A New Experimental Approach for Urban Soundscape Characterization Based on Sound Manipulation : A Pilot Study

G. Lafay[a], M. Rossignol[b], N. Misdariis[b], M. Lagrange[a] and J.F. Petiot[a]
[a]Ecole Centrale de Nantes, IRCCyN, 1 rue de la noe, 44321 Nantes, France
[b]IRCAM, 1 Place Igor-Stravinsky, 75004 Paris, France
gregoire.lafay@irccyn.ec-nantes.fr

In this paper, we aim at better understanding how human mental representations are structured in the specific case of the perception of urban soundscapes. This task is traditionally studied using questionnaires, surveys or categorization tasks followed by a lexical analysis. In contrast, we propose a new experimental approach to tackle this aim. In this approach, the subject is asked to manipulate sound events and textures within a dedicated computer environment in order to recreate two complex urban soundscapes, one ideal and the other not ideal. Subjects have access to a sound data set which has been designed and structured based upon perceptual considerations, and may alter the physical parameters of the selected sound samples. In order to achieve this, we use an audio-digital environment and a web audio interface for sound mining developed for the purposes of this study. The latter allows subjects to explore a sound database without resorting to text. By focusing on the auditory modality during the experimental process, this new paradigm potentially allows the subject to be better put in context and provides a more detailed description of the actual mental representations. In the light of the results presented in this paper, it seems that it also reduces the potential bias of only using verbalization during the experiment.

1 Introduction

Even if the city has always been a noisy environment, whatever the times, our perception of this noise has evolved. It is in the 80s that the association between noise and pollution has been the strongest. Noise was considered as an overall degradation of the quality of life [1]. In response, ”anti- noise ” legislation took place and planed to fight the noise by reducing its intensity level. But the problem remains, and for good reason, noise is a subjective phenomenon, that depends on the ”listener appreciation”. The noise is a matter of context. It may entertain as well as disturb or annoy. Improving the sound environment by only focusing on acoustical parameters that are by definition objective, is not enough. Besides, this acoustical approach do not tackle directly the issue of improving the soundscape, as it is focused on isolated ”negative sounds” and not interconnected ”positive sounds”. To summarize, a pleasant city is not a silent city.

Understanding the noise requires a methodology that differs from the more traditional approach of psychophysics, as noise is more ”a cognitive object than a physical object” [2]. For those reasons a new concept of soundscape had been introduced by M. Schafer [3]. A commonly agreed definition of the ”soundscape” has been given by Truax who worked into the *World Soundscape Project*: ”an environment of sound (sonic environment) with emphasis on the way it is perceived and understood by the individual, or by a society” (Truax quoted in [4]). It’s a subject centered approach that considers sound environment from the listener point a view. It assumes that the qualitative evaluation of a soundscape mainly depends on the context and the knowledges of the community or the subject that experiment the environment. The question is no longer when the noise is annoying, but why the sound is annoying.

Since its introduction, this approach has been widely used to study urban sonic environment [5]. Many sociological or acoustical psycho-cognitive studies developed a database of qualitative descriptors and sounds to better understand our sound environment. One major today’s challenge is to connect these perceptual data derived from questionnaires, categorization task and psycho-linguistic studies to acoustic measurements in order to establish an effective policy of noise reduction adapted to each situation [6]. To this end, we propose a new experimental approach to study soundscape perception. In order to investigate what could be the nature of an ideal urban environment, we ask subjects to reconstruct two soundscapes, one qualified as ”ideal” (in which you would like to live) the others ”not ideal” (in which you would not live), from a dedicated urban sound data set. By analysing the reconstruction process, we are able to objectify the mental representations of an ideal (resp. not ideal) sound environment. We will first introduce the experimental paradigm as well as the technological devices on which it relies, then we will detail the creation of our sound data set, and finally we will present the results of a pilot study.

2 Experimental Protocol

2.1 Paradigm of the proposed approach

Prevent psycho-cognitive studies addressing mental representation of sound environment use linguistic resources to objectify mental categories. Two types of approach has been widely used: The **sorting** tasks, which ask the subject to sort sounds and label the classes (categorization tasks), and the **describing** tasks which ask the subject to describe a sound via a adapted questionnaire (usually free or semi structured questionnaire). Both approach considers verbal data as input, and thus rely on psycho-linguistic analysis to objectify generic and meaningful mental categories from individual names or labels.

In this paper, the proposed approach is also a psycho-cognitive one, as we also consider subjective literal data, but joined with objective data assumed to be directly linked to the ”reality of the word”. To do so, we adopt a reconstruction approach. The subject is asked to recreate a complex sound environment by choosing sounds in a sound data set, and modifying the physical properties of them (sound level, time positioning). The selection process is made without any text-written help, see Section 2.3 for more details. The subject must name each selected sound. After the creation process, the subject has to give a global name to the scene and freely comment its creation. Doing so, we obtain 2 types of data: the **objective data** that depends on the sound data set used and thus are non ambiguous and controllable by the experimenters, and the **subjective data** that depends on the subject. For the **Subjective data** we have the **name** given by the subjects to the sounds they selected, the **title** and the **free comments**. **Objective data** are composed of 1) **numerical data** that are the audio control parameters values set by the subjects to reconstruct the scene, and 2) non ambiguous verbal data called **tag**. A **tag** is the ”real name” of a sound element, which is given by the experimenter who record it. The **tags** of the sounds act as a ground-truth compared to the **name** given by subject.

Mixing both objective and subjective data allows us to obtain subject-centered results which are meaning full and easily analysable, as the two may always be weighted relative to each other to give potentially lesser non ambiguous results.

2.2 Creation of the sound bank

Special care has been taken to construct the sound data set as it represents the "sonic world" on which subjects will rely to recreate a desired soundscape. The goal was to propose a sound data set that 1) is representative of the diversity of all the sounds that populated the urban environment, 2) is able to offer several variants of a same sound, and 3) could be quickly explored in order to supplies to the time constraints of the experiment. Moreover, as our *tags* entirely depend on the sound data set nomenclature, a consistent and, as much as possible, generic typology of the different classes of urban sounds had to be designed. To this sense, the sound data set structure was motivated by practical and perceptual considerations.

We decided to divide the sound data set in two parts of perceptively distinct sounds: the **events** and the **textures**. Several studies have pointed out that this two types of sounds provoke two distinct cognitive process. Maffiolo [7] showed that *event sequences* lead to a descriptive (semantic) analysis, which mainly rely on the identification of the present sound events, whereas *amorphous sequences*, which are sequences without salient events that can be regarded as textures ("traffic hubbub", "street hubbub"), lead to a holistic analysis, depending on the acoustical properties of the sequence. McDermott and Simoncelli [8] recently showed that sound textures perception mainly depends on the analysis of simple statistical properties of the acoustical signal. For those reasons we consider a soundscape as "a skeleton of events on a bed of textures" [9].

Several studies address the difficulty of obtaining a generic classification of all the elements that populated our sound environment [4][10] , as it is difficult to regroup them under consistent and non ambiguous classes. Indeed, sound classes as "traffic sound", "human voice" or more precisely classes as "cars" or "sound of a man" may regroup a large variety of sounds, depending of the study context. Considering those remarks, we design a hierarchical sound data set of urban environmental sounds adapted for our approach. In this hierarchy, classes of top level represent large categories of concept as "sounds of nature" or "motorised transportation sounds". The deeper the class is in our typology and the lesser is the diversity of the sound that belong to it. Each class of the leaf level ("car passing", "man yelling"), regroup several exemplars of what could be regarded as perceptively close sounds. In order to find a typology of urban sounds with perceptual inspired class names, we did a bibliographical review of several paper addressing mental categories of urban environmental sounds [2][4][7][10][11][12] [13][14] .

It is important to note that the subject only interact with **leaf classes** to recreate the soundscape, and never with **a particular exemplar** belonging to those different leaf classes.

2.3 Simulating a soundscape

To simulate a soundscape we use a web software called *Simscene* adapted to this study. *Simscene* allows the user to manipulate sound classes of events and textures ("man yelling", "car passing", "rain", "parc hubbub") in order to create a targeted soundscape. In order to free the subjects from the a priori fixed nomenclature of a sound data set, the selection interface of *Simscene* has been designed not to use a text based keyword search. The selection interface allows the subject to explore the data base by listening to the sounds. To do so, we dispatch the leaf classes of our typology in a 2D space, each leaf class being represented by a circle, see figure 1. The space configuration of the circles depends on the hierarchical structure of each top-classes of the typology. Circles representing leaf classes which belong to the same class are packed together, and so on until the top-classes. There is no correlation between the spatial configuration and the top-classes.

The subject explore the data set by clicking on the circles. When a circle is clicked, the subject hears a sound prototype of the leaf class represented by the circle. By listening to the different prototypes the subject may select a sound class.

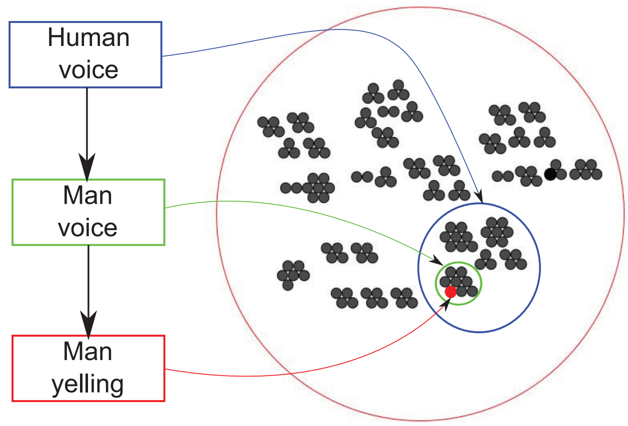

Figure 1: Sound selection Interface of *Simscene* with illustration of the urban sound data set hierarchy

After having selected a sound class, a time lime populated with sound events or textures belonging to the selected class is displayed. The subject may then adjusts control parameters as average/variance of time intervals between events, or average/variance of sound levels to create the scene. As subjects interact with sound classes, control parameters have been designed to manipulate sequence of sounds and not unique sound. The *Simscene* environment designed for this study can be accessed via the url: `http://217.70.189.118/soundthings/SceneSynth-SSF/`.

3 The Pilot study

3.1 Task and collected data

Subjects were asked to successively create two urban soundscapes. The first must be ideal (ie. the favourite urban soundscape of the subject), the second not ideal (ie. the worst urban soundscape of the subject). No instructions were given concerning sounds to choice or audio parameters to apply. The only constraint was to recreate a **physically plausible** urban soundscape. In other words, unrealistic situations as "a dog barking every 10 milliseconds" were forbidden.

Two types of data were collected: The **verbal** data and the **numerical** data. For the **verbal** data, on the one hand we consider the **subjective** data that depends of the subject evaluation, from which we analyse the **title** of the soundscape, the **names** given by the subject to each chosen sounds, the *free comment* describing the reconstruction process and the *missing sounds* that the subject did not find in the sound data set. The **numerical** data are the **audio control parameters** from which we will analyse the **relative sound level** of each sound, and the **time intervals** between sound events.

3.2 Participant

Participants were 10 french subjects individually selected from volunteers. They were 7 males and 3 females and were about the same age (M : 24.1, STD 1.7). All of them are urban dwellers.

3.3 Apparatus

Subjects have access to a sound data set of 483 urban environmental sounds including 381 sound events and 102 textures. Among them, 260 events and 72 textures were recorded. 121 sound events and 30 textures which proved to be particularly difficult to record came from existing sound banks. All recordings were performed using a shotgun microphone *audio technica* AT8035 connected to a *ZOOM* H4n recorder. We chose to use a shotgun microphone (highly directive) in order to isolate sound recordings from undesired events. All the sounds were normalized to the same peak level.

The experiment was run with one subject at a time, in the audiometric test booth of the IRCAM french institute[1]. Audio was presented in monophonic to each participant via *Yamaha MSP 5* speakers (active speaker), on *Macintosh Mac Book Pro* type computer, connected to a *RME FireFace 800* sound card. The software (SimScene) was located on a distant server and loaded via Google-Chrome navigator in a Linux operating system. At the end of each task, data were automatically collected server side. One experimenter was always present to give instructions and answer queries if needed. Experiment lasts about one hour and a half for each subject.

4 Results of the Pilot study

4.1 Objective and Subjective verbal data analysis

Selected sound analysis

For the ideal scenes, subjects used 51 event classes and 27 texture classes. For not-ideal scenes we have 96 event classes and 32 texture classes used. This result indicates that not-ideal scene are composed of a larger number of sound sources than ideal scene. Table 1 displays the number of texture and event classes used in average by the subjects.

We first look at the *tag* (see section 3.1) of the sounds chosen by the subjects. We regroup the *tags* based on their position in our typology. We chose the semantic hierarchical

Table 1: Average and standard deviation of the number of event and texture classes used by each subjects

	Events	Textures
Ideal scenes	5.1;2.4	2.7;1.1
not-ideal scenes	9.6;3.1	3.2;1.8

levels that are the most meaning full. Figure 2 and 3 display the results for the ideal and not ideal scenes. It is found that the most frequent events chosen by the subjects for not-ideal scene are: ”alarm” sounds and ”horn” (35% of occurrences), ”construction work” sounds (22% of occurrences) and ”traffic” sounds (13.5% of occurrences). The majority of the textures are related to ”construction work” sounds: (*construction work hubbub* : 28% of occurrences , *work vehicles* : 19% of occurrences) and ”human voices” (25% of occurrences). For ideal scenes, events are mainly sounds of human (*footsteps* : 21.6% of the occurrences, and *human voices* 19.6% of occurrences), sounds of ”birds” (11.8% of occurrences), ”sounds of bicycle” (11.8% of the occurrences) and ”bells” (11.8% of occurrences). For textures, sounds of human are still very presents (*voice*: 33.3% of occurrences). Some results may be considered as counter-intuitive. For the ideals scenes, we observe the presence of ”alarm”, ”construction work” and ”traffic” sounds events. Traffic textures (”traffic hubbub”) are also well represented (25.9% of occurrences). The same applies to the not-ideal scenes where textures of ”human voices” and ”fountains” are observed.

To check if those results are due to a miss identification of the chosen sounds by the subject, we look at the *names* (see section 3.1) given by the subject to each selected sounds. To achieve the linguistic analysis we rely on the following rule: We link a *tag* to a *name* when the latter is explicitly referring to a sound source (”foot step”, ”man calling someone”), or a sound ”background” (”courtyard”, ”street atmosphere”). For *names* that cannot be explicitly linked to a *tag*, we check if they belong to the same lexical field.

1. If we detect a common lexical field along the *names*, we regroup them under a same designation
2. If we do not detect a common lexical field, we eliminate the isolated *name* of the analysis. These sounds are then referred to as ”untreated”

We show the results of the linguistic analysis in Figures 4 and 5 for respectively the ideal scenes and the not-ideal scenes. In general, the events were well identified with 90% of non ambiguous associations between *names* and *tags* for the ideal scenes, and 92% for the not-ideal scenes. The identification is more difficult for textures with only 48% of non ambiguous associations between *names* and *tags* for the ideal scenes and 47% for not-ideal scenes. We observe that more than 25% of the textures names do not refer to a sound source, but to a global description of the all texture. These observations are in line with those obtained by Maffiolo ([7]) which states that the *sequences of events* cause semantic analysis (ie. identification of sound sources), while amorphous sequences are subject to a holistic treatment which relies on acoustical considerations, and thus does not promote the objectification of the names of the sound sources.

[1] Ircam: (*Institut de Recherche et Coordination Acoustique/Musique*) www.ircam.fr/

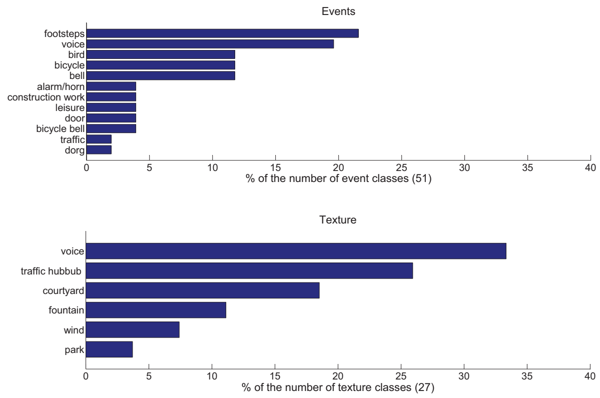

Figure 2: *Tags* chosen by the subjects for the ideal scenes (percentage of the total amount *tags* used in ideal scenes)

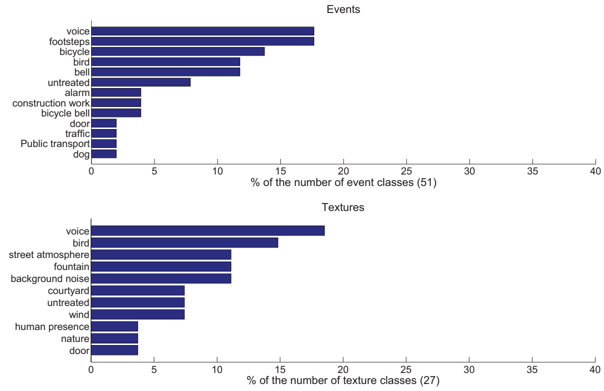

Figure 4: *Names* given by the subjects for the ideal scenes (percentage of the total amount of *names* used in ideal scenes)

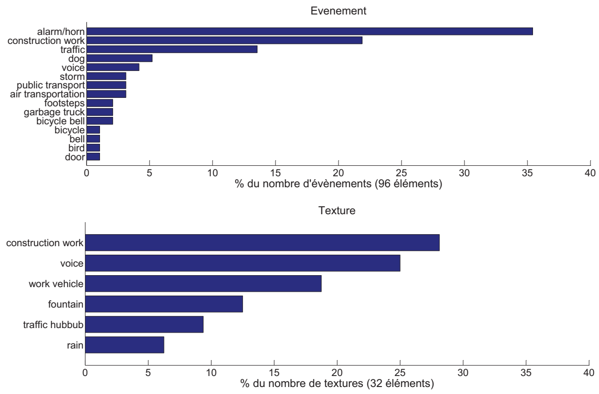

Figure 3: *Tags* chosen by the subjects for the not-ideal scenes (percentage of the total amount of *tags* used in not-ideal scenes)

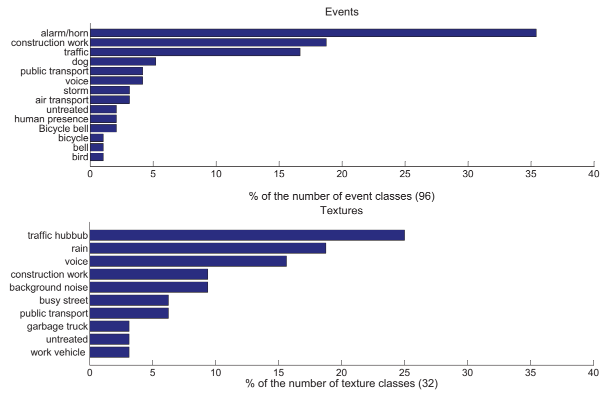

Figure 5: Names given by the subjects for the not-ideal scenes (percentage of the total amount of *names* used in not-ideal scenes)

We also note that several names given to textures for ideal as well as not-ideal scenes directly refer to what could be considered as sound events (birds, public transport). This is most probably due to the fact that it is very difficult to record a scene without any sound events. Some of them, specially birds sounds, occur in the texture sequences. It is interesting to note that it is those unwanted events that lead the identification process of the textures, even if they are very poorly represented in the textures. This underlines the fact that semantic value of a soundscape mainly depends on the identifiable sound events.

Regarding textures of the ideal scenes, we see a significant decrease of the number of sounds related to ”traffic hubbub” between the *tags* and *names*. Traffic sounds are designated by the subject with generic names: ”street atmosphere” (”ambiance de rue”), or ”background” (”bruit de fond”), indicating that sounds related to urban traffic are indeed accepted if they are part of the background of the scene. There is a notable difference between *tags* and *names* for not-ideal textures. Many sound tagged ”construction work” has been identified as urban traffic sounds.

Tags analysis reveals the presence of sound events belonging to the *alarm/horn* category (sounds used by one subject) and *construction work* (sounds used by 2 subjects) for the ideal scenes. By analysing the free comments given by the subjects we note that the alarm sounds have been conscientiously chosen, and are considered as ”discreet and nice to hear”. According to the subjects, *construction work* sounds have been chosen for realistic considerations.

We propose to compare our results to those obtained with a similar study driven by Guastavino [12], in which she asked subjects to describe their *ideal urban sound environment*. Thanks to a psycho-linguistic analysis she derived categories of *positive* and *negative* sounds. The results we obtain with ideal scenes are consistent with those of Guastavino. In both cases the main represented categories are: human sounds (*foot step* and *voice*) and sound of nature (*animals*, *weather*). The difference lays in the appreciation of some categories. In both studies, sound categories of ”traffic”, ”construction work”, ”alarm/horn” are used to characterize ideal cities. In the case of Guastavino, terms relative to those categories was found in negative construction (”with no cars” or ”with less traffic”), as they give rise to negative judgments. In our case, the analysis of the free comments made by the subjects of their scenes showed that those sounds were used to make the reconstructed environment plausible. Several subjects clearly indicated that they found distant sounds of ”traffic” and ”construction work” pleasant.

Missing sounds analysis

Based on the analysis of descriptions we identify 23 references to missing sounds, 14 for ideal scenes and 9 for not-ideal scenes. Of those sounds, 3 refer to musical sounds and 10 refer to sounds well present in the database that the subjects did not found. As musical sounds were deliberately not integrate in the data base typology, this leave us with 10 well missing sounds. This small number tend to shows that the proposed corpus was representative of a urban sound environment. Note that 50% of the subjects spontaneously specified in the free comments that the diversity of the corpus was sufficient for the ideal scenes (20% subjects), the not-ideal scenes (20% subjects), or both (10% subject).

Title analysis

We perform here the analysis of the titles given by the subject to the reconstructed scenes from which we seek to identify categories of geographic location. As a specify title may contains several distinct semantic entities, one title may be put in several geographical categories. From the lexical analysis of the 10 ideal scenes titles, we derive 4 categories: ”park” (4 subjects) ”pedestrian space” (3 subjects), ”courtyard” (2 subjects) and ”quiet street” (2 subjects). Titles categories refer to urban area where traffic noise is absent or limited, but where human sounds are well present. From the analysis of the not-ideal titles, we identify four geographical categories: ”street” (4 subjects) , ”Boulevard” (2 subjects) , ”avenue” (2 subjects) and ”crossroads” (2 subjects). 3 subjects titles clearly indicates the presence of construction work sounds, and three others the presence of traffic sounds.

4.2 Numerical data analysis

Relative sound levels and time intervals

In this section we address the relative sound levels and the time intervals fixed by the subjects via the control parameters. Table 2 displays the means and standard deviations of the events and textures relative sound levels for ideal and not-ideal scenes. It should be noted that all sounds have an initial relative level fixed to 0 dB SPL, and subjects may only reduce the average sound level. We find that the average levels for ideal scenes, are lower than those of not-ideal scenes. Similarly textures relative levels are lower than events levels (*t-test*: between ideal events and textures: $p = 0.02$; between not ideal events and textures: $p = 5.10^{-4}$).

Table 3 shows the averages and standard deviations of the time intervals between events. Note that this parameters only applied for events sounds. Given the length of the synthesized scene (60 seconds), we note that the average time intervals between events are not significantly different for the ideal and not ideal scenes (*t-test*: $p = 0.59$).

Results indicate that if the ideal city has an overall level lower than the not-ideal city, it remains an active environment in term of sound event apparition.

Table 2: Average and standard deviation of the events and textures relative sound levels for the ideal and not-ideal scenes

	Events(dB)	Textures(dB)
Ideal scenes	-6.1;7	-10.3;8.1
Not-ideal scenes	-1.2;3.2	- 4.1;5.5

Table 3: Average and standard deviation of the time intervals between events for the ideal and not-ideal scenes

	Events
Ideal scenes	17;16
Not-ideal scenes	18.5;16.5

5 Discussion and perspective

In this paper, a new experimental protocol used to objectify mental representations of urban soundscapes has been introduced. To perform the experiment, we created a corpus of environmental sounds based on a typology established from a dedicated literature review. To free ourselves from the influence of a priori nomenclature, as well as the lack of lexical words to describe adequately the acoustic phenomena [15],[12], we have developed an interface which allows subject to explore a environmental sound data set, structured according to perceptual considerations, by listening to the sounds themselves.

The comparison between our results and those of Guastavino [12] shows that our approach provides consistent results. However, we note a difference. While the approach of Guastavino allows to consider the mental representations of an urban sound environment as a whole, our experience helps to refine the category analysis thanks to a pre established typology.

The pilot study shows that the proposed protocol provides us with ecologically viable data. Indeed, motivated by a realism concern, some subjects have voluntarily placed in their ideal scenes sound that could be a priori considered as unpleasant events. We also see this trend for the textures of the ideal scenes from which 25.9% are sounds relative to ”traffic hubbub”. Several subjects stated in their free comments that they deliberately deleted sound elements of their soundscape in order not to overload the final sound environment, particularly concerning the ideal scenes. This fact reinforces the ecological validity of our experimental protocol. This deletion is in fact the result of an awareness, by the subject of the sonic context of an urban environment. This awareness is encouraged by the experimental paradigm. We believe that the reconstruction process acts as a mirror, allowing the subject to adjust its responses during the experiment, and thus provides meaningful data. Subjects do not manipulate isolated sounds, but interconnected sounds as in the ”real word”.

The analysis of numerical data shows a significant difference between the average sound levels of an ideal urban environment and that of a not-ideal environment. Nevertheless none of the environments present a lower activity than the other in terms of occurrences of sound sources.

To conclude, the pilot study described in this paper validate the potential of this new experimental paradigm as well as many of the design choices that were made. In order to reach conclusive results for the many research points discussed in this section, future word will focus on the development of this experiment at a larger scale.

Acknowledgments

Research project partly funded by ANR-11-JS03-005-01.

References

[1] Manon Raimbault. Simulation des ambiances sonores urbaines: intégration des aspects qualitatifs. *Ph. D. dissertation, Université du Maine, France*, 2002.

[2] Catherine Guastavino. Etude sémantique et acoustique de la perception des basses fréquences dans l'environnement sonore urbain. *Semantic and acoustic study of lowfrequency noises perception in urban sound environment"), Ph. D. dissertation, Université Paris*, 6, 2003.

[3] R. M. Schafer. *The Soundscape, Our sonic environment and the tuning of the world*. Rochester, Vermont, destiny books edition, 1977.

[4] Maria Niessen, Caroline Cance, and Danièle Dubois. Categories for soundscape: toward a hybrid classification. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 2010, page 5816–5829, 2010.

[5] William J. Davies, Mags D. Adams, Neil S. Bruce, Rebecca Cain, Angus Carlyle, Peter Cusack, Deborah A. Hall, Ken I. Hume, Amy Irwin, Paul Jennings, Melissa Marselle, Christopher J. Plack, and John Poxon. Perception of soundscapes: An interdisciplinary approach. *Applied Acoustics*, 74(2):224–231, February 2013.

[6] Brigitte Schulte-Fortkamp. Soundscape-focusing on resources. In *Proceedings of Meetings on Acoustics*, volume 19, page 040117, 2013.

[7] Valérie Maffiolo. De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain. *"Semantic and acoustic characterization of urban environmental sound quality" Ph. D. dissertation, Université du Maine, France*, 1999.

[8] Josh H. McDermott and Eero P. Simoncelli. Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis. *Neuron*, 71(5):926–940, September 2011.

[9] Israel Nelken and Alain de Cheveigné. An ear for statistics. *Nature neuroscience*, 16(4):381–382, 2013.

[10] A.L. Brown, Jian Kang, and Truls Gjestland. Towards standardization in soundscape preference assessment. *Applied Acoustics*, 72(6):387–392, May 2011.

[11] Manon Raimbault and Danièle Dubois. Urban soundscapes: Experiences and knowledge. *Cities*, 22(5):339–350, October 2005.

[12] Catherine Guastavino. The ideal urban soundscape: Investigatng the sound quality of french cities. *Acta Acustica United with Acustica*, 92:945–951, 2006.

[13] Catherine Guastavino. Categorization of environmental sounds. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 61(1):54, 2007.

[14] A. L. Brown. Advancing the concepts of soundscapes and soundscape planning. In *Acoustics 2011: Annual Conference of the Australian Acoustical Society*, 2011.

[15] Danièle Dubois, Catherine Guastavino, and Manon Raimbault. A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. *Acta Acustica united with Acustica*, 92(6):865–874, 2006.