A method to collect representative samples of urban soundscapes

Julien Tardieu, Cynthia Magnen
MSHS-T USR3414, University of Toulouse and CNRS, Toulouse, France.

Marie-Mandarine Colle-Quesada, Nathalie Spanghero-Gaillard
LERASS-CPST EA827, University of Toulouse, Toulouse, France.

Pascal Gaillard
CLLE UMR5263, University of Toulouse and CNRS, Toulouse, France.

Summary

Many studies of urban soundscapes have been performed with different aims ranging from auditory perception, to evaluation of noise annoyance, sound design for public places, speech learning, and urban planning, among others. A typical first step of most studies involves collecting soundscape samples, which has one major limitation: it is a selection from the scientist's own representation of the urban soundscape being studied. This paper addresses this issue with a method based on online questionnaire surveys. In this study, we collected French linguistic data through a large-scale questionnaire survey to find the shared knowledge about soundscapes in three types of urban situations at different times during the day. From a linguistic analysis of the responses, we elaborated profiles to be used to guide the recording of soundscapes or the selection in a database. These profiles consist of a distribution of different sound sources and human activities that is different for each type of situation. The proposed method and results as well as possible future applications are likewise presented.

PACS no. 43.90+v, 43.66.+y

1. Introduction

As introduced in the 70's by Canadian composer Murray Schafer [1], a soundscape is a complex concept that refers not only to the sonic environment (i.e., different sounds occurring in a given space configuration) but also to the multiple interactions that emerge from this sonic environment: interactions between sounds, between sounds and space, as well as between sounds, space and humans through perception and actions. In the case of urban soundscapes, many studies dealing with perception have been carried out with different aims ranging from noise annoyance [2-7], auditory cognition [8-12], sonification in public places [13], speech learning [14,15], urban planning [16], among many others. In most studies, soundscape perception is addressed in two different ways: *in situ* or in the laboratory. *In situ* studies have been using different methods such as interviews with passers-by [9,17] or soundwalks with selected participants either in groups or individually (for a review of related works see [18]). In the laboratory, one important issue is the type of technical equipment used to record soundscapes and to play the audio files. In [19], it has been demonstrated that ambisonics [20] was the more ecologically valid recording and sound reproduction technique for soundscape research as it provides the best results in terms of sound immersion.

As regards the laboratory research on the perception of urban soundscapes, the initial step consists of selecting samples to build a set of stimuli according to research purposes. These samples can be selected from sound databases, new recordings in the field, or from artificial soundscapes created with individual sounds. While this initial collection of samples is crucial insofar as it will have a strong impact on the validity of the study, there is one major limitation: it is a selection and therefore it is the result of a choice from the scientist's own assessment of what is representative of the urban soundscape being studied. Indeed, this selection is based on an individual mental representation of soundscapes. In [9] for example, ambisonic recordings were made in several train stations with the aim to collect representative samples of the different types of places that make up a train station (halls, platforms, waiting rooms, etc.). In this study, the

sample typicality was judged by the authors according to *a priori* choices based upon preliminary observations. To make this selection less arbitrary, it would be relevant to choose samples according to a shared representation of soundscapes, i.e., a knowledge that is shared between a group of individuals in a given language, a given culture and a given place at a given period of time. A representative sample of an urban soundscape would be an example that is as close as possible to this shared representation, i.e. very typical.

In addition to research purposes, this issue is also of interest in the field of sound design: in video games or movies soundtracks, sound designers often need to choose samples of soundscapes to reflect the sonic environment of a given situation. Usually these samples are chosen from large databases or recorded by the sound designers themselves. For example, for a pedestrian street in a large city at a given period of time, it is necessary to get a sample that will be recognized undoubtedly. In other words, it is necessary to know what is a highly representative soundscape sample of this situation.

In this paper, we suggest a method to better apprehend what is a typical soundscape sample of three different street configurations at different times during the day. This method, introduced in a previous work with individual interviews [14,15], relies on a questionnaire survey performed online in which participants are asked to describe what they usually hear in each urban situation. Each participant provides his/her own individual representation of the soundscape in each situation, leading to a representation that is shared between all the participants. The remainder of this article describes the questionnaire we created and the analysis of the results.

2. Questionnaire survey

2.1. Goal

The goal of this survey is to collect the knowledge shared by a group of French participants concerning the soundscape of some typical urban situations. Three different typical urban situations were chosen by the authors and are described in the following section.

2.2. Structure of the questionnaire

The questionnaire was written in French and was entitled *"Qu'entendez-vous en ville?"*, that can be translated into *"What do you hear in town?"*. In the first page, an introductory text presented the purpose of the questionnaire and informed the participant that all the collected data would remain anonymous. Then the participant was asked about his/her age and gender, and whether he/she was living in a city with more than 10000 inhabitants. The second page of the questionnaire was the main part of the survey. The participant was first given the following instruction (in its original formulation in French and then translated to English):

"CONSIGNE : décrivez précisément ce que vous entendez dans chacune des situations proposées ci-dessous" - *"INSTRUCTION: describe precisely what you hear in each of the situations presented below"*.

Then, the three situations were described and the participant was asked to describe precisely what he/she usually hears in each situation, using the keyboard of his/her computer.

Situation 1

"Imaginez que vous êtes dans une rue semi-piétonne en ville, entre 12h et 14h un jour de semaine, il fait beau. Décrivez précisément ce que vous entendez."

"Imagine yourself in a semi-pedestrian[2] street in town, between 12am and 2pm during the week, the weather is nice. Describe precisely what you hear."

Situation 2

"Vous marchez sur le trottoir d'une rue non-piétonne en ville, le matin vers 8 h, en semaine, il fait beau. Décrivez précisément ce que vous entendez."

"You are walking on the sidewalk of a non-pedestrian street in town, in the morning around 8am, during the week, the weather is nice. Describe precisely what you hear."

Situation 3

"Vous êtes dans la même rue que celle décrite en situation 2. Il est maintenant 21h, il fait toujours beau. Décrivez précisément ce que vous entendez."

"You are in the same street as the one described in situation 2. It is now 9pm, the weather is still nice. Describe precisely what you hear."

These three situations were chosen according to the authors' own experience of streets in town, with the aim to have three typical situations that are representative of what is usually found in cities larger than 10000 inhabitants. The situations were written precisely in order to help the participant to imagine being in each situation, as well as to induce some activities in which the participant

[2] A pedestrian street with limited access to some vehicles.

could be involved. The first situation represents the situation in which most people are when they have lunch in town during the week. The second situation represents a typical morning situation when walking in a street (e.g. going to work). Then the third situation represents the situation in which people are when they go out in the evening (e.g. to have dinner or a drink with friends). The indication about the nice weather was to avoid any description of weather related sounds (rain, wind, etc.).

2.3. Participants

60 participants (39 female, median age 31 years old) were recruited online to participate for this survey via the French volunteer database http://expesciences.risc.cnrs.fr/. The survey took place in May 2013. 49 participants reported to live in a city with more than 10000 inhabitants, and out of the remaining 11 participants, 7 reported to go in a big city between 2 to 7 times a week, 4 between 5 times a year to 2 times a month. 12 French regions were represented (out of 22), with a majority of participants originally from Ile-de-France and Midi-Pyrénées. 53 participants reported to have normal hearing, and 7 participants reported to have hearing losses.

2.4. Results

2.4.1. Linguistic data and analysis

Data consisted of verbal descriptions written in French by the participants on the keyboard of their computer. Since participants were not limited in their descriptions, the length of the description across all the participants varied from 1 to 85 words. In total, 1409 words, 989 words and 818 words were analyzed for situation 1, 2 and 3, respectively.

In addition to the length, participants used different types of linguistic expressions to describe the soundscapes: lists of sounds referring to sources for example (e.g. *cars, car horns*), but also full sentences with nouns, verbs and adjectives. The analysis used in this study is a semantic analysis based on functional linguistics [21,22]: a selection of forms corresponding to units of signification (words or groups of words). Starting from participants' descriptions of soundscapes, this analysis unfolds as follows:

1. Extract forms referring to sounds. 229 forms were extracted for situation 1, 152 for situation 2 and 175 for situation 3.
2. Group them according to semantic links, and deduce some categories. A form can be in two different categories if it refers to both.
3. Count the number of forms in each category.

This procedure is exemplified in Table 1 with the descriptions made by two participants of the situation 1 (translated from French). This table also presents the semantic categories deduced from the semantic links between the forms, grouped in two macro categories: sound sources and human activities. These categories are consistent with the ones found in previous works on urban soundscapes [19, 23] and train stations [9]. The results of this linguistic analysis are presented in Figure 1 for the sound sources and in Figure 2 for the human activities, respectively. These figures show the number of forms (words or groups of words) in each semantic category as a percentage of the total number of forms used by the 60 participants. The following sections describe the result for each situation.

	Sound sources						
Participant	Traffic	Warning signals	Phones	Music	Animals	Weather	Doors
1	motorbikes, cars, buses, planes far away	ambulance or police sirens			dog barks	wind	door slamming
2	cars at slow, trains, planes far away		people using mobile phones	music from cars	birds		
Total	2	1	1	1	2	1	1

	Human activities				
Participant	Walking	Conversations	Vocal sounds	Restaurants/Bars	Shops
1	footsteps	conversations	laughs, child shouts	cutlery sounds	
2		discussions at a cafe terrace		discussions at a cafe terrace	music from shops
Total	1	2	1	2	1

Table 1. Example of linguistic analysis with the descriptions made by 2 participants of the situation 1.

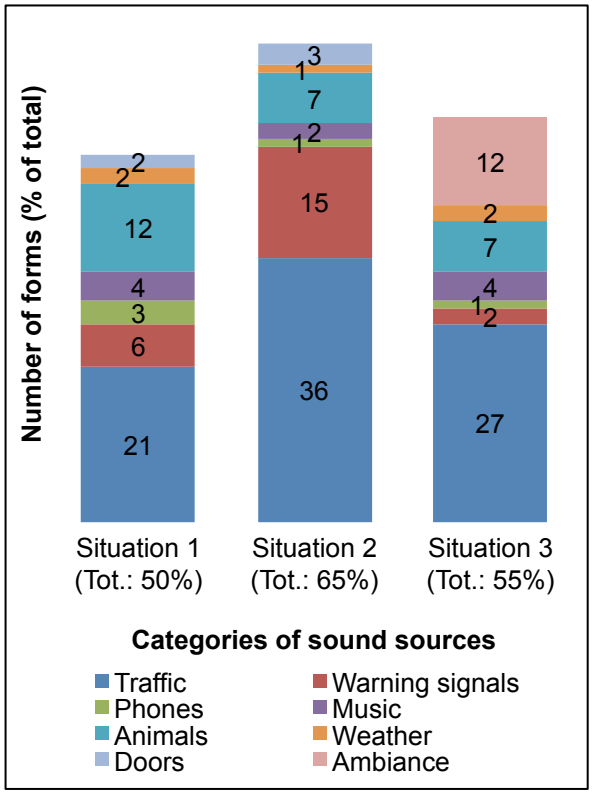

Figure 1. Number of forms (words or groups of words) in each category of sound sources and for each situation (% of total number of forms in each situation)

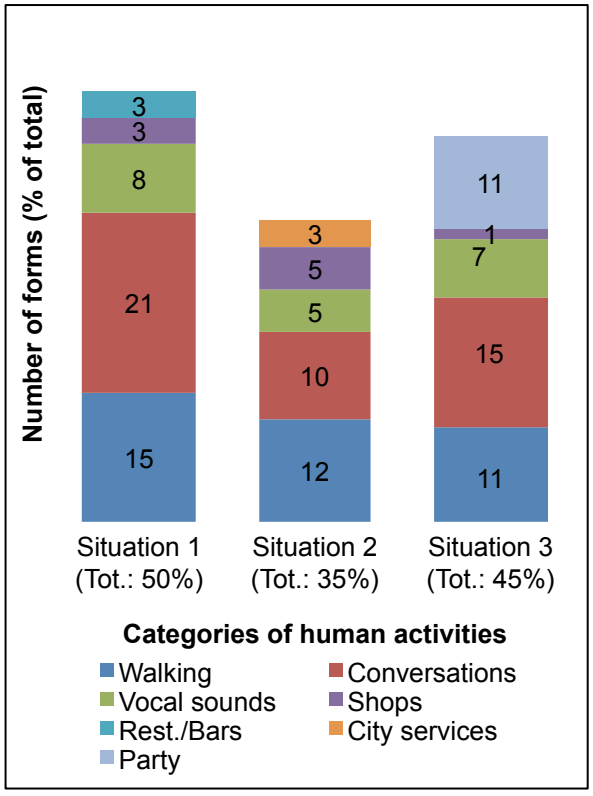

Figure 2. Number of forms (words or groups of words) in each category of human activities and for each situation (% of total number forms in each situation)

2.4.2. Distributions of categories

Figures 1 and 2 show that for situation 1, 21% of the 229 forms describing this situation and collected in this analysis (see previous section) concerned the traffic, 21% concerned conversations, 15 % walking, 12% animals, 8% vocal sounds, 6% warning sounds and 4% music (for the largest numbers of forms). As introduced in this paper, this distribution of descriptions along the different categories represents the typical soundscape of situation 1 for the 60 participants who answered the questionnaire. These figures also show different distributions for the two other situations, which demonstrates the specificities of each situation in the way they are described by the participants. It is also shown that situation 3 differs from the two other ones since two categories of descriptions were used only for this situation: forms about the ambiance (Figure 1) and forms about parties (Figure 2). Situation 2 is also the only one to be described with forms about the sounds of city services (see Figure 2). In the descriptions of situation 2, Figures 1 and 2 indicate a majority of sound sources in the description compared to human activities (65% versus 35%), whereas for situations 1 and 2 the distribution is more balanced.

2.4.3. Differences between categories

In addition to the differences in the distribution of each category between the 3 situations, it is also necessary to focus on the forms used in each category and the type of sound sources they refer to. Considering only the main categories used by the participants, these differences are summarized as follows (see also Table 2):

- **Walking**. Situations 1 and 2 are described mainly in terms of *footsteps sounds*, whereas in situation 3 participants talked about *passing pedestrians* much more than *footsteps*.
- **Conversation**. The three situations are described in terms of *conversations while walking* or *with a phone*. Situation 1 differs since it is the only one described also with forms about *the babble* (i.e., mix of many voices in the background with no intelligibility). Situations 1 and 3 are also described with forms about *conversations taking place at a café terrace*.
- **Traffic**. The linguistic data shows that participants referred to two types of traffic: either the *car traffic* as a global event or *cars passing* as individual events. For situations 1 and 3, the data shows a balance between the two types of traffic descriptions added to *bikes*

passing for situation 1 only. In Situation 2, the words about global traffic are predominant compared to isolated car passing by, added to *bus traffic*, *braking* and *acceleration sounds* and *traffic jam*.

- **Warning sounds**. For the situation 1, this category includes *car horns being far away* and *bike bells*, whereas for situation 2 car horns are described as being close.

The implication of these results in the selection of soundscape samples is discussed in the following section.

	Walking	Conversations	Traffic	Warning sounds
Situation 1	Footsteps	At a terrace, while walking, with phone, with babble	Car traffic = Cars passing, bikes	Car horns far away, bike bells
Situation 2	Footsteps	Walking, with phone (no babble)	Car traffic > car passing, braking, accelerating, traffic jam	Car horns close
Situation 3	Pedestrians passing by	At a terrace, walking, with phone (no babble)	Car traffic = Cars passing	-

Table 2. Specificities of forms (words or groups of words) used by the 60 participants to describe each situation for the main categories of sound sources and human activities.

2.5. Guidelines to select soundscape samples

As presented in the previous sections, participants' descriptions of what they usually hear in 3 typical urban situations referred to different semantic categories of words or groups of words, each situation having its own distribution of these categories as well as its specificities in the type of words used in each category. These results can be applied to soundscape sample selection in at least two ways. (1) These different distributions of sound sources and human activities can be used to guide future soundscape recordings in the field. For example, to record a typical soundscape of a semi-pedestrian street between noon and 2pm (situation 1 in this study), the place and the moment to record can be chosen accordingly to the distribution of sound sources and human activities provided in section 2.4. (2) These results can also be used to guide the selection of soundscapes samples either from a database (for a detailed review of available database see [24]) or from a soundscape design tool such as [25-27]. This selection can be performed by comparing the distribution of sound sources and human activities that are present in the sample with the distributions revealed in the shared representation in section 2.4. In addition, the specificities of each situation listed in section 2.4.3 provide some indications about the distance there should be between sounds and listening or recording points. Therefore, this method of selection will not only be helpful to select typical soundscapes but also to evaluate the typicality of a soundscape sample. To do so, further investigations need to be carried out in the annotation of soundscape samples in order to evaluate the distribution of the sound sources and human activities that are present in the sample.

3. Conclusions

This work presented a method to select urban soundscapes samples (as stimuli for an experiment on perception for example) based on a shared representation of soundscapes. This method was implemented in the case of three urban situations corresponding to three typical street configurations in a city larger than 10000 inhabitants. In an online questionnaire survey, 60 participants were asked to imagine themselves in each situation and to express what they hear. The resulting linguistic data revealed the semantic categories employed by the participants to describe the situations, grouped in two macro categories: sound sources (e.g., traffic or animals) and human activities (e.g., walking or conversations). Different distributions of these categories were found for each situation as well as specificities of each situation in the type of sound sources each category refer to. The data collected in this study provide guidelines that can be used in urban soundscape research to collect, select or even create urban soundscapes. This study suggests further investigations in soundscape annotation as well as future applications of this method to different types of soundscapes.

Acknowledgement

This study is part of the CIESS project (http://petra.univ-tlse2.fr/ciess/) funded by the French research agency, (Agence Nationale de la Recherche, ANR).

References

[1] Schafer RM. The tuning of the world. Random House Inc.; 1977.

[2] Brambilla, G., Gallo, V., Asdrubali, F., and D'Alessandro, F., 2013, "The perceived quality of soundscape in three urban parks in Rome.," J. Acoust. Soc. Am., 134(1), pp. 832–9.

[3] De Coensel, B., Boes, M., Oldoni, D., and Botteldooren, D., 2013, "Characterizing the soundscape of tranquil urban spaces," Meetings on Acoustics, Montreal, p. 040052 (19).

[4] Jeon, J. Y., Lee, P. J., You, J., and Kang, J., 2010, "Perceptual assessment of quality of urban soundscapes with combined noise sources and water sounds.," J. Acoust. Soc. Am., 127(3), pp. 1357–66.

[5] Kang, J., and Zhang, M., 2010, "Semantic differential analysis of the soundscape in urban open public spaces," Build. Environ., 45(1), pp. 150–157.

[6] Yang, W., and Kang, J., 2005, "Acoustic comfort evaluation in urban open public spaces," Appl. Acoust., 66(2), pp. 211–229.

[7] Botteldooren, D., De Coensel, B., and De Muer, T., 2006, "The temporal structure of urban soundscapes," J. Sound Vib., 292(1-2), pp. 105–123.

[8] Dubois, D., Guastavino, C., and Raimbault, M., 2006, "A Cognitive Approach to Urban Soundscapes: Using Verbal Data to Access Everyday Life Auditory Categories," Acta Acust. united with Acust., 92(6), pp. 865–874.

[9] Tardieu, J., Susini, P., Poisson, F., Lazareff, P., and McAdams, S., 2008, "Perceptual study of soundscapes in train stations," Appl. Acoust., 69(12), pp. 1224–1239.

[10] Botteldooren, D., and Coensel, B. De, 2009, "The role of saliency, attention and source identification in soundscape research," Inter-Noise 2009.

[11] Nielbo, F. L., Steele, D., and Guastavino, C., 2013, "Investigating soundscape affordances through activity appropriateness," Meetings on Acoustics, Montreal, p. 040059 (19).

[12] Oldoni, D., De Coensel, B., Boes, M., Rademaker, M., De Baets, B., Van Renterghem, T., and Botteldooren, D., 2013, "A computational model of auditory attention for use in soundscape research.," J. Acoust. Soc. Am., 134(1), pp. 852–61.

[13] Tardieu, J., Susini, P., Poisson, F., Kawakami, H., and McAdams, S., 2009, "The design and evaluation of an auditory way-finding system in a train station," Appl. Acoust., 70(9), pp. 1183–1193.

[14] Colle-Quesada, M.-M., & Gaillard, P. (2012). How to create soundscapes in order to make sense for the hearer: a questionnaire approach. In SCANDLE: acoustic SCene ANalysis (Ed.). Plymouth (UK).

[15] Colle-Quesada, M.-M., & Spanghero-Gaillard, N. (2014). Étude de l'impact du contexte sonore environnemental sur la compréhension d'un dialogue verbal : une démarche expérimentale. Cel. de Rech. En Linguistique, 51–67.

[16] Raimbault, M., and Dubois, D., 2005, "Urban soundscapes: Experiences and knowledge," Cities, 22(5), pp. 339–350.

[17] Raimbault, M., Lavandier, C., and Bérengier, M., 2003, "Ambient sound assessment of urban environments: field studies in two French cities," Appl. Acoust., 64(12), pp. 1241–1256.

[18] Jeon, J. Y., Hong, J. Y., and Lee, P. J., 2013, "Soundwalk approach to identify urban soundscapes individually.," J. Acoust. Soc. Am., 134(1), pp. 803–12.

[19] Guastavino, C., Katz, B. F. G., Polack, J.-D., Levitin, D. J., and Dubois, D., 2005, "Ecological Validity of Soundscape Reproduction," Acta Acust. united with Acust., 91(2), pp. 333–341(9).

[20] Gerzon, M. A., 1985, "Ambisonics in Multichannel Broadcasting and Video," J. Audio Eng. Soc, 33(11), pp. 859–871.

[21] Martinet, A. (1985). *Syntaxe générale*, Paris : Armand Colin.

[22] Pottier, B. (1987). *Théorie et analyse en linguistique*. Paris: Hachette.

[23] Maffiolo V, Dubois D, David S, Castellengo M, Polack J-D., 1998. Loudness and pleasantness in structuration of urban soundscapes. Inter-noise 1998.

[24] Shafiro, V., and Gygi, B., 2004, "How to select stimuli for environmental sound research and where to find them," Behav. Res. Methods, Instruments, Comput., 36(4), pp. 590–598.

[25] Accolti, E., and Miyara, F., 2015, "Method for generating realistic sound stimuli with given characteristics by controlled combination of audio recordings.," J. Acoust. Soc. Am., 137(1), p. EL85.

[26] Janer, J., Kersten, S., Schirosa, M., and Roma, G., 2011, "An online platform for interactive soundscapes with user-contributed content," AES 41st Int. Conf. on Audio for Games.

[27] Lafay, G., Rossignol, M., Misdariis, N., Lagrange, M., and Petiot, J.-F., 2014, "A New Experimental Approach for Urban Soundscape Characterization Based on Sound Manipulation : A Pilot Study," ISMA 2014.