied to the sounds to adjust the level. We calculated the average of the ratio coefficients across the six participants for each sound. The average ratio coefficients and the corresponding standard deviation for each sound are detailed in Table 4.2 p.40. The average ratio coefficients are used for all the experiments to change the level of sounds corresponding to their "ecological level".

Procedure for the causal uncertainty experiment

Participants had first to read the instructions (see Appendix A.). They were explained that they would have to listen to sounds recorded in different kitchens. We gave this information in order to avoid individual differences due to a possible recognition of the context during the experiment. For example, some participants could guess that the sounds were kitchen sounds, and thus identify more easily the sounds, while some others not.
Participants were simply asked to indicate the cause of each sound by typing a "noun" and a "verb", following Ballas's procedure [1]. Participants were asked not to employ metaphoric descriptions and to make simple descriptions.
The experiment was divided into three steps. Firstly, participants were provided with 5 examples of sounds (sounds 25, 26, 63, 71, 100, not appearing in the following of the experiment) to get accustomed to the interface. Secondly participants heard all the remaining 96 sounds. The order of presentation was randomized for each particpant. Thirdly participants heard every sound, indicating for each sound the cause of the sound. For this part, participants had only two trials.
The verbal descriptions were written using a computer keyboard.

3.1.3 Results and analysis

The results of two participants were excluded from the analyses because they have taken too much time to do the experiment. We collected the verbalizations of the other 27 participants for each 96 sounds.

Examples of verbalizations

We present here two verbalizations translated from the French, corresponding to the sounds 1 and 72:

Sound 1

- *icecube falling into a glass* / glaçon qui tombe dans verre
- *aspirine falling into a glass* / aspirine tombe dans un verre
- *money pour* / monnaie verse
- ...

Sound 72

- *whipped cream spray can empty* / crème chantilly, bombe vider
- *gas open* / gaz allumer
- *match scrape* / allumette gratter
- ...

see Table 4.2 p. 40 for the description of the sounds.

Analysis of the verbalizations

Overall, the participants used 523 different nouns and 289 different verbs. "Water" was the most cited noun (141 occurrences), and "closing" was the most cited verb. 47 nouns and 44 verbs were cited by more than 14 participants. 295 nouns and 117 verbs were cited only once.
Before calculating the causal uncertainty for each sound, three sorters have analyzed the 96*27 descriptions. The goal of this analysis was to sort the verbalizations for each sound into categories of similar objects and similar actions. Two of the three sorters are the experimenters, called "expert" sorters A and B, and the other sorter is a student in psychology familiar with analyses of verbalizations, but not aware of the goals of the study. We call her the "independent" sorter C. We separated the coding of the "object" and of the "action" in order to calculate two Hcu: one related to the "object" and the other to the "action".
The three sorters used the same set of written instructions (see Appendix A 4.2 p.36). These instructions explained how to do the task and what to consider as equivalent or different categories of actions and objects. We defined categories as "usual categories" of objects or actions, at a basic level

of categorization (see our review of Rosch studies in the first part of the deliverable). We considered an "action" as a physical action, i.e. the process of an agent on an object.
Firstly the task was to analyze what is the action and what is the object (or the subject of the action) in the description. Secondly, the sorters grouped on the one hand the verbalizations describing identical actions, and on the other hand identical objects. Sorters were asked to write the number "one" in a shared column when the descriptions belonged to the same group of similar actions. In a same way, they had to write "one" when the descriptions belonged to the same group of similar objects. An example of this coding is expounded in Appendix A 4.2 p.36. We used this coding in order to compute the Hcu.
After this analysis, for each sound i we have obtained the proportion of verbalizations matching a category j of an identified object and the proportion of responses matching a category k of an identified action. We have finally two sets of data concerning the action and the object or subject of the action.

Calculation of the Hcu

We compute two different Hcu: one for the object of the subjects of the action, and one for the action. The Hcu related to "object of the action" is calculated as :

$$Hcu_i^{object} = \sum_j^n p_{ij} log_2 p_{ij} \tag{3.2}$$

where Hcu_i^{object} is the measure of causal uncertainty related to "object" for a sound i, p_{ij} is the proportion of responses for a sound i matching a category j of similar objects corresponding to cause of the sound , and n is the number of categories of similar objects corresponding to cause of the sound i.
The Hcu related to the "action" is calculated as :

$$Hcu_i^{action} = \sum_k^n p_{ik} log_2 p_{ik} \tag{3.3}$$

where Hcu_i^{action} is the measure of causal uncertainty related to "action" for a sound i, p_{ik} is the proportion of responses for a sound i matching a category k of similar actions corresponding to cause of the sound, and m is the number of categories of similar actions corresponding to cause of the sound i.

Results

This task was very long. Each sorter has spent three days to achieve the coding for "object" and "action". They have found difficult to use always the same criteria during the coding, especially for "action".

Hcu measured for the objects The three distributions of the Hcu measured for the objects have the same profile: see Figure 3.1. They all have a median around 2.8 and a standard deviation around 0.9/1. The values of the HCU are distributed between 0 and 4.5[1]. The correlation between the two sorters A and B are weaker (0.86), than the correlation between independent sorter C and the sorter A (0.91). Nevertheless, these results show that the instructions allowed to obtain similar results for the three sorters.
For each sound, the standard deviations across the three sorters are distributed between 0 and 1.4, with a mean value of 0.31.

[1]With 27 participants, the maximum value is 4.75

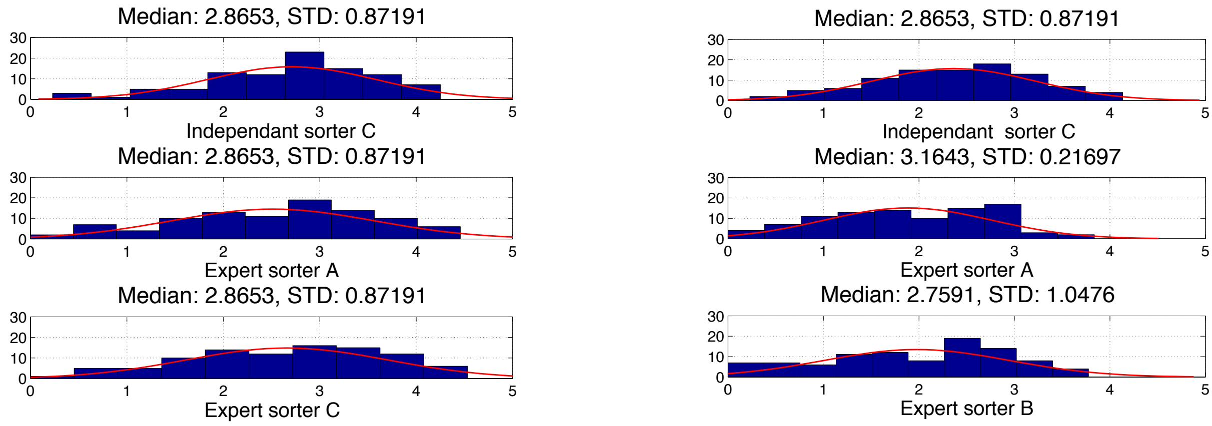

Figure 3.1: Distributions of Hcu for the three sorters: left panel = Hcu for objects; right panel = Hcu for actions.

Results of the Hcu measured for the actions The distributions of these data are dissimilar, medians of the distribution are different (A: 2.2, B: 3.1,C: 2.7) like the standard deviations (A: 0.2,B: 0.2,C: 1.0): see the right panel of Figure 3.1.
The correlation between the sorters are weak (0.67, 0.71, 0.74). Sorters A and B are more correlated together (.74) than with the independent sorter C (A vs C :0.67 and B vs C : 0.71). There are some large differences observed for particular sounds. The instructions given for categories of similar actions seem to be more difficult to use for the expert sorters A and B as well for the independent sorter C. Indeed the instructions given for "action" were less detailed than "object" because of the lack of knowledge on the representation of actions comparing to the objects. The standard deviations across the three sorters are distributed between 0 and 1.9 with the mean at 0.47. These errors are homogeneously spread.

3.1.4 Conclusion and selection of the sounds

The aim of this experiment is to select sounds with a homogeneous distribution of Hcu values

Both measures of Hcu (object and action) contribute to the causal uncertainty of a sound. If we observe the relation between the measures of Hcu_{object} and Hcu_{action} for each sorter, these measures are not independent. For sorter A, the correlation between the two Hcu is 0.61, for sorter B: 0.69, and for sorter C: 0.61. Therefore, we have decided to add both values Hcu for each sound to define a *global* Hcu value:

$$Hcu_i^{global} = Hcu_i^{object} + Hcu_i^{action} \tag{3.4}$$

To select a relevant subset of sounds, we have first removed the sounds for which the variability of Hcu (object or action) values across sorters if too important. For each type of Hcu, we have removed the 10 sounds with the most important standard deviation across the sorters. In total 18 sounds (2 sounds were common to the "object" and "action") were removed from the data: sounds with numbers {3, 4, 11, 027, 28, 36, 39, 46, 48, 50, 55, 65, 70, 73, 75, 76, 78, 93, 96}, see Table 4.1 p. 41 for the description of all the sounds.
To select the sounds, we now consider the median of the Hcu values across the three sorters. The distribution of the median of the global Hcu (figure 3.2) for the 78 remaining sounds indicate a high occurrence of sounds with values of the median of Hcu_{global} between 3 and 7. We therefore remove sounds inside 4 intervals of values of the median bewteen 3 an 7 ({3,4}, {4,5}, {5,6}, {6,7}) in order to smooth the distribution of the data. Finally, we select 60 sounds for the corpus[2]. The correlations of the values of Hcu_{global} on these 60 sounds between the three sorters are high (A vs C : .93, B vs

[2] We removed sounds 34 from interval ({3,4}; sounds 13, 32, 45, 90 from {4,5}; sounds 2, 14, 21, 31, 33, 37, 39, 86, 87 from {5,6}, and sounds 5, 42, 62, 69 from {6,7}

C : .93 and A vs B .96). These correlations indicate that the measure of global Hcu for the corpus of sounds is reliable as a mean to relate causal uncertainty and results of classification. In the remaining of the document, the global Hcu will simply be called Hcu.

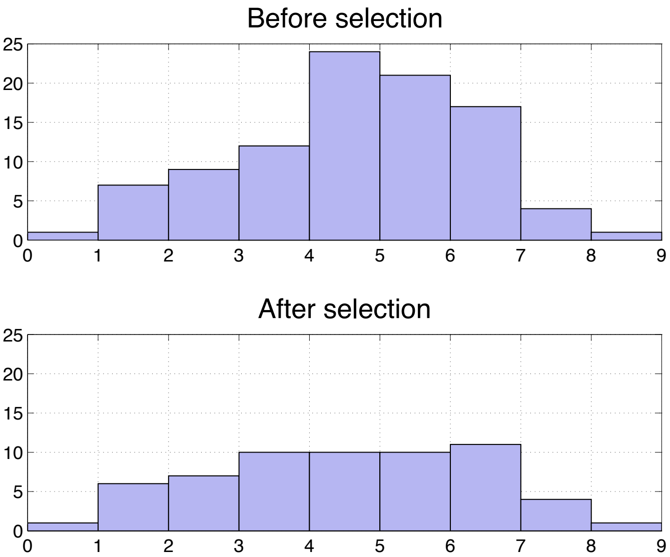

Figure 3.2: Upper panel: distribution of values of Hcu Global for the 78 sounds; Lower panel: distribution of values of Hcu Global for the 60 sounds remaining after selection

3.2 Experiment 1.2: free classification of environmental sounds

3.2.1 Experimental protocol

Participants 30 participants (12 women, 18 men) volunteered as listeners and were paid for their participation. They were aged from 19 to 64 years old (median: 32 years old). All reported having normal hearing. None of these participants had previously taken part to experiment 1.1. The participants were preliminary selected on the basis of questionnaires they had filled in previous experiments. After the experiment, the participants had to answer a questionnaire about their sound expertise. From their answers, we labeled each participant as "expert" or "naive". We defined as experts those fulfilling the following requirements:

- Being a professional musician, or having a major musical education,
- Being a professional artist, regularly working with sounds (sound installations, performances, etc.),
- Being a professional or semi-professional sound engineer or recording engineer,
- Being a scientist working in the fields of sound perception, acoustics or sound signal processing.

15 expert and 15 naive participants made the experiment. We defined these two groups because of our Hypothesis 5 (relation between the expertise and the type of similarity used to group together sounds, see section 2.2 p. 2.2).

Stimuli The 60 stimuli selected from the results of the experiment 1.1, see Table 4.2 p. 43.

Apparatus The same hardware equipment as in experiment 1.1 was used in experiment 1.2. However, the software used to run the experiment and to implement the graphical interface was Matlab 7.0.4.

Procedure The procedure had three steps.
First, the participants sat alone in a sound attenuated booth in front of a computer display. They were

all given written instructions explaining the sorting task (see Appendix D.). They saw a white screen on which red dots labeled from 1 to 60 were drawn, each dot corresponding to a sound. The labeling was different for each subject. They could hear the sound by double-clicking on a dot. Participants were asked to move the dots in order to group together the sounds. They were allowed to form as many groups as they wished and to put as many sounds in each group as they desired.
After they had made the categories, they had to describe to the experimenter, for each category they had made, the properties, shared by the sounds, that they have used to make this category (see Appendix B. for the verbatim of the instructions).
Finally, they were told that we had identified three strategies usually used to group together sounds. They were given a written description of these strategies, and required to indicate, for each category, if they thought to have used one of these strategies. These described strategies correspond to the three aforementioned types of similarities . In addition two others answers allowing participants to indicate the irrelevance of the proposed strategies are indicated (1.ãcoustical similarities; 2.Ẽvent similarities; 3.S̃emantic similarities; 4.Õther reasons; 5.Ño similarity).

3.2.2 Analyses

Raw results

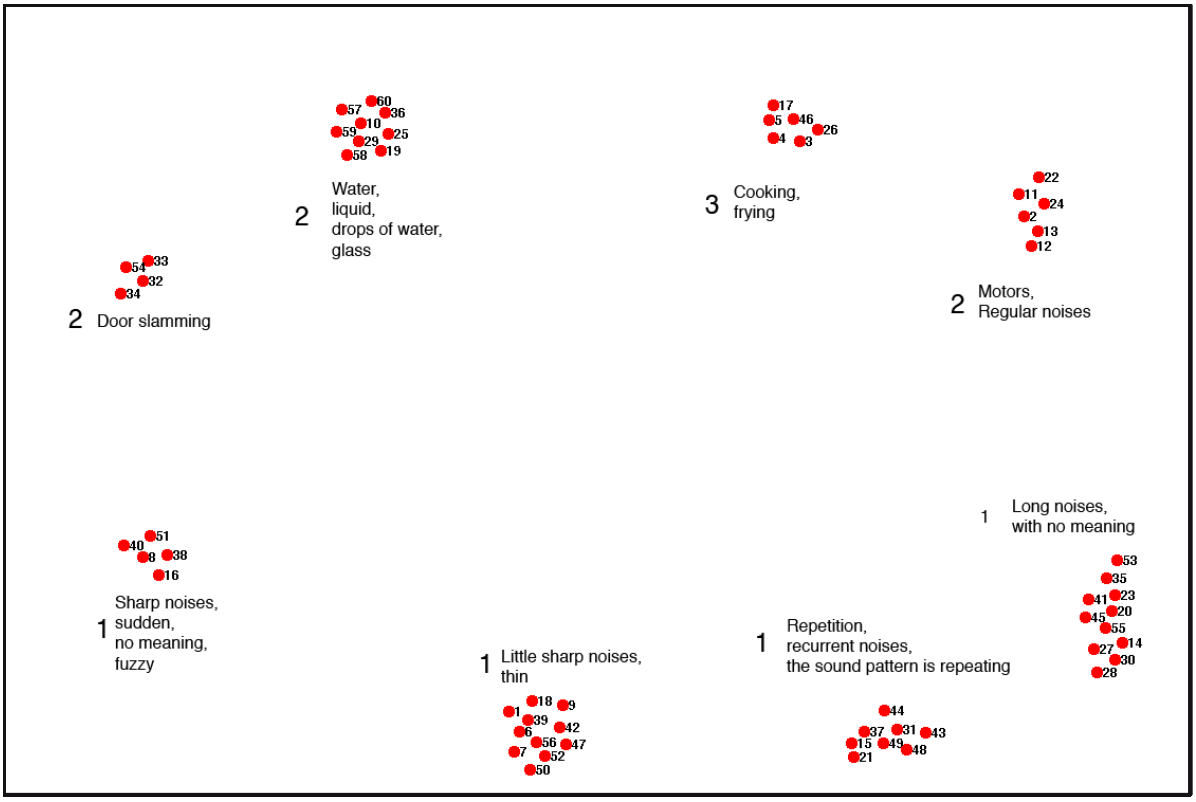

Figure 3.3: Example of of classifications made by a participant. The participant is an "expert" participant. The numbers associated to descriptions indicate the n^o of the answer choose among the 5 propositions.

On average, the participants made 11.5 classes (from 3 to 34). They reported having difficulties to describe the categories. They also reported that the proposed grouping strategies were most of the times relevant to their own strategies.

h!

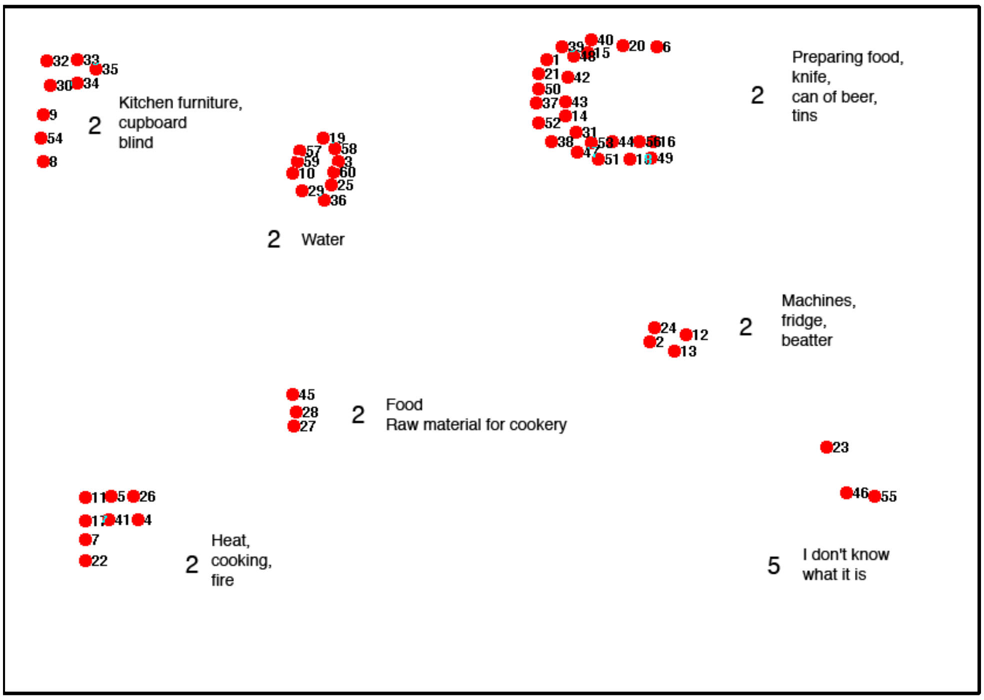

Figure 3.4: Example of of classifications made by a participant. The participant is a "naive" participant. The numbers associated to descriptions indicate the n^o of the answer choose among the 5 propositions.

Figures 3.3 and 3.4 draws examples of classifications made by two participants. Figures 3.3 is a classification made by an expert participant, Figure 3.4 is the classification made by a naive participant. The expert participant made classes of sounds which were acoustically similar (e.g. group of sharp sounds) or made by similar events (e.g. group of door slamming sounds). The naive participant made mainly groups sounds caused by similar physical events.
Besides, this latter example illustrates why we directly asked the participants to indicate which type of similarity they used to group together the sounds, rather than analyzing ourselves the verbalisations. For instance, the descriptions of some categories (e.g. "food") may appear to correspond to semantic similarities of the sounds: the general idea of food, possibly including different materials and different actions. However, the participant choose to code this category with "2" (similarity of physical events). This may indicate that participants were not able to understand the proposed strategies, and prevent us from using their similarity coding. But when we asked this participant why she coded this category as "2", she answered that she was not thinking of the general idea of food, but that she had grouped together sounds in this category, because she had identified for all these sounds a precise texture of the material (that she named "food"), indeed corresponding to similarities of events. Therefore, it appears that analyzing the verbalizations to decide to which type of similarity they correspond would have been a very difficult task and thus would have given fuzzy data. Asking directly the partipants to report themselves their strategy seems conversely a safer method.
On average, participants have grouped together 32.2 % of the sounds according to acoustical similarities, 45 % of the sounds according to the similarities of their physical cause, 12.5 % of the sounds because of semantic similarities, 3.3 % of the sounds for other reasons, and 0.8 % of the sounds because they did not how know to group them. This is coherent with hypothesis H1: participants have used different strategies to make the classes.

3.2.3 Recoding the results

From this experiment, we coded two sets of data. The first set of data (proximity data) codes the partitions of the sounds made by the participants. The second set of data (similarity criteria) codes the criteria used to classify the sounds.

Proximity data

It has been remarked that classification data actually amount in proximity data for large sets of stimuli [4, 9, 11]. The data for each participant consisted of an incidence matrix, i.e. a matrix in which a *one* indicates that the two sounds have been classified together and a *zero* that they have been classed in different groups. The individual incidence matrices (coding the set partitions of each subject) are summed. A co-occurrence matrix is then obtained. The co-occurrence matrix represents how many participants have placed each pair of sounds in the same category. This can be interpreted as a proximity matrix [6]. Moreover, 100 % of the 34220 triplets formed with the 60 sounds follows the triangular inequality. This confirms that the co-occurence matrix is a matrix of distances.

Types of similarity

For each participant, we collected for each sound the criteria used to group it together with other sounds in a category. This results in a 30*60 matrix, consisting in, for each sound/subject combination, the index of the type of similarity used to group this sound together into a category. This is the matrix of similarity criteria.
Since we are interested in assessing the influence of the Hcu values on the types of similarity used to group the sounds, we need to have a regular scale of Hcu. We decided to define a scale dividing the range of Hcu values for our sounds (0-8.2; 8.2 is the maximum of Hcu, each sound belongs to a specific category of actions and objects) into five intervals (0-1.6, 1.6-3.3, 3.3-4.9, 4.9-6.6, 6.6-8.2). This scale was chosen so that the size of the intervals is greater than the standard deviation of the Hcu measures, and that there is more than one sound in each interval of Hcu. Figure 3.5 draws the distribution of sounds in each Hcu interval. The first interval Hcu 1 has the less sounds (five sounds), and the fourth interval Hcu4 has the most sounds: 20.

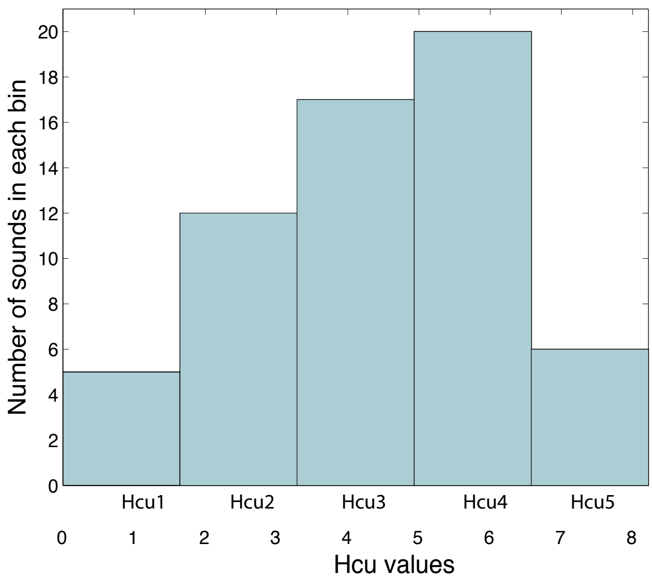

Figure 3.5: Distribution of sounds within the Hcu intervals.

Then, we recoded the results according to the following procedure: for each subject (each line of the matrix of similarity indexes), we count the ratio of sounds in each interval of Hcu which were grouped together according to each of the five types of similarities. We then obtained for each type of similarity a 30*5 matrix (30 subjects * 5 Hcu intervals): matrix of similarity ratio.

3.2.4 Analysis of variance

To analyse the influence of the expertise of the participants, and of the causal uncertainty of the sounds, onto the strategies (the type of similarity) used the participants classify the sounds, we performed five analyses of variance. Indeed, our experiment can be formalized as a one-between one-within repeated measure experiment, with the expertise ($\mathcal{E}$, with 2 modalities) of the participants as the between subject experimental factor, and the Hcu interval ($\mathcal{HCU}$, with five modalities) as the within subject experimental factor. The dependant variable are the similarity ratio for each type of similarity: $\mathcal{S}_{30} < \mathcal{E}_2 > *\mathcal{HCU}_5 \rightarrow C_{acoustical}, C_{event}, C_{semantic}, C_{other}, C_{unknown}$. The five ANOVA tables are reported in Appendix E.

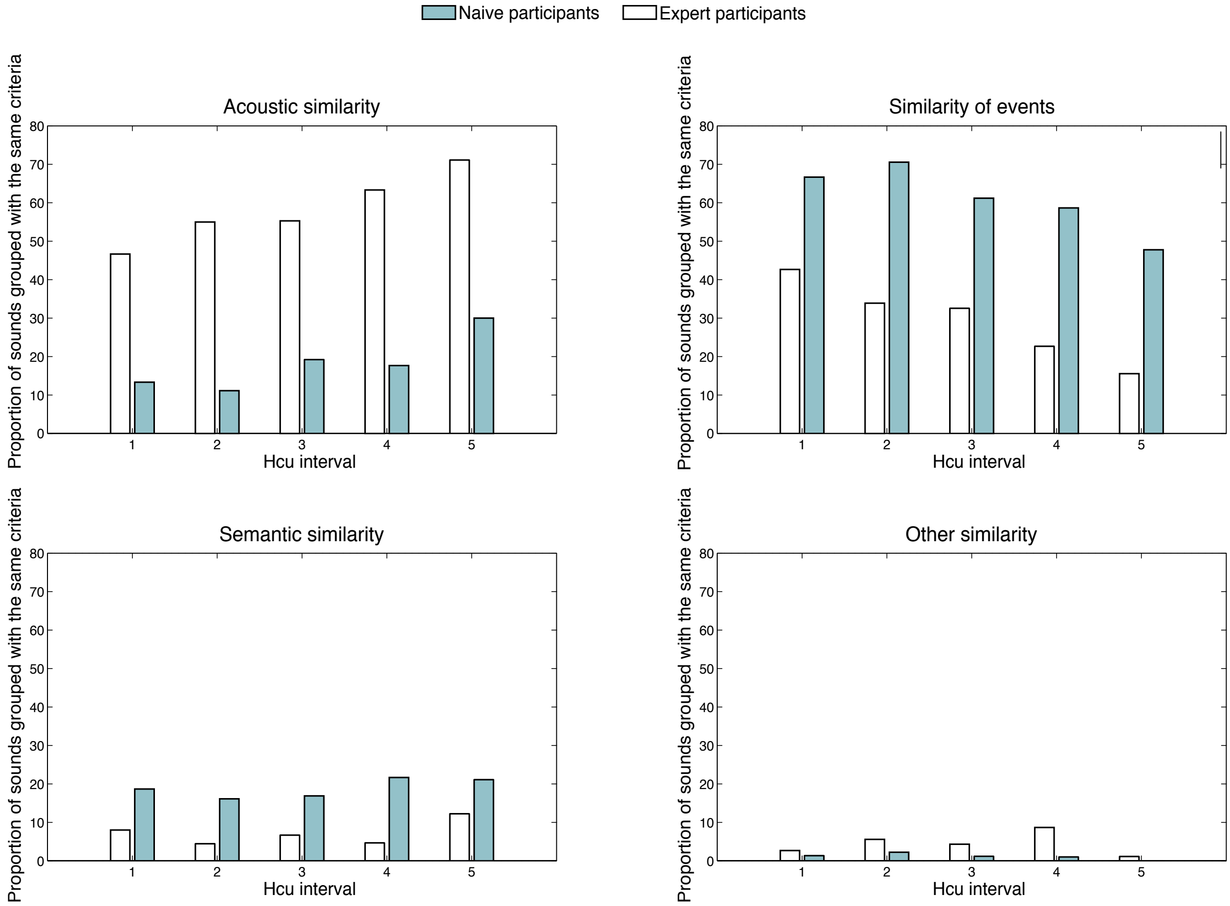

Figure 3.6: Influence of the Hcu interval values and of the expertise of the participants on the similarity used to group together the sounds.

Analyzing the influence of the experimental factors on each similarity ratio Figure 3.6 summarizes the influence of the Hcu interval values and of the expertise of the participants on the ratios of each type of similarity used by the participants to group together the sounds into categories.
Table 4.3 in Appendix E. indicates that both the principal effects of the expertise of the participants, and of the Hcu interval on the ratio of sounds grouped together because of acoustical similarities are highly significant. Examining the upper left panel of Figure 3.6 first shows a clear effect of

the expertise of the participants: expert participants have grouped together sounds because of their *acoustical similarities* for much more sounds than the naive participants. This figure shows as well clearly an effect of the Hcu intervals: within the Hcu intervals corresponding to the higher Hcu values (i.e. sounds with a high causal uncertainty), much more sounds are grouped together because of their acoustical similarities than within the intervals of low Hcu values (i.e. sounds with a weak causal uncertainty). The interaction between these factors is not significant.

Table 4.4 in Appendix E. indicates that the principal effects of the expertise of the participants, and of the Hcu intervals on the ratio of sounds grouped together because of the *similarities of the events* causing the sound are as well highly significant. The upper right panel of Figure 3.6 displays however a reversed pattern: naive participants have used the similarities of the events causing the sounds for grouping much more sounds than expert participants. Much more sounds within the low intervals of Hcu are grouped together because of the similarities of the event than within the high Hcu intervals. Again, the interaction between the two factors is not significant.

Table 4.5 in Appendix E. indicates that none of the experimental factors (expertise, Hcu interval), neither their interaction, has a significant effect on the ratio of sounds grouped together because of their *semantic similarity.* Yet the lower left panel of Figure 3.6 apparently shows an effect of the expertise of the participants (much more naive participants having used this criteria than the expert ones). Carefully examining the results shows that the differences between naive and expert participants are not significant in this case because of the high variability of ratios.

Table 4.6 in Appendix E. shows that only the influence of the Hcu interval on the ratio of sounds grouped together because of *another criteria* is moderately significant. However, as shown by the lower right panel of Figure 3.6, only a few sounds (3.4 % on average for all the sounds) have been grouped using this criteria. It is therefore uneasy to interpret this effect.

Similarly, only 0.8 % of the sounds have been grouped together because the participants *did not know* how to classify them.

Discussion This analysis clearly shows that naive participants and experts participants use different strategies to do the classification: while naive participants spontaneously (they received no particular instruction) group together sounds mainly because they identified them as caused by the same physical event, expert participants spontaneously group together sounds because of their acoustical similarities. This effect is strong. It may be concluded that judging the sounds according to their acoustical properties require having been trained (implicitly or explicitly) to do so. This is coherent with the notice made by several authors [12]: when listeners have to describe a sound, they describe mainly the cause of the sound (when they are able to identify it), and not the sound itself (see the first part of this document). This is also coherent with our hypotheses H1 and H5.

The other major conclusion of this analysis is that the causal uncertainty of the sounds influences a lot the strategy used to group together the sounds. Sounds with a high causal uncertainty, i.e. sounds the cause of which are hard to identify precisely are grouped together according to acoustical similarities much more easily than sounds the cause of which are easy to identify. This is coherent with our hypotheses H2. However, when the sounds are identifiable, naive listeners group them together mainly because of the similarities of the events that they have identified, while expert listeners use both event and acoustical criteria to group them together. Further more, the ANOVA indicates than these two factors are independent. This is slightly different from our hypothesis H3: naive listeners use acoustical criteria to group together the sounds only when they are not able to identify the cause of the sound. When they are able to identify the cause of the sounds, they tend to use mainly the similarities of the identified sound events.

The high variability of the ratios of sounds for which participants have indicated having used a semantic criteria seems to indicate that they did not understand precisely what we meant. Indeed, the definition that we gave was rather fuzzy, and negative (not acoustical, not related to the event, but more "abstract").

However, the low ratio of sounds "other" for which participants have indicated having used another criteria that those that we proposed seem to indicate that these criteria were relevant to their strategy.

3.2.5 "Focused" cluster analysis

In the last sections, we have examined the influence of the causal uncertainty onto the type of similarity used to group together the sounds. In this section, we will provide graphical representations of the proximity data which will allow to unveil the organization of sound events.
Overal, the proximity data, for all the participants and all the sounds, follow the ultrametric inequality for only 74 % of the 34220 triplets that can be formed among 60 sounds. This means these data will hardly be represented by a tree representation. We may assume that this results from the variety of strategies used by the participants.
The analyses of variance have shown that naive participants have in majority group together sounds because of the similarities of the events causing the sounds. On the other hand, expert participants, because of the skill, have elaborated strategies to group together sounds on the basis of their acoustical similarities. Since the organization of sound events properties is more likely to be hierarchical than the organization of acoustic properties (see the review), the idea is therefore to do two separated analyses: one for the expert, and the other for the naive listeners. Thus, we analyzed the structure of the classification on the co-occurence matrices for these two groups of participants corresponding to the "expert group" and the "naive group" presented in paragraph A.
However, since we know now that both Hcu and expertise influence the strategy used by the participants, we can select subsets of sounds corresponding to a specific strategy. Then we focused on sounds with high Hcu ($Hcu4$ and $Hcu5$ Figure 3.5 p. 21) for the expert group, and sounds with low Hcu ($Hcu1$ and $Hcu2$ fig. 3.5 p. 21) for the naive group. This analysis is presented in paragraph B.

Introduction to hierarchical clustering

Cluster analyses allow to represent proximity (or, conversely, dissimilarity) data by a tree representation (a *dendrogram*) as long as the data follow two conditions:

- They follow the triangular inequality: for each triplet of items A, B, C, $d(A,C) \leq d(A,B)+d(BC)$, where $d(A,B)$ is the dissimilarity between item A and B
- They follow the ultrametric inequality: $d(A,C) \leq max(d(A,B), d(BC))$,

Proximity data obtained from sorting experiments always follow the triangular inequality (by definition, they are distances). However, the ultrametric inequality is rarely true for all the triplets that can be made of the proximity data. Therefore, to be able to have a proper tree representation of such data, it is required to approximate the proximity matrix by an ultrametric proximity matrix.
We have used the algorithms provided by Hubert et al. [5]. These algorithms generate the best-fitting ultrametric distances minimizing the least square criterion called L_2norm (see equation 3.5) between distances in proximity matrix and ultrametric distances, see [5] for the detail of the algorithm. These algorithms use a heuristic search strategy using iterative projection to locate the best-fitting ultrametric distance in the L_2norm. This L_2norm corresponds to :

$$L_2 = \sum_{i<j}(p_{ij} - u_{ij})^2 \tag{3.5}$$

where p_{ij} is the input proximity between the objects i and j and u_{ij} is the corresponding ultrametric.
The variance accounted for (VAF) the ultrametric distance matrix representing the proximity matrix is given by :

$$VAF = 1 - \frac{\sum_{i<j}(p_{ij} - u^*_{ij})^2}{\sum_{i<j}(p_{ij} - \bar{p})} \tag{3.6}$$

where $\bar{p}$ is the mean off-diagonal proximity and and u^*_{ij} is the best-fitting ultrametric for a pair of objects i and j. The variation of the VAF is between 0 and 1, 1 for a perfect fitting.
In hierarchical tree like Figures 3.8 and 3.9, the proximity distance between two sounds is represented

by an ultrametric distance within a hierarchical tree. The ultrametric distance between two sounds is represented by the height of the fusion of the higher point where the two sounds are joined within the tree. More the point of fusion is high, more these two sounds are dissimilar.

A) Selecting participants

We focus here on hierarchical clustering analyses of the proximity matrices for "naive" and "expert" groups.

Expert participants 78 % of the triplets formed from the sub-matrix of proximity data for the expert participants follow the ultrametric inequality. The hierarchical clustering representation of classes for expert group is provided in Figure 4.1 p. 50. This representation is not accurate (VAF = 0.7011) and is very hard to explain because of the great variability between expert participants. There are however classes associated to liquid sounds, gas sounds and motor sounds (see Figure 4.1 p. 50). The class concerning solid sounds is more confused and mixes different criteria : acoustical and event similarities.

Naive participants Conversely, 85 % of the triplets formed from the sub-matrix of proximity data for the naive participants follow the ultrametric inequality, and the VAF is better (VAF=0.8502), indicating that the tree representation for these date is more accurate than for all the data. The Figure 3.7 p. 27 explains the structure of the classification. Descriptions for all theses sounds are detailed in appendix B in Table 4.1 p. 41.
We have found mainly 10 classes :

- **Class A** This is the class of liquid sounds, with sub classes of "flow of water" (sounds 99, 101, 98, 47) "drips of liquid" (sounds 17, 97) and "pouring liquid" (sounds 7, 61)
- **Class B** This class is also related to liquid sounds but corresponds to "bubbling liquid" (sounds 49, 81)
- **Class C** These sounds are gas sounds, (sounds 9; 29, 12, 72) including a "spray" sound. Sound 8 is dissimilar from the other ones and seems to be between liquid and gas sounds. It is a sound of "boiling water"
- **Class D** The class is related to electric sounds, like motor sounds (sounds 6, 19, 20), and other electric sounds like refrigerator sound (sound 44), micro wave sound (sound 18) and electronic sound like "bip bip" from micro wave (sound 16).
- **Class E** All these sound are "hitting" sounds like "ice cube in a glass" (sound 1), "shocked glasses" (sound 10) and "spoon in a cup" (sound 77), "beating eggs" (sound 79) and "put a cup on table" (sound 30).
- **Class F** This is the class of friction sounds, sounds of "cutting" or "peeling" (sounds, 23, 40, 84, 83), "sharpening" (sounds 22,91), "(un)screwing" (sound 56, 64).
- **Class G** These two sounds are like "open a can" (sound 67) or "switch on" (sound 95). These are friction sounds, but different from class F
- **Class H** This class is related to "crumpling / crushing" sounds (sounds 51, 52, 94, 68, 82, 85, 89)
- **Class I** These sounds are "rubbing" sounds like "closing door" (sounds 92, 57, 58) "closing a chair" (sound 54) and "closing a drawer"
- **Class J** These two sounds are "scraping" sounds (sounds 24, 88).

Hence, this representation indicates a distinction between liquid sounds (classes A, B), gas sounds (class C), electric sounds (class D) and other big classes of solid sounds (classes E, F, G, H, I, J). This distinction is in agreement with Gaver's taxonomy (Deliverable 4.1 part one). Inside each class there are also major distinctions:

- For the liquid we observed three classes related to "flowing", "pouring" "dripping" sounds.
- The gas category is more related to "flow of gas" due to the selected sounds. It was hard to find different types of gas sounds from our sound libraries.

- The category of solid sounds is the most prominent, because of the great variety of sound sources. Inside this category, we found the distinction between impact sounds (classes E, H) and friction sounds (classes F, G, I, J).

These results show that "naive" participants with no expertise in sound can group sounds according to the representation of the physical production of the sounds (event similarity). We found the different categories like Unvirona's taxonomy based on physical phenomena. This result has confirmed our Hypothesis 4 on the hierarchical organization of sounds events.

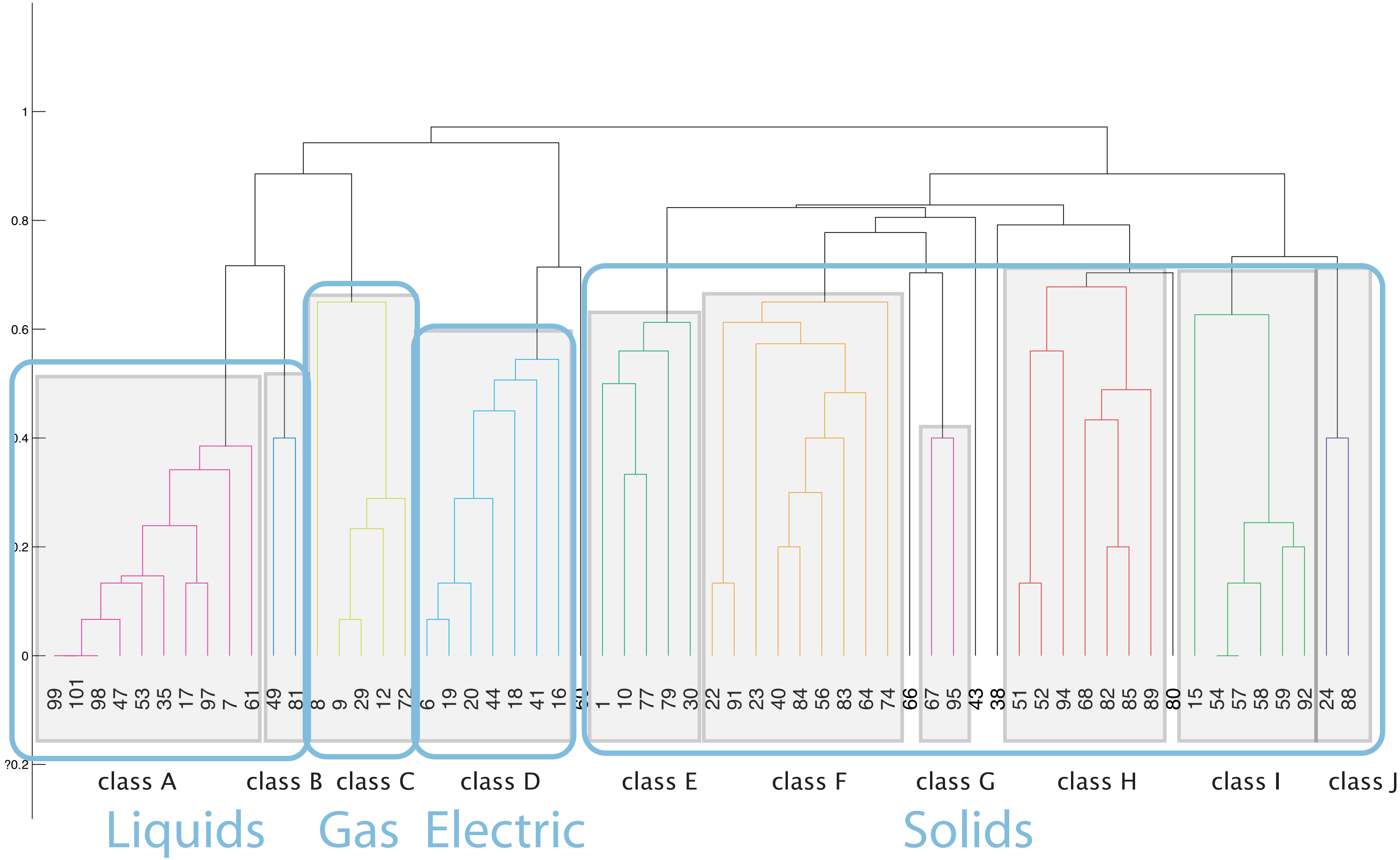

Figure 3.7: Cluster analysis of the proximity matrix of the naive group for the 60 sounds using Hubert et al. algorithm [5]

B) Selecting participants and sounds

Since the causal uncertainty has also an influence on the classification strategy, we are also able to focus the cluster analyses only on sounds with a certain causal uncertainty. To represent the organization of acoustical properties, we have therefore focused on sounds with high causal uncertainty ($Hcu4-5$) for the expert participants, and on sounds with a low causal uncertainty ($Hcu1-2$) for naive participants. Hence, we have constructed two sub co-occurence matrices corresponding to the 26 sounds within intervals $Hcu4$ and $Hcu5$ for expert group and 17 sounds with intervals $Hcu1$ and $Hcu2$ for naive group.
The VAF for the ultrametric representation of proximity matrix of the expert group vs $Hcu4-5$ is 0.6741 and for the naive group vs $Hcu1-2$ is 0.9683. The representation of the data of the expert group is less accurate than the naive group with a strong VAF. The ultrametric representation of proximity matrix of the expert group $Hcu4-5$ is thus not accurate and produced an unstable hierarchical tree. In this case the classes formed by this structure should be analyzed with circumspection.

Classification of the naive group / Hcu1-2 The hierarchical representation of the proximity matrix for sounds within intervals of $Hcu1-2$ is plotted on the Figure 3.8. We have observed 7 classes. These classes are highly contrasted due to a high VAF (VAF=0.9683). Descriptions for all theses sounds are detailed in appendix B in Table 4.1 p. 41.

Class A All these sounds are liquid sounds as "sink flow", "hands in water"

Class B The sound is produced by "pouring cereals"

Class C These two sounds are "gas combustion" sounds produced by "gas" and "match"

Class D These three sounds correspond to sounds of "knife" : "cutting bread", "removing knife from case", "sharpening knife"

Class E These sounds are produced by a shock of similar materials : "porcelain" and "glass",

Class F The two sounds correspond to electric sound sources : "bip bip micro wave" and "refrigerator",

Class G These three sounds are "closing door" sounds from "cupboard door" and "microwave door".

All the classes are directly related to two types of similarity, physical event similarity (Classes A, C, E and F) and semantic similarity (classes D and G). These results are coherent with the ANOVA analyses.

Classification of the expert group / Hcu4-5 The hierarchical representation of the proximity matrix for sounds within intervals of $Hcu4-5$ is plotted on the Figure 3.9. We have observed 7 main classes. Some sounds do not belong to these classes because of their high height of fusion in the ultrametric tree. The VAF is not accurate for these representation (VAF=0.6741), that why we observed low contrasted classes with relatively high height of fusion in the ultrametric tree. In order to explain the classes we used the verbalizations of the participants.

Class A These sounds are "repetitive" sounds with a "pulsation" as participants said.

Class B These sounds are "short" sounds and "crackling sounds", "hissing" sounds,

Class C These sounds are "discontinuous sounds", "short sounds", "crackling sounds", "crumpling sounds"

Class D These sounds are "small", "short", "dry",

Class E These sounds are "short", "metallic", "impact" sounds,

Class F These sounds are "crackling", "crumpling" sounds,

Class G These sounds are "loud" "continuous" "cyclic/regular".

The classes are more difficult to explain but the similarity involved for class formation is related to complex acoustical criteria like morphological properties or timbre properties of sounds, typically like class A, class G. We need further acoustical analyses in order to understand which acoustical properties have been used to form classes.

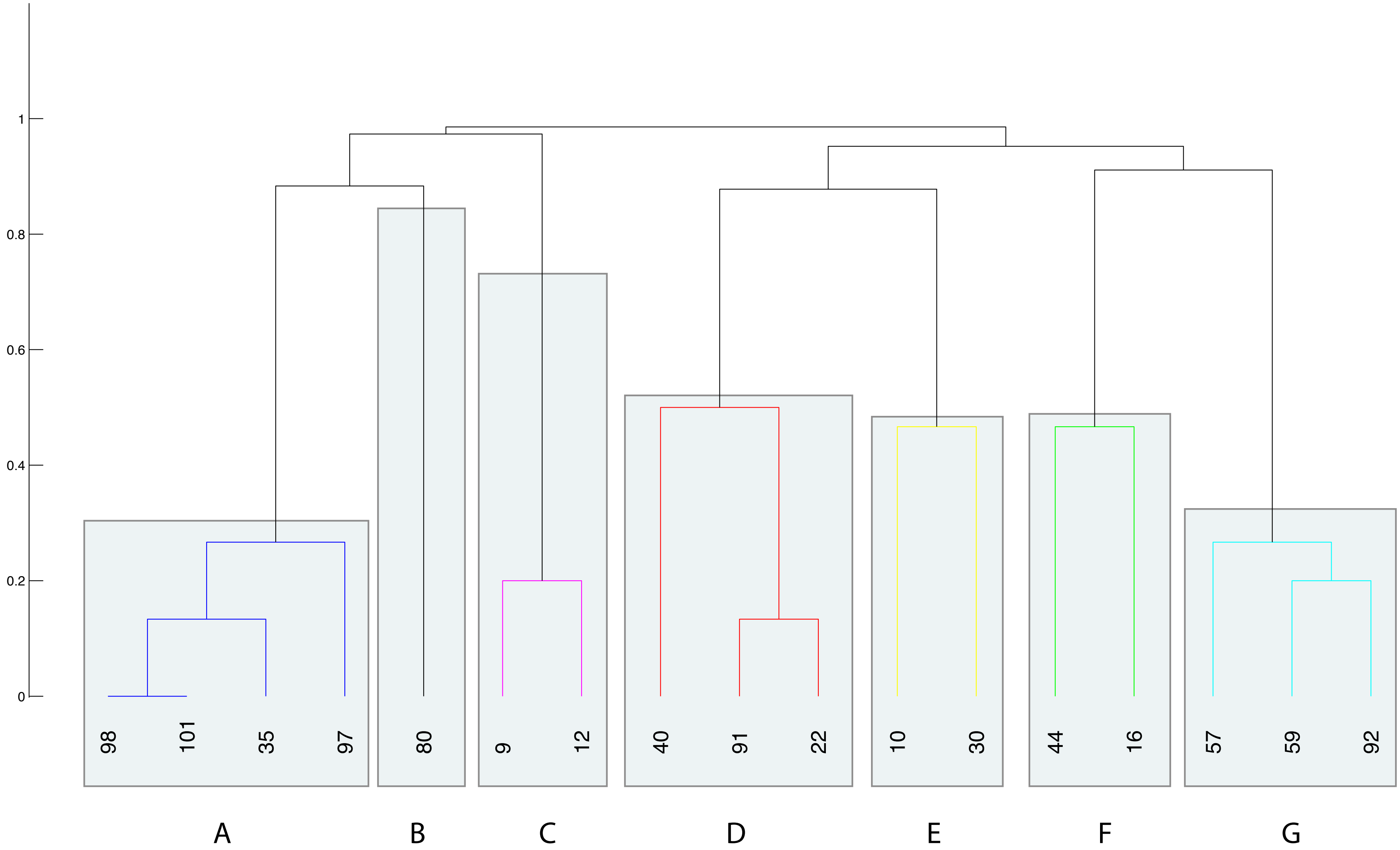

Figure 3.8: Cluster analysis of the proximity matrix of the naive group with sounds within Hcu1-2 using Hubert et al. algorithm [5]

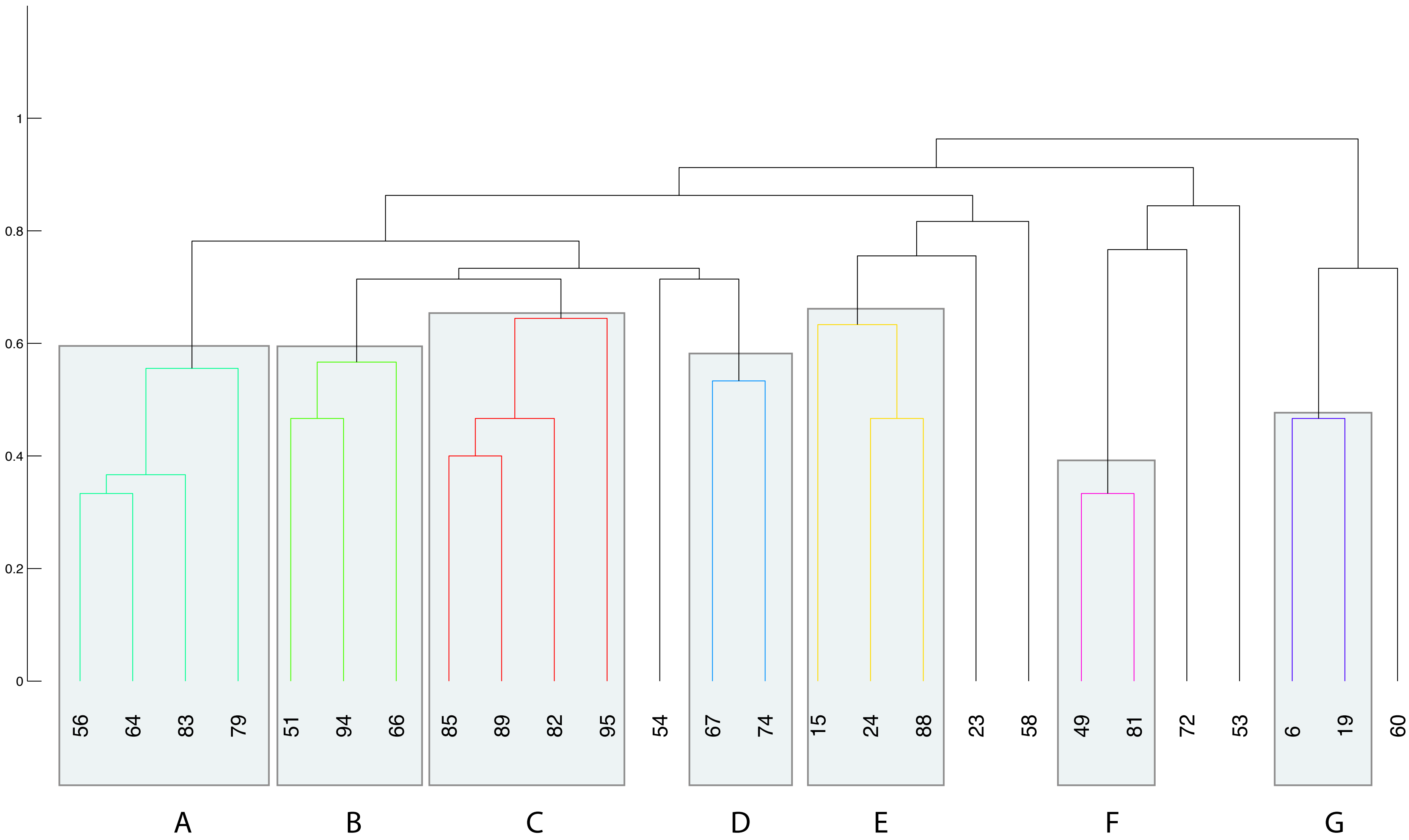

Figure 3.9: Cluster analysis of the proximity matrix of the expert group with sounds within Hcu4-5 using Hubert et al. algorithm [5]

Discussion The cluster analyses reported in this section lead to several conclusions.
Firstly, they show that, overall, the sorting data for all the sounds and all the participants are only poorly represented by a hierarchical representation. This is not surprising, since the Anova results had already demonstrated that participants had used a number of different strategies to group together the sounds, using different types of similarities and properties. But this conclusion has to draw our attention to the fact that, to explore the organization of the sound event properties, we must *select* the sounds and the participants. Indeed, the "focused" analyses reported in this paragraph (based on subsets of the data focused on categories of sounds and expertise) actually simulate what would have happend if we would have studied only the descriptions of naive participants and well identified sounds on the one hand, and only expert participants and poorly identified sounds on the other hand. In this latter case, the representation illustrates the organization of the perceived acoustical properties while on the former case the representation illustrates the organization of the properties of the events causing the sounds. For the organization based on different acoustical properties, we have to find a specific analysis to represent this classification.
Secondly the cluster analysis based only on the data of the naive participants have allowed to confirm the hierarchical organization of the sound events (and is therefore coherent with our hypothesis H4). This organization confirms also the results of Guyot (distinction between "natural" and "non natural" modes of production, for example class D in Figure 3.7), and Gaver's intuition. Figure 3.10 summarizes these results. This schema has however not to be thought of as being exhaustive: since, we have used a large variety of different sounds (from electronic beep to water dripping), we do not have used enough sounds within each class to explore the details of the classification.

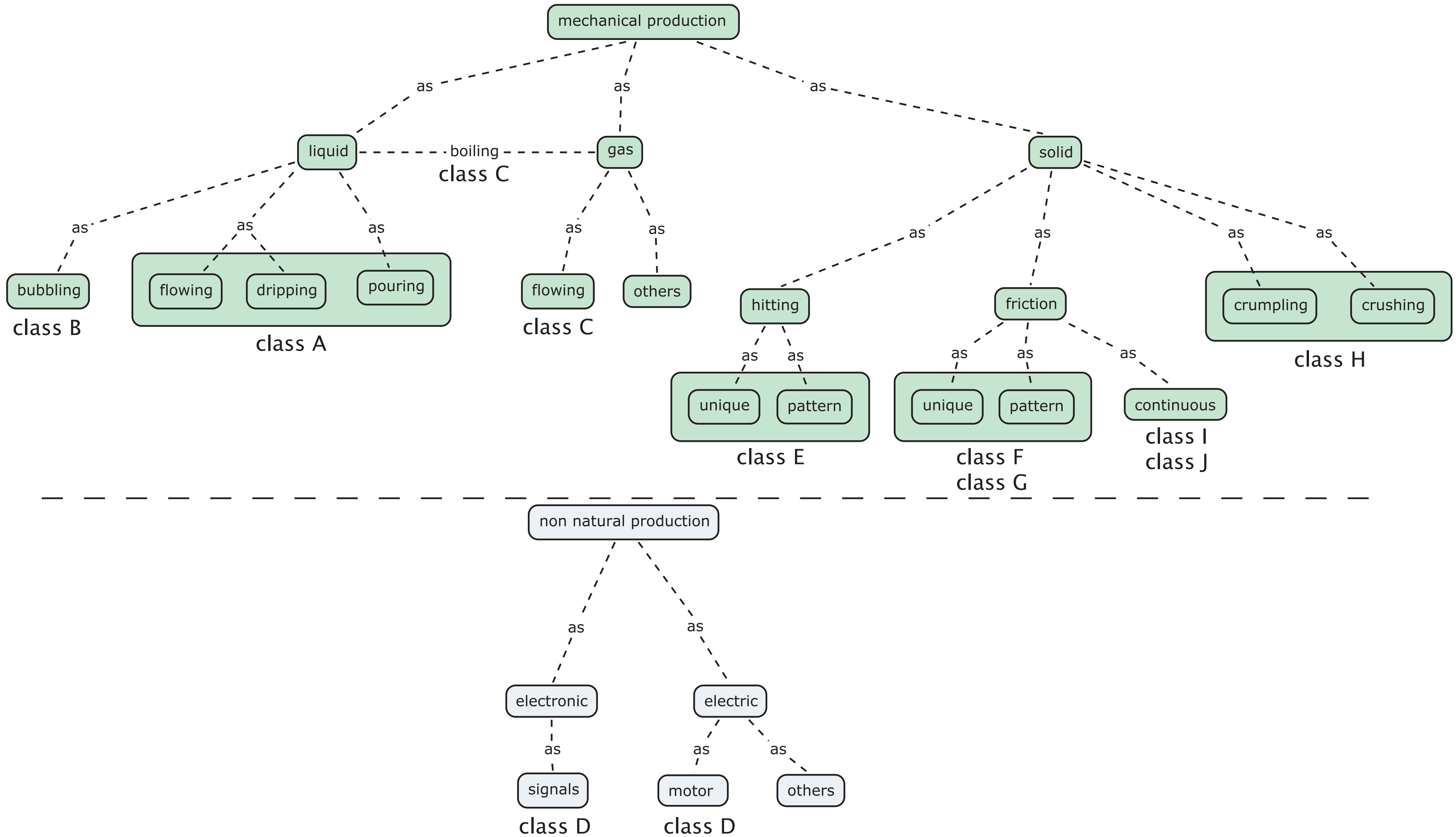

Figure 3.10: Summary of the organization of sound events provided by the "focused" cluster analyses.

4 General discussion

4.1 Back to the hypotheses: summary of the main results

In the first part of this document we have examined the results of four classification experiments reported in the literature. This has lead us to formulate five different hypotheses:

- H1. When listeners are required to group together sounds "freely", they may group together the sounds according to different kinds of similarities (acoustical, event, semantic),
- H2. When the sounds are not identifiable, listeners can only group together sounds which are acoustically similar,
- H3. When the sounds are identified, listeners can choose to group together the similar sound events, or to group together sounds because of property related to their knowledge of the source
- H4. The categories of sound events are hierarchically organized
- H5. Listeners with an expertise in analysing sounds will tend to group together sounds using acoustical similarities more often than naive listeners.

We have then made two experiments, aiming at testing these hypotheses.
The results of experiment 1.2 are coherent with H1. Participants have reported having used different strategies to classify the proposed set of sounds. They were able to label their own strategies by using the items of a list describing the main strategies that we had identified in the available literature. In majority, they have reported strategies corresponding to different types of similarities used to group together the sounds: acoustical similarity, similarity of the event causing the sound, and, to a lesser extent, semantic similarity.
The results of experiment 1.2 go along with the overall idea expressed by H2 and H3, although modulating their detailed formulation. Indeed, experiment 1.2 has shown that sounds with a high causal uncertainty tend to be grouped together mainly on the basis of acoustic similarities, while sounds with a low causal uncertainty tend to be mainly grouped together on the basis of similarities of event causing the sound. Moreover, there seems to be a direct numerical relationship between the causal uncertainty and the proportion of sounds grouped together on the basis of these factors. However, even when the sounds have a very high uncertainty (i.e. are almost unidentifiable), some participants will however try to group them together on the basis of semantic or event similarities.
Indeed, experiments 1.1 and 1.2 have shown a very strong effect of the expertise of the participants, coherent with H5: participants that we identified as "expert" tend to group together sounds on the basis of the acoustical properties of the sounds, while participants that we identified as "naive" tend to base their classification on the properties of the sound events.
Analyzing separately the data of expert and naive participants show that the sorting data of the former can not be represented properly by a hierarchical representation, while the sorting data of the latter fit well with such a representation. Furthermore, the tree representation of these data demonstrate that the properties of sound events are hierarchically organized, following the proposal of Guyot (and thus Gaver), and is therefore in agreement with H4.

4.2 Discussing the results

Classifying sounds: different strategies, different similarities The intuition gained from the review of published experimental "free classification" of everyday sounds is confirmed by our experiments: there are a lot of possibilities available to a listener to classify a set of sounds. When listeners are not provided with any hint, they actually have first to decide on which basis they will do the classification

(this latter remark has also been suggested by the participants' comments). Thus "free" classification experiment implies gathering the results of possibly very different strategies. Different participants will use different strategies, and a participant may also use different strategies to group together different sounds.
The different strategies that we had defined from the review of the literature, and that we proposed to the participants of our experiments, appeared relevant to describe the strategies of the listeners. Yet it is not very clear whether "semantic" similarity was well understood by the participants. It appears however a very clear distinction between strategies based on the acoustical properties, and strategies based on the properties of the physical event causing the sound.

Causal uncertainty of sounds One of the factors that we suspected of having a great influence on the strategy used to group together sounds, was the ease with which the sounds can be identified. The experiments reported in this document clearly show a relationship between the measure of causal uncertainty (borrowed from the work of Ballas [1]), and the type of similarity used to group together the sounds. The analyses of variance of the results of Experiment 1.2 (classification) have indeed shown that the causal uncertainty influences very clearly how listeners classify the sounds: sounds with high values of Hcu, i.e. sounds with many possible are grouped together preferentially because of acoustical similarities. In comparison, sounds with low values of Hcu, i.e. sounds with few possible causes, are grouped together using similarities of the events causing the sounds, or, to a lesser extent, semantic similarity (similarity based on the meanings attributed to the sounds).
The measure of Hcu requires to analyze the verbalizations provided by the listeners, and thus requires a huge amount of time and effort, to define very precisely how to analyze these verbalizations, and to do the analysis. By having three judges doing independently the analysis, we were able to assess the reproducibility (and thus the reliability) of this measure. The reliability of this measure, although not fully accurate, was nevertheless sufficient enough to pursue the analyses, and to demonstrate interesting conclusions. Particularly, it can be concluded that it is required to work with identified sounds to explore the perceptual organization of sound events.
But the time required to get an usable measure of Hcu may prevent experimenters from using this measure systematically to select the sounds. It may be thereby valuable to provide an easiest method to assess the causal uncertainty of the sounds. For instance, asking the listeners how well "they can form a mental image of the cause of the sounds" may be an idea to investigate. This task was cited by Balls as a factor influencing the identification of the sounds, and it requires little use of the language, neither a thorough analysis of verbal descriptions.

Expertise Experiment 1.2 shows also very clearly that the expertise of the participants influences greatly how they form the categories. This effect is very strong, and moreover, independent from the causal uncertainty of the sounds. Expertise do no interact with the causal uncertainty, but is rather added to the effect of the Hcu. This result is very important since, to our knowledge, the results provided in the literature were obtained without selecting the participants.
Thus on the one hand, expert participants use their skill to group together the sounds: they use their analysis of the acoustical properties of the sounds (what we defined in the first part of the deliverable *musical listening*). On the other hand, naive participants classify the causes of the sounds (when able to identify them): this is what we defined in the first part of this deliverable as *everyday listening*. Classes for the expert participants are related to complex compounded acoustic similarities, like morphological and timbre properties. However, expert participants do not use necessary the same acoustical properties.

Organization of sound events: first bricks of a model of classification Focusing on only the data of the naive participants have allowed us to explore the organization of the perceived properties of the sound events, and to provide the skeleton for a model of classification of everyday sounds. We have found classes of solid, liquid, gas and electric sounds, mirroring the different types of physical production of these sounds. Within these classes has also appeared a hierarchical organization. However, because we are studying a large variety of different sounds, we have only a few sounds in each category, and we are therefore not able to get into more details of the internal organization of each category.

CLOSED project These different results are important for the continuation of our project. Compared to the classification of physical phenomena provided by Univerona (see Deliverable 4.1 part one), our classification of perceived sound events shows that we find a common structure. The next step is to find the relations inside each class of sound events and the relation between these categories and the basic elements of sound interaction introduced in Deliverable 3.1.

Appendix A
Verbatim of the experimental instructions provided to the participants in experiment 1.1

Adjustment of ecological sound levels

You will hear sounds recorded in different kitchens, these sounds have been recorded at different levels. These levels do not correspond necessarily to the real levels of sounds in a realistic environment. Your task is to adjust the levels of the sounds corresponding to how you can perceive them in your kitchen if you stay approximatively at one meter of the source. During the experiment, each sound is described with a short sentence in order to identify it.
First, you will listen to all the sounds, in order to get familiarized with them. During the second part of the experiment, you will hear each time a pair of sounds composed by a reference sound and the sound to be adjusted in level. You will adjust the level of the second sound of the pair as you could hear it in your kitchen in comparison to the first sound, the reference. To do so, you will move a cursor that modify the level of the second sound. Each time you move the cursor, the sound reference and the second sound are played sequentially.

Experiment 1.1 : causal uncertainty

You will listen to sounds recorded in different kitchens. These sounds are played at different levels as you can hear them in your kitchen. We ask you to simply indicate what is for you, the cause of the sound, using this form :

- A "noun" and a "verb"
  For example, if you hear "a car accelerating in a street", you may describe the cause of the sound like a combination of the noun " car " and the verb "accelerate".
- "car" "accelerate"

In any case, you should use simple descriptions and not use metaphoric descriptions, indicating at each time : NOUN and VERB
The experiment is divided in three different steps :

1) You will train yourself to use the computer interface with few sounds
2) You will listen to all the sounds
3) For each sound, you will indicate the cause of the sound

Instructions for the coding of the verbalizations

You task is to analyze the descriptions formed by a "noun" (e.g. "balloon", "plastic-balloon") and a "verb" (ex. "hit", "play") for example : "car driven". There are 96 items and for each item, 27 verbalizations. The 96 items with for each item their 27 verbalizations are presented in a table Excel with a column "noun" and a column "verb" like this example:

Item 89	noun	verb
description 1	balloon	hit
description 2	ball	hit
description 3	plastic balloon	play
...	...	...
description 27	boxing bag	hit

For each of the 96 items, you have to form groups of objects and groups of similar actions through the 27 descriptions. You can do as many groups of objets and actions you want if necessary. For example :

			action				object	
Item 89	**noun**	**verb**	Hit	Play			Balloon	Bag
description 1	balloon	hit	1				1	
description 2	ball	hit	1				1	
description 3	plastic balloon	play		1			1	
...	...	...						
description 27	boxing bag	hit	1					1

You will have to write the number "one" when the descriptions belong to the same group of similar actions. In a same way, you will have to write the number "one" when the descriptions belong to the same group of similar objects.
Your task is to analyze what is the action and what is the object or the subject of the action that are described. After, you will have to group on the one hand the verbalizations describing identical actions, and on the other hand identical objects.
Sometimes, verbal groups or noun groups are composed by several words. These words can help you to understand what object is suggested (for example "rail / railways"). In the contrary case (for example "file / bulb"), consider the first word.
We have defined different rules for grouping actions and objects.
Warning : you will have to determine the actions and objects using the whole description and not only based on noun or verb.

Instructions for the objects

EQUIVALENCE :

- If two synonymies are used, objects are considered as equivalent. For example, we consider "stone" and "peble" as equivalent.
- If for two quoted objects, one is a particular type of the second, we consider both as equivalent. For example, "cellular phone" is a particular type of "phone", we consider "cellular phone" and "phone" as equivalent or "car" and "cabriolet" as equivalent.
- If for two quoted objects, one is the part of the other, we consider both as equivalent. For example, "pencil" and "lead" or "television" and "screen" are respectively considered as equivalent because "lead" is a part of a "pencil" and "screen" is part of "television".
- If a description quotes the material of an object quoted in other description, objects of these descriptions are consider as equivalent. For example, "wood" and "bough" are considered as equivalent.

For the objects : DIFFERENCE

- We consider as different objects, objects in their usual sense. For example, a "phone" and a "computer" are different objects.
- If a verbalisation indicates a category of objects, and an other description an object of this category, these two descriptions are considered as different. For example, "computer" and "computer

equipement" are considered as different, "computer equipement" is a too broad category and can indicate an other thing than "computer".

- If several verbalizations describe too large categories of objects (even if they share the same words), these verbalisations are considered as different. For example, "vehicle" and "transport" should be considered as different objects. And if two verbalizations used the same expression "transport", these verbalizations should be considered as different objects.
- If an object quoted in a description is a part of several different objects, it should be consider as a different object. For example, "keys" and "computer" and "phone" should be consider as three different objects.

For example, among these six verbalizations, two are considered as equivalent, "car" and "cabriolet", others are considered as describing different objets, even if the same expression "transport" is used two times.

Item 65						
description 1	transport	1				
description 2	car		1			
description 3	motorcycle			1		
description 4	transport				1	
description 5	cabriolet		1			
description 6	vehicle					1
	Total	1	2	1	1	1

Instructions for the actions

We consider an "action" as a physical action, i.e. the process of an agent on an object. For example, the action "to stick something", as in the expression "to stick a poster".

EQUIVALENCE

- If two synonymies are used, the actions are considered as equivalent.
- If two verbalizations describe, one an action and the other an action that is a manner of doing this action, these two actions are considered as equivalent. For example, "to howl" and "to shout" are considered as equivalent, because "to howl" is a manner of "to shout".
- If two verbalizations describe, one an action that can be decomposed in a sequence of actions and the other an action of the sequence, these actions can be consider as equivalent. For example "to phone" and "to compose a telephone number" are considered as equivalent.
- If an action decribed by a verbalization is a consequence of an other action, both actions are considered as equivalent. For example, "to polish a mirror" and "to rub a rag against a mirror" are verbalizations considered as equivalent actions.
- Verb describing a common use of an object, without ambiguity, should be consider as describing an usual action of this object. For example, "to phone" and "using a phone" are consider as equivalent.

DIFFERENCE

- If a verbalization describes an action and an other verbalization a category of actions, or an abstract function, these actions describe by these verbalizations are considered as different. For example "to make sport" and "to hit a ball" are considered as different, like "to make noise" and "hit a box".
- If two verbalizations describe too large categories of actions, these two verbalizations should be considered as describing different actions, even is the verbalizations are the same. For example "to make sport" and "to have a sports activity" should be consider as different actions.

Remark : For nouns sometimes "nnn" is written and for verbs "vvv", this notation indicates that a person didn't know how to answer. You should consider a group of actions or a group of objects for each "nnn" or "vvv" written for an item.

Appendix B: Description of the corpus

Column 1 is the index of the sound used during the experiments. Column 2 called ”Source” is the source of the sounds. The two first letters correspond to the sound library, HE for Audio Hollywood Edge Edition I,II and III, SI for Audio Sound Ideas 6000, BB for Data Blue Box Audio Wav, SS for Audio SoundScan V2 Vol.61 SFX ToolBox. The first number corresponds to the CD number and for BB and SS **:** indicates the number of folders and subfolders of the CD, the second number corresponds to the CD tack. Columns 3 and 4 are the descriptions of the sound, column 5 is the mean of coefficient of ecological adjustment and SD the standard deviation.

Description of the corpus of sounds					
Sound	**Source**	**Description (English)**	**Description (French)**	**Mean of the EL**	**SD of the EL**
1	HE, 16, 2	ice cubes in a glass without water	glaçons dans un verre sans eau	0,325	0,139
2	HE, 25, 9	air conditioning	air conditionné	0,192	0,145
3	HE, 4, 11 *	drops in water	gouttes à gouttes dans de l'eau	0,418	0,207
4	HE, 5, 24 *	boiling water	eau qui bouillonne	0,167	0,067
5	HE, 25, 27 *	closing a dishwasher door	porte d'un lave-vaisselle qui se referme	0,390	0,251
6	HE, 25, 28 *	dishwasher on	Lave-vaisselle en marche	0,257	0,118
7	HE, 25, 30 *	coffee maker with filter on	machine à café avec filtre en marche	0,212	0,092
8	HE, 25, 31 *	water boiling in a pan	eau qui bout dans une casserole	0,173	0,086
9	HE, 25, 32 *	gas open and furnace on	mise en route du gaz d'un four et allumage	0,317	0,165
10	HE, 16, 32 *	champagn cup shocked	verre de champagne que l'on a cogné	0,156	0,055
11	HE, 25, 33 *	furnace on, hot thermostat	allumage du four à gaz à un thermostat important	0,163	0,038
12	HE, 16, 35 *	striking and igniting a match	grattage et allumage d'une allumette	0,185	0,089
13	HE, 25, 35 *	opening and closing a furnace	ouverture et fermeture d'un four/grille	0,475	0,127
14	HE, 25, 37 *	lowering the toaster compartment	grille-pain, baisse le compartiment	0,237	0,123
15	HE, 25, 37 *	ejection of the toaster comportment	éjection du compartiment d'un grille-pain	0,453	0,115
16	HE, 25, 39 *	bip bip micro wave	bib bip micro-onde	0,479	0,206
17	HE, 4, 38 *	agitating hands in water	mains qui agitent un eau	0,421	0,115
18	HE, 25, 39 *	micro wave on	mise en route d'une micro-onde	0,332	0,080
19	HE, 25, 42 *	food processor	robot ménager	0,587	0,233
20	HE, 25, 43 *	mixer on	mixeur en route	0,658	0,246
21	HE, 25, 45 *	electric press citrus fruits	presse agrume électrique	0,414	0,111
22	HE, 25, 48 *	knife remove from his case	couteau que l'on sort d'un étui	0,240	0,174
23	HE, 25, 48 *	cuting foods with a knife	découpage aliment avec un couteau	0,223	0,121
24	HE, 25, 50 *	scraping a metal pan	raclement casserole en métal	0,275	0,050
25	HE, 25, 51 *	closing a refrigerator door	fermeture d'une porte d'un réfrigérateur	0,302	0,112
26	HE, 16, 60 *	pop up from a toaster	éjection du grille-pain	0,079	0,023
27	HE, 16, 61 *	closing a refrigerator door	fermeture d'une porte d'un réfrigérateur	0,189	0,076
28	HE, 16, 62 *	compressor noise of a refrigerator	bruit de compresseur d'un réfrigérateur	0,110	0,052
29	HE, 16, 66 *	gas open of a furnace	ouverture du gaz d'un four	0,255	0,170
30	HE, 25, 70 *	puting a bowl on a table	pose un bol sur une table	0,553	0,163
31	HE, 25, 70 *	puting a bowl on a table	pose un bol sur une table	0,256	0,110
32	HE, 25, 71 *	closing a door cupboard	fermeture d'une porte d'un placard	0,379	0,093
33	HE, 25, 72 *	closing a door cupboard	fermeture d'une porte d'un placard	0,430	0,117
34	HE, 16, 88 *	turning on a faucet	ouverture d'un robinet	0,781	0,210
35	HE, 25, 89 *	sink to empty	évier qui se vide	0,327	0,194
36	SS, 12:5, 1-5	pouring wine into a glass	versement de vin dans un verre	0,246	0,100
37	SS, 12:1, 1-2	bottle shocked	bouteille que l'on cogne	0,433	0,190
38	SS, 12:1, 1-7	bottle top	bouchon d'une bouteille	0,090	0,039
39	SS, 11:4, 1-2	puting bowl on a saucer	bol que l'on pose sur une soucoupe	0,332	0,155
40	SS, 10:2, 1-2	cutting bread	pain que l'on découpe	0,122	0,073
41	BB, 10:1:3, 2-1 *	coffe maker is whistling	cafetière qui siffle	0,407	0,108
42	BB, 10:1:3, 3-1 *	coffee maker with filter on	machine à café avec filtre en marche	0,418	0,155
43	SS, 12:9, 1-1	removing a cork stopper	bouchon en liège que l'on enlève	0,149	0,096
44	SS, 10:1, 1-1 *	refrigerator noise	bruit de réfrigérateur	0,193	0,103
45	SS, 11:2, 1-15	dish noise	bruit de vaisselle	0,517	0,119
46	SS, 10:3, 4-1	removing a metal top from a kettle	couvercle en métal que l'on enlève d'une bouilloire	0,317	0,122
47	SS, 10:3, 4-2	pouring water into a metal kettle	eau versé dans une bouilloire en métal	0,239	0,131
48	SS, 10:3, 4-4	closing a kettle	fermeture d'une bouilloire	0,372	0,172
49	BB, 10:1:3, 8-1 *	cooking noise with fat	bruit de cuisson avec de la graisse	0,186	0,133
50	BB, 10:1:3, 9-1 *	sound of end of cooking, micro wave	son de fin de cuisson d'un micro-onde	0,465	0,240
51	SI, 6010, 13-4	crushing a paper bag	sac en papier que l'on écrase	0,315	0,220
52	SI, 6010, 14-2	crumpling a plastic bag	sac en plastique que l'on froisse	0,381	0,159
53	SI, 6010, 41-1	big bubble inside a metal kettle	grosse bulle dans une bouilloire en métal	0,268	0,139
54	SI, 6010, 82-7	shuting a wood chair	chaise en bois que l'on plie	0,770	0,166
55	SI, 6015, 83-2	shaking water in a basin	eau agitée dans cuve métal	0,307	0,167
56	SI, 6015, 88-1	unscrewing a stopper	bouchon que l'on devisse	0,099	0,054
57	SI, 6018, 10-2	closing a door cupboard	fermeture d'une porte d'un placard	0,460	0,166
58	SI, 6018, 30-3	closing a door	fermeture d'une porte	0,553	0,256
59	SI, 6018, 34-5	open a drawer with castors	ouverture d'un tiroir monté sur une glissière	0,385	0,153
60	SI, 6020, 2-2	unrolling a blind	store que l'on déroule	0,378	0,144

Continued on next Page...

Sound	Source	Description (English)	Description (French)	Mean of the EL	SD of the EL
61	SI, 6020, 6-2 *	pouring a drink into a glass	versement d'une boisson dans un verre	0,207	0,133
62	SI, 6020, 9-2	screwing the bottle top	couvercle d'une bouteille que l'on visse	0,299	0,186
63	SI, 6020, 10-4	taking off the bottle top	enlève le bouchon d'une bouteille	0,304	0,093
64	SI, 6020, 12-2	screwing a bottle stopper	visse un bouchon sur une bouteille	0,124	0,032
65	SI, 6020, 14-4 *	several sprays from an atomizer	plusieurs jets d'un spray atomiseur	0,157	0,087
66	SI, 6020, 16-2 *	evacuating air from a crushed bottle	évacuation d'air par une bouteille écrasée	0,158	0,120
67	SI, 6020, 22-1	opening a metallic can	ouverture d'une cannette métallique	0,413	0,155
68	SI, 6020, 24 1	crushing a metallic can	écrasement d'une cannette métallique	0,214	0,056
69	SI, 6020, 33-5	closing the top of an aerosol bomb	fermeture d'un bouchon de bombe aérosol	0,376	0,120
70	SI, 6020, 35-1	spray from an aerosol	jet d'un spray aérosol	0,290	0,115
71	SI, 6020, 35-2	irregular spray from an aerosol	jet irrégulier d'un spray aérosol	0,281	0,103
72	SI, 6020, 35-5	spray from an aerosol	jet d'un spray aérosol	0,350	0,151
73	SI, 6020, 59-4	puting a porcelain lid on a pan	pose d'un couvercle en porcelaine sur une casserole	0,453	0,082
74	SI, 6020, 64-1	removing the top of a plastic container	enlève le couvercle d'un récipient en plastique	0,109	0,038
75	SI, 6020, 64-4	closing the top of a plastic container	fermeture du couvercle d'un récipient en plastique	0,267	0,133
76	SI, 6020, 65-4	removing the metallic lid of a pan	enlève le couvercle métallique d'une casserole	0,294	0,151
77	SI, 6020, 68-2 *	turning a spoon inside an empty cup	cuillère que l'on tourne dans une tasse vide	0,390	0,142
78	SI, 6020, 70-1 *	hand washing-up	lavage de la vaisselle à la main	0,775	0,228
79	SI, 6020, 79-1 *	beating eggs inside a container	battage des œufs dans un récipient	0,479	0,185
80	SI, 6020, 81-3	pouring cereal into a bowl	versement de céréales dans un bol	0,368	0,090
81	SI, 6020, 81-4 *	pouring milk on cereal in a bowl	versement de lait sur des céréales dans un bol	0,092	0,047
82	SI, 6020, 82-3 *	egg open in two parts	œuf que l'on ouvre en 2	0,237	0,087
83	SI, 6020, 84-1 *	grating carrots	râpage manuel de carottes	0,094	0,040
84	SI, 6020, 85-3 *	cuting vegetable with a knife	découpe de légumes avec un couteau	0,315	0,165
85	SI, 6020, 88-1	pulling out vegetable sprays	arrachage des feuilles d'un légume	0,153	0,069
86	SI, 6020, 89-2	cuting salad in two parts	salade que l'on découpe en 2	0,116	0,115
87	SI, 6020, 91-1 *	gas noise of a furnace	bruit de gaz d'un four	0,517	0,124
88	SI, 6020, 93-5	garbage top falling	couvercle d'une poubelle qui tombe	0,532	0,356
89	SI, 6020, 99-2	grinding salt mechanically	sel moulu mécaniquement	0,169	0,050
90	SI, 6021, 4-2	put a top on a container	pose d'un couvercle sur un recipient	0,233	0,150
91	SI, 6021, 14-1	knife sharpening	couteau que l'on aiguise	0,369	0,175
92	SI, 6021, 21-2	closing the micro wave door	fermeture de la porte d'un micro-onde	0,473	0,162
93	SI, 6021, 25-2	mixer on	mixeur mis en route	0,646	0,214
94	SI, 6021, 35-1	unrolling absorbing paper, detaching a sheet	papier absorbant que l'on déroule, détache une feuille	0,290	0,122
95	SI, 6021, 54-3	lamp switch	interrupteur d'une lampe	0,245	0,203
96	SI, 6021, 69-1 *	drops in a container	gouttes à gouttes dans un récipient	0,355	0,152
97	SI, 6021, 69-2 *	drops in a container	gouttes à gouttes dans un récipient	0,199	0,087
98	SI, 6021, 76-2 *	water runing in a sink	eau qui coule dans un évier	0,478	0,055
99	SI, 6021, 77-2 *	filling a sink with water	remplissage d'un évier avec de l'eau	0,500	0,000
100	SI, 6021, 78 2 *	empyting a sink	évier qui se vide	0,315	0,129
101	SS, 4:1, 1-2 *	flow of water and stop	écoulement d'eau puis arrêt	0,328	0,140

Table 4.1:

Appendix C: Hcu for object and action

Values of the causal uncertainty of the 60 sounds for each sorter (A, B and C) for "object" and "action".

Sound #	Sorter A, Object	Sorter B, Object	Sorter C, Object	Sorter A, Action	Sorter B, Action	Sorter C, Action
1	2.4851019e+00	2.1482008e+00	2.2198700e+00	2.1534493e+00	2.0649175e+00	1.5012404e+00
2	3.2501027e+00	2.7703020e+00	3.1299133e+00	3.2319462e+00	2.6946781e+00	2.4988649e+00
3	2.2917952e+00	1.1212766e+00	1.3386406e+00	1.4693089e+00	9.0148287e-01	1.4005014e+00
4	1.5894982e+00	4.5502066e-01	1.4693089e+00	2.2853814e-01	0.0000000e+00	-0.0000000e+00
5	3.0036204e+00	3.2454472e+00	2.9039924e+00	3.0036204e+00	3.0036204e+00	2.7450852e+00
6	3.4541684e+00	3.1858309e+00	3.0413815e+00	2.0331516e+00	2.9440041e+00	2.8600260e+00
7	2.4896557e+00	2.0843054e+00	2.0843054e+00	2.4389866e+00	1.8571530e+00	2.0308551e+00
8	3.0204830e+00	2.1812998e+00	2.9464089e+00	2.1072257e+00	1.2645665e+00	1.6045671e+00
9	1.4133913e+00	6.0529121e-01	1.1904925e+00	6.7936528e-01	0.0000000e+00	3.8094659e-01
10	2.3104431e+00	1.9008634e+00	2.0331516e+00	1.3386406e+00	8.2740880e-01	8.2740880e-01
11	2.1399483e+00	2.3719728e+00	1.9782991e+00	2.8419712e+00	1.8570447e+00	1.1212766e+00
12	1.5534593e+00	1.4793852e+00	1.5534593e+00	1.4005014e+00	4.5502066e-01	1.1625337e+00
13	2.8222649e+00	2.9756616e+00	2.5675301e+00	1.6068637e+00	1.3264273e+00	2.4875067e+00
14	3.0656416e+00	2.9211921e+00	3.2454472e+00	2.1986079e+00	2.1245338e+00	3.1283635e+00
15	2.9751191e+00	2.7171264e+00	3.1615708e+00	2.9184501e+00	2.5465300e+00	2.4073291e+00
16	7.2537593e-01	5.0325833e-01	1.3195213e+00	1.5025343e+00	1.9819978e+00	6.0518658e-01
17	2.6248958e+00	2.8434193e+00	3.0837980e+00	2.1180931e+00	1.0423443e+00	8.2498026e-01
18	2.0062579e+00	1.6955038e+00	1.9782991e+00	2.9160453e+00	1.8338497e+00	1.7824859e+00
19	3.4956282e+00	3.5826103e+00	3.3278755e+00	2.2201260e+00	2.6047487e+00	3.3935954e+00
20	2.6489392e+00	1.9582292e+00	1.8109294e+00	2.3793703e+00	1.2615444e+00	1.6772390e+00
21	3.1707801e+00	2.7352829e+00	3.0982558e+00	1.8570447e+00	2.2147737e+00	2.8730862e+00
22	4.5502066e-01	4.5502066e-01	4.5502066e-01	2.1485772e+00	2.1685062e+00	1.4536770e+00
23	3.3474801e+00	4.0006457e+00	4.0747198e+00	2.7066294e+00	2.2222749e+00	2.3667243e+00
24	3.1434144e+00	3.7360069e+00	3.4019495e+00	2.8139108e+00	2.8139108e+00	2.6150935e+00
27	3.0036204e+00	3.4760236e+00	4.5326653e+00	1.1911691e+00	1.1911691e+00	2.3369537e+00
28	2.8652745e+00	2.7875017e+00	3.3097190e+00	2.9636087e+00	0.0000000e+00	3.0930073e+00
29	2.2140224e+00	1.5154242e+00	1.7550118e+00	2.5019645e+00	2.4193819e+00	1.7729583e+00
30	1.3386406e+00	1.3386406e+00	1.4793852e+00	6.7936528e-01	6.7936528e-01	2.2853814e-01
31	3.7821222e+00	3.7821222e+00	3.8789066e+00	2.0843054e+00	2.0102313e+00	1.8424786e+00
32	2.7995546e+00	2.3342117e+00	2.8879848e+00	2.0544205e+00	2.3612138e+00	2.2275234e+00
33	3.2817602e+00	3.1158698e+00	3.6566844e+00	2.5027158e+00	2.4710583e+00	2.5880914e+00
34	2.5591759e+00	1.7247565e+00	1.7247565e+00	1.9311188e+00	1.8570447e+00	1.5766083e+00
35	1.3773523e+00	6.7936528e-01	6.7936528e-01	4.5502066e-01	7.2537593e-01	2.2853814e-01
36	2.4995596e+00	2.7315842e+00	2.4691335e+00	1.6663716e+00	1.6663716e+00	0.0000000e+00
37	3.8380398e+00	3.7360069e+00	3.1443711e+00	1.9582292e+00	1.3773523e+00	1.4693089e+00
38	2.6901243e+00	2.1151457e+00	1.4793852e+00	2.2189525e+00	1.4307567e+00	2.0708044e+00
39	2.8139108e+00	2.8879848e+00	2.7269287e+00	2.9464089e+00	2.7254806e+00	2.2472863e+00
40	1.0472025e+00	1.0472025e+00	1.0472025e+00	9.0148287e-01	4.5502066e-01	4.5502066e-01
41	2.6028239e+00	2.3507168e+00	2.4844072e+00	2.5305224e+00	1.9782991e+00	1.8100810e+00
42	3.4397106e+00	3.2915625e+00	3.6158176e+00	3.3060203e+00	3.1283635e+00	2.8320672e+00
43	3.0497357e+00	2.0331516e+00	2.1072257e+00	1.8918146e+00	2.6999266e+00	2.6575101e+00
44	4.5502066e-01	4.5502066e-01	4.5502066e-01	1.0472025e+00	1.3713054e+00	6.7936528e-01
45	2.1581048e+00	2.1343361e+00	2.1343361e+00	2.4110278e+00	2.3369537e+00	2.3369537e+00
46	3.5427002e+00	3.4080531e+00	4.0747198e+00	2.5305224e+00	1.6174570e+00	3.3763956e+00
47	3.0263307e+00	3.1283635e+00	2.4566005e+00	1.3773523e+00	1.2571630e+00	2.2687665e+00
48	2.6653217e+00	2.5451324e+00	4.0043444e+00	1.5894982e+00	1.0472025e+00	1.8109294e+00
49	3.2878638e+00	3.8659986e+00	3.5245438e+00	2.6793653e+00	1.3952348e+00	1.4215542e+00
50	1.9560863e+00	2.2853814e-01	1.8341244e+00	2.4691335e+00	1.0377724e+00	4.5502066e-01
51	2.9915675e+00	2.7171264e+00	3.3060203e+00	2.9440041e+00	3.3837931e+00	3.0180781e+00
52	1.8804564e+00	1.8380398e+00	2.4467982e+00	2.7230133e+00	2.3342117e+00	2.2140224e+00
53	2.9304014e+00	2.8056582e+00	2.7967110e+00	3.0212343e+00	2.6501126e+00	2.3369537e+00
54	3.4360119e+00	3.5878588e+00	3.4541684e+00	2.8019595e+00	2.9076911e+00	3.0874968e+00
55	2.4564484e+00	2.4564484e+00	2.8879848e+00	3.3800943e+00	1.4793852e+00	1.2645665e+00
56	3.7323082e+00	3.9265717e+00	3.8789066e+00	3.1056532e+00	2.5441251e+00	2.4118829e+00
57	2.2062674e+00	1.4793852e+00	1.5534593e+00	1.0472025e+00	1.2645665e+00	1.4793852e+00
58	2.9751191e+00	2.9854639e+00	2.7912005e+00	2.6489392e+00	2.5007911e+00	2.9780664e+00
59	1.6772390e+00	1.6068637e+00	1.4005014e+00	1.4693089e+00	1.7597756e+00	1.6068637e+00
60	4.2084102e+00	4.2824843e+00	3.6023166e+00	3.5319412e+00	3.0278804e+00	3.4504697e+00
61	2.2140224e+00	2.8269709e+00	2.1765360e+00	1.6031650e+00	1.2615444e+00	2.2853814e-01
62	3.2780615e+00	3.0640918e+00	3.3278755e+00	3.3982508e+00	3.1842811e+00	3.5085362e+00
64	3.2863140e+00	3.4215542e+00	2.6356549e+00	3.2636037e+00	2.5164222e+00	2.5056632e+00
65	3.7178505e+00	2.3362590e+00	1.8581097e+00	2.5584812e+00	1.8581097e+00	1.7840357e+00
66	4.2508267e+00	4.3845171e+00	4.4585912e+00	4.1343361e+00	3.8380398e+00	3.6060153e+00
67	3.4019495e+00	3.3278755e+00	3.4019495e+00	2.3806017e+00	2.9010450e+00	2.9010450e+00
68	3.0134227e+00	2.9393486e+00	2.4337381e+00	2.7875017e+00	1.9311188e+00	1.4133913e+00
69	3.8804564e+00	3.2356449e+00	4.3565583e+00	2.0331516e+00	2.8443761e+00	2.8093569e+00
70	2.9765167e+00	2.7134277e+00	2.3559653e+00	3.8100810e+00	1.9390388e+00	2.4746049e+00
72	3.0413815e+00	2.5441251e+00	2.4239358e+00	3.2999167e+00	2.4980098e+00	2.5773324e+00
73	1.9008634e+00	1.9008634e+00	4.0747198e+00	2.2726819e+00	6.7936528e-01	1.3773523e+00
74	3.9302704e+00	4.0323032e+00	4.5047065e+00	2.7423432e+00	2.6067394e+00	3.3800943e+00
75	2.9341001e+00	2.6197490e+00	4.0141468e+00	1.9008634e+00	1.9008634e+00	2.9575051e+00
76	1.8570447e+00	2.0692546e+00	3.8561963e+00	2.9039924e+00	9.4516961e-01	2.5689782e+00
77	2.4564484e+00	2.1601521e+00	2.2601376e+00	3.1713732e+00	1.8918146e+00	2.3285995e+00
78	2.8652745e+00	3.2282475e+00	2.6768980e+00	3.8659986e+00	2.7990121e+00	2.2880964e+00

Continued on next Page...

Sound #	Sorter A, Object	Sorter B, Object	Sorter C, Object	Sorter A, Action	Sorter B, Action	Sorter C, Action
79	3.5697023e+00	3.3023215e+00	2.9464089e+00	2.6325553e+00	1.8109294e+00	2.1399483e+00
80	1.9787966e+00	1.6174570e+00	1.4693089e+00	1.1904925e+00	1.0884596e+00	-0.0000000e+00
81	4.1622949e+00	4.4585912e+00	4.1343361e+00	3.3603881e+00	2.8299183e+00	3.2303964e+00
82	4.1063773e+00	4.3565583e+00	4.2824843e+00	2.9319512e+00	2.9319512e+00	3.2599050e+00
83	3.7080482e+00	4.1804514e+00	4.0747198e+00	3.0967060e+00	1.7144117e+00	1.9204681e+00
84	3.1299133e+00	3.7467660e+00	3.7467660e+00	1.7550118e+00	6.7936528e-01	9.0148287e-01
85	3.5841601e+00	3.2999167e+00	3.4676694e+00	3.8380398e+00	2.5480798e+00	2.7703020e+00
86	2.7134277e+00	3.0097240e+00	2.5373207e+00	3.3278755e+00	2.4036303e+00	1.7565744e+00
87	2.8019595e+00	2.3410666e+00	2.8707851e+00	3.0505907e+00	2.7315842e+00	2.6015925e+00
88	3.2454472e+00	3.0497356e+00	3.3460319e+00	3.1117568e+00	2.7693453e+00	2.5591759e+00
89	3.7027997e+00	3.4397107e+00	3.3195213e+00	3.7080482e+00	2.9946732e+00	2.1912104e+00
90	2.4579982e+00	1.9078221e+00	3.1283635e+00	1.9590776e+00	1.1904925e+00	2.7398367e+00
91	2.2853814e-01	0.0000000e+00	-0.0000000e+00	7.2537593e-01	0.0000000e+00	0.0000000e+00
92	1.8571530e+00	1.7368553e+00	2.0331516e+00	6.7936528e-01	9.4516961e-01	6.0529121e-01
93	2.9749669e+00	2.3246841e+00	2.7066294e+00	3.4504697e+00	1.8565600e+00	2.8250462e+00
94	3.8561963e+00	3.3460319e+00	3.2719579e+00	2.9039924e+00	2.7226369e+00	2.7315842e+00
95	3.8048325e+00	3.8048325e+00	3.8048325e+00	3.6484318e+00	3.5546515e+00	3.7768737e+00
96	1.6214298e+00	6.7936528e-01	1.8424786e+00	2.4436420e+00	1.0377724e+00	1.1749947e+00
97	2.3369537e+00	1.3386406e+00	1.3386406e+00	2.9076911e+00	1.6031650e+00	1.6955038e+00
98	1.4215542e+00	9.7312841e-01	6.0529121e-01	1.6395862e+00	9.6827025e-01	6.9128987e-01
99	2.2167644e+00	2.4352879e+00	2.0882208e+00	1.4708767e+00	1.0860486e+00	1.0860486e+00
101	1.4534670e+00	1.0472025e+00	4.5502066e-01	2.0386667e+00	7.2537593e-01	1.0493985e+00

Table 4.2:

Appendix D.
Verbatim of the experimental instructions provided to the participants in experiment 1.2

(Translated from the French). The participants were given the instructions for each step one by one, without knowing the goal of the next step. Particularly, they were not aware that the will have to describe their categories.

Step 1: free classification

"Classification of kitchen sounds

Introduction
The goal of this experiment is to classify a set of sounds, recorded in different kitchens. Your task is therefore to make classes from this set of sounds. The sounds are played by the loudspeakers. The levels of these sounds correspond to their levels in your kitchen.

Goal
Your task is to make as many classes as you wish, with as many sounds as you wish in each class. Use your own criteria to make the classes.

Outline of the experiment

- After a short training phase, you will have to make classes with the 60 sounds displayed randomly on the computer screen.
- The sounds are represented by red dots.
- When you double click on a red dot, the sound is played, and the dot becomes pink. When you click outside the dot, it becomes red again.
- When you click on a sound and move the mouse, you move the dot. You have to move the dots to form the classes.
- To select several sounds, you have to click on the background and draw a rectangle surrounding the dots you want to select. In this case, the dots become green. Then, you can move the selected sound by clicking on the green dots and moving the mouse.
- When you have finished to classify the sounds, please ask the experimenter to record your results.

Note:
Do not spend more than 40 minutes for this task."

Step 2: description of the classes

"Now that you have made classes of sounds, please describe these classes to the experimenter.
For each of the classes you made, you will describe the properties shared by the sounds within the class. These properties have to explain why you have grouped together these sounds in this category."

Step 3: coding the descriptions

”Now that you have described the properties shared by the sounds within each of the classes, please answer the following question, by choosing one of the answers proposed below:
According to you, the sounds within this class:
1. Are similar because they ”sound” the same way ?
For instance, the sounds may be grouped together because:
• they all are low in pitch,
• they all are short rhythmic sounds,
• they all have a very low level,
• etc. . . .
2. Are similar because they all have been caused by the same physical event ?
For instance, the sounds may be grouped together because:
• they all are sounds made by impacts,
• they all are made by water drips,
• they all are made by wooden objects,
• they all are made by electrical motors,
• etc.
3. Are similar for more abstract reason ?
For instance, the sounds may be grouped together because:
• they all happen during breakfast,
• they all happen in a restaurant kitchen,
• they all sounds related to food preparation,
• etc.
4. Are similar for some other reasons ?
5. You do not know why you made this class ”

Appendix E.
ANOVA tables for experiment 1.2

Source	df	Sum of squares	Mean squares	F-value	p-value	Corrected p-value (GG)
$\mathcal{E}$ Expertise	1	6.108	6.108	21.603	0.0001**	
$\mathcal{S}(\mathcal{E})$ Subjects(Expertise)	28	7.917	0.283			
$\mathcal{HCU}$ Intervals of Hcu	4	0.588	0.147	10.830	0.0001	0.0001**
$\mathcal{HCU} * \mathcal{E}$ Intervals of Hcu*Expertise	4	0.085	0.021	1.573	0.1863	0.2027 n.s.
$\mathcal{S}(\mathcal{E}) * \mathcal{HCU}$	112	1.521	0.014			
df: degree of freedom; GG: Geisser-Greenhouse corrected p-value for sphericity violation ** p<0.01; n.s.: not significant						

Table 4.3: Anova table for experiment 1.2. The dependant variable is the ratio of sounds grouped together because of their acoustic similarity.

Source	df	Sum of squares	Mean squares	F-value	p-value	Corrected p-value (GG)
$\mathcal{E}$ Expertise	1	3.697	3.697	11.527	0.0001**	
$\mathcal{S}(\mathcal{E})$ Subjects(Expertise)	28	8.981	0.321			
$\mathcal{HCU}$ Intervals of Hcu	4	01.271	0.318	19.014	0.0001	0.0001**
$\mathcal{HCU} * \mathcal{E}$ Intervals of Hcu*Expertise	4	0.084	0.021	1.262	0.2890	0.2926 n.s.
$\mathcal{S}(\mathcal{E}) * \mathcal{HCU}$	112	1.872	0.017			
df: degree of freedom; GG: Geisser-Greenhouse corrected p-value for sphericity violation ** p<0.01; n.s.: not significant						

Table 4.4: Anova table for experiment 1.2. The dependant variable is the ratio of sounds grouped together because of their event similarity.

Source	df	Sum of squares	Mean squares	F-value	p-value	Corrected p-value (GG)
$\mathcal{E}$ Expertise	1	0.512	0.512	2.389	0.1334 n.s.	
$\mathcal{S}(\mathcal{E})$ Subjects(Expertise)	28	5.996	0.214			
$\mathcal{HCU}$ Intervals of Hcu	4	0.068	0.017	1.741	0.1460	0.1693 n.s.
$\mathcal{HCU} * \mathcal{E}$ Intervals of Hcu*Expertise	4	0.029	0.007	0.755	0.5565	0.5134 n.s.
$\mathcal{S}(\mathcal{E}) * \mathcal{HCU}$	112	1.092	0.010			
df: degree of freedom; GG: Geisser-Greenhouse corrected p-value for sphericity violation ** p<0.01; n.s.: not significant						

Table 4.5: Anova table for experiment 1.2. The dependant variable is the ratio of sounds grouped together because of the similarity of their semantic similarity

Source	df	Sum of squares	Mean squares	F-value	p-value	Corrected p-value (GG)
$\mathcal{E}$ Expertise	1	0.039	0.039	2.438	0.1296 n.s.	
$\mathcal{S}(\mathcal{E})$ Subjects(Expertise)	28	0.452	0.016			
$\mathcal{HCU}$ Intervals of Hcu	4	0.036	0.009	3.234	0.0149	0.0319*
$\mathcal{HCU} * \mathcal{E}$ Intervals of Hcu*Expertise	4	0.022	0.006	2.013	0.0974	0.1261 n.s.
$\mathcal{S}(\mathcal{E}) * \mathcal{HCU}$	112	0.311	0.003			
df: degree of freedom; GG: Geisser-Greenhouse corrected p-value for sphericity violation ** p<0.01; * p<0.05; n.s.: not significant						

Table 4.6: Anova table for experiment 1.2. The dependant variable is the ratio of sounds grouped together because of another kind of similarity

Source	df	Sum of squares	Mean squares	F-value	p-value	Corrected p-value (GG)
$\mathcal{E}$ Expertise	1	0.00001067	0.00001067	0.012	0.9140 n.s.	
$\mathcal{S}(\mathcal{E})$ Subjects(Expertise)	28	0.025	0.001			
$\mathcal{HCU}$ Intervals of Hcu	4	0.003	0.001	1.783	0.1371	0.1582 n.s.
$\mathcal{HCU} * \mathcal{E}$ Intervals of Hcu*Expertise	4	0.001	0.0003723	0.788	0.5351	0.5009 n.s.
$\mathcal{S}(\mathcal{E}) * \mathcal{HCU}$	112	0.053	0.0004723			
df: degree of freedom; GG: Geisser-Greenhouse corrected p-value for sphericity violation ** p<0.01; n.s.: not significant						

Table 4.7: Anova table for experiment 1.2. The dependant variable is the ratio of sounds grouped together without any reason.

Appendix F.
hierarchical clustering for expert group. Experiment 1.2

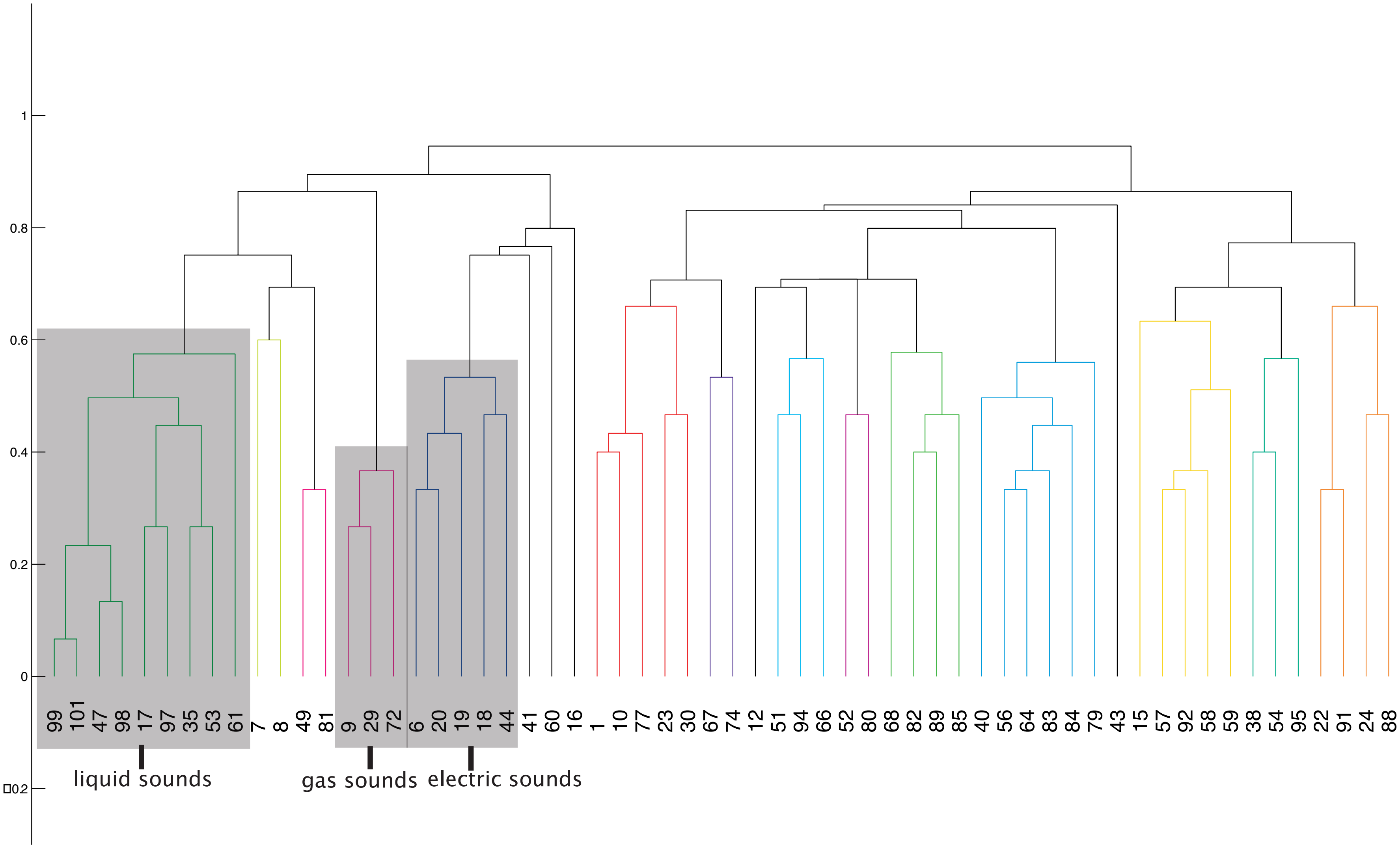

Figure 4.1: Cluster analysis of the proximity matrix of the group of expert for the 60 sounds using Hubert et al. algorithm [5]

Bibliography

[1] James A. Ballas. Common Factors in the Identification of an Assortment of Brief Everyday Sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2):250–267, 1993.

[2] Frédérique Guyot. *Étude de la perception sonore en terme de reconnaissance et d'appréciation qualitative: une approche par la catégorisation.* PhD thesis, Université du Maine, 1996.

[3] Yannick Gérard. *Mémoire sémantique et sons de l'environnement.* PhD thesis, Université de Bourgogne, 2004.

[4] John M. Hajda, Roger A. Kendall, Edward C. Carterette, and Michael L. Harshberger. Methodological issues in timbre research. In Irène Deliège and John Sloboda, editors, *Perception and cognition of music*, chapter 12, pages 253–306. Psychology Press, 1997.

[5] LJ Hubert, P Arabie, and J Meulman. *The Structural Representation of Proximity Matrices with MATLAB.* Society for Industrial and Applied Mathematics; American Statistical Association, 2006.

[6] Joseph B. Kruskal and Myron Wish. *Multidimensional scaling.* Sage University Paper series on Quantitative Applications in the Social Sciences 07-011. Sage Publications, 1978.

[7] Michael M. Marcell, Diane Borella, Michael Greene, Elisabeth Kerr, and Summer Rogers. Confrontation naming of environmental sounds. *Journal of clinical and experimental neuropsychology*, 22(6):830–864, 2000.

[8] Eleanor Rosch. Principles of categorization. In Eleanor Rosch and Barbara B. Lloyd, editors, *Cognition and Categorization*, chapter 2, pages 27–48. Lawrence Erlbaum Associates, 1978.

[9] Gary P. Scavone, Stephen Lakatos, and Colin R. Harbke. The sonic mapper: An interactive program for obtaining similairty ratings with auditory stimuli. In *Proceedings of the 2002 International Conference on Auditory Display (ICAD02)*, Kyoto, Japan, July 2002.

[10] B. Smith. PsiExp: an environment for psychoacoustic experimentation using the IRCAM musical workstation. In *Society for Music Perception and Cognition Conference '95.* University of Berkeley, 1995.

[11] P. Susini, S. McAdams, S. Winsberg, I. Perry, S. Vieillard, and X. Rodet. Characterizing the sound quality of air-conditioning noise. *Applied Acoustics*, 65(8):763–790, 2004.

[12] Nancy Jean Vanderveer. *Ecological acoustics: human perception of environmental sounds.* PhD thesis, Cornell University, 1979.