ACTA ACUSTICA UNITED WITH ACUSTICA
Vol. 91 (2005) 333 – 341

Ecological Validity of Soundscape Reproduction

Catherine Guastavino[1], Brian F.G. Katz[2], Jean-Dominique Polack[3], Daniel J. Levitin[1], Danièle Dubois[4]

[1]: Centre For Interdisciplinary Research on Music Media and Technology and McGill University, Department of Psychology, 1205 Dr. Penfield Avenue, Montreal, QC, H3A 1B1, Canada. Catherine.Guastavino@mcgill.ca

[2]: LIMSI-CNRS, Perception Située, BP 133, F91403, Orsay, France

[3]: LAM (Laboratoire d'Acoustique Musicale), Université Paris 6-CNRS, 11 rue de Lourmel, F75015 Paris

[4]: LAM, Equipe Langage Cognition Pratiques, Université Paris 6-CNRS, France

Summary
We introduce a methodology based on linguistic exploration of verbal data to investigate the influence of reproduction method on cognitive processing of environmental sounds in laboratory conditions. Three experiments were carried out to explore the ecological validity of reproduction systems. The reference study consisted of interviews conducted in actual environments, which were also recorded simultaneously. The recordings were used for two listening tests, the first one using stereophonic reproduction and the second one using multichannel reproduction. The comparison of the verbal data collected in the different contexts sketches some theoretical and methodological issues concerning the reproduction of everyday life scenes in laboratory conditions. The linguistic analyses indicate that the "same" acoustic phenomenon gives rise to different cognitive representations, depending on the spatial presentation of the stimuli. It follows that the quality of the reproduction system must be adapted to specific properties of mental representations (here, spatial immersion vs. source identification). On methodological grounds, the analysis of spontaneous language representations gives access to cognitive representations elaborated in real life situations and in experimental conditions. The comparison of the linguistic exploration can then be used as a psycholinguistic measure of the ecological validity of experimental settings.

PACS no. 43.38.Md, 43.66.Lj, 43.50.Lj, 43.50.Rq

1. Introduction

A fundamental aim of psychoacoustics is to better understand how acoustic phenomena are perceived and represented at a cognitive level by individuals. Mental representations of sounds cannot be observed directly, but one way to study these representations empirically is through language, specifically, by analyzing how people talk about their sensory experiences.

1.1. From discourse to cognition

From what is being said and how is it being said, linguistics and psycholinguistics aim at deriving relevant inferences about how people process and conceptualize sensory-perceptual experiences through the use of discourse analysis techniques [1, 2]. However, the relationship between language and cognition is complex, since words and concepts cannot be mapped with a one-to-one correspondence [3, 4]. The accuracy of the inferences relies therefore on the elaboration of both linguistic and cognitive theory [5]. Furthermore, the lack of basic lexicalized terms [6] or *a priori* established categories for acoustic phenomena calls into question the relationship between words and representations in the auditory modality. If visual objects can often be described by simple lexical devices, there are few single words on which people agree as spontaneous descriptions of sounds and noises[1] [7]. Discourse analysis conducted on free descriptions of domestic noises (see [8] for a review) and environmental sounds [9, 10, 11, 12, 13] reveals a large variety of linguistic devices. Therefore, a linguistic analysis of complex phrasings, rather than a lexical analysis of words, is necessary to infer properties of mental representations of acoustic phenomena. There is strong evidence that linguistic constraints contribute to the elaboration of cognitive categories (see [5] for examples in the auditory and olfactory domain). The present study relies on the analysis of the psycholinguistic processes that mediate between individ-

Received 25 November 2003,
accepted 28 September 2004.

[1] With the exception of specific terms for human sounds and music.

ual representations of sensory non-verbal experiences and shared conceptual cognitive representations, conveyed in language.

1.2. From everyday situations to laboratory conditions

The notion of ecological validity was introduced by Gibson [14] in the visual domain, to express the need to study perception under ecological conditions, *i.e.* to take contextual and environmental cues into consideration. An experimental protocol is ecologically valid if the participants react, to some extent, as if they were in a natural situation. In other words, the laboratory must be like life in regards to the context of the question asked. From a methodological viewpoint, instructions must be given to direct the subjects' response strategy towards an everyday situation in order to enable the reactivation of cognitive processes elaborated in actual situations. Furthermore, the stimuli must be selected and presented in such a way that subjects will recognize and treat the test samples as natural or potentially familiar experiences. The present research investigates the influence of the reproduction method, and more specifically spatial presentation, on the cognitive processing of sound material during listening tests. Multichannel spatial sound reproduction offers new possibilities for the study of auditory perception and cognition in complex sonic environments [15]. However, stereophonic or even monophonic reproduction are still widely used in psychoacoustic research. Maffiolo and Vogel [12] conducted a set of studies to determine the stereophonic recording setup most adapted to urban environments. Subjects listened to pairs of recordings of outdoor environments carried out with setups differing in spacing, angle and directivity of microphones. The task was to select the version which sounded the most like their everyday life experience. Results indicate a strong preference for widely spaced microphones. This highest rated configuration used cardioids with an angle of 100° and a separation of 60 cm. Based on these findings, a pair of microphones with the same spacing was employed in the present research. Omnidirectional microphones were chosen because of their improved response at low frequencies due to the fact that low frequency content was an important issue in this study.

1.3. Overview

Three experiments were carried out to investigate ecological validity of soundscape reproduction in laboratory conditions. The same open questionnaire was used in three listening contexts. The reference study (Experiment 1) consisted of 42 interviews carried out in actual environments at locations identified as representative of city noises in previous studies [16, 13]. The verbal data collected in real-life situations were then compared with that collected during two listening tests to test the ecological validity of experimental settings in an acoustically damped room. Experiment 2 used stereophonic reproduction, with 29 subjects listening to six different soundscapes. Experiment 3 used multichannel Ambisonics reproduction with a separate group of 27 subjects listening to five different soundscapes.

Comparison of spontaneous descriptions collected in different contexts raises theoretical and methodological issues concerning the reproduction of familiar auditory scenes, as the same acoustic phenomenon gives rise to different cognitive objects depending on the "quality" of the reproduction system, and more specifically on the spatial presentation.

2. Experiment 1: Field survey

2.1. Method

2.1.1. Subjects

Forty-two interviews were conducted in two large French cities (Paris and Nantes). The participants, aged 16 to 59, served without pay. They were selected on the basis of familiarity with urban soundscapes: they had all been living or working in a large city for at least five years.

2.1.2. Questionnaire

The same semi-structured questionnaire was used in all three experiments. It was constructed in a progressive way, starting with very general questions about the appraisal of urban soundscapes and ending with more specific ones about transportation noise and low frequency components. Ten open questions were formulated in French with very general terms ('*feel, be affected*') in order not to influence the judgment or confine the answers in predefined categories.

2.1.3. Sound material

The locations were selected from a list of public places identified as representative of city noises in previous research. Maffiolo *et al.* [16] conducted a survey in which Parisians were asked to freely describe the sounds of Paris. In a similar study, Raimbault [13] investigated typical soundscapes for the city of Nantes. Examples of such places include the Place du Commerce (Nantes) and the Boulevard Sébastopol (Paris). The interviews were carried out in 8 representative soundscapes, presented in Table I. The soundscapes were recorded simultaneously with the interviews for use in Experiments 2 and 3. The average sound pressure level (15 min Leq) ranged between 74 dB SPL and 78 dB SPL (mean, 75.5; SD, 2.4).

2.1.4. Procedure

Subjects were questioned about their appraisal of the sound environment with an oral version of the open questionnaire used in all three experiment.

2.2. Results

Two broad categories were derived from the linguistic analysis: *source events*, which can be attributed to an identified source and agent (*truck, bus* "*Camions, bus*"), and *background noise of the city* ("*Bruit de fond de*

Table I. Description of sound material.

Sample	Exp.	Place	Description
1	1-2	Downtown, Nantes	Open near major road
2	1-2	Montparnasse, Paris 15	Park near road and train station
3	1-2	Dupleix, Paris 7	Quiet street close to schoolyard
4	1-2-3	Beaubourg, Paris 3	Walled-in plaza in pedestrian area
5	1-2-3	Sebastopol, Paris 4	Sidewalk of major road
6	1-2-3	Montorgueil, Paris 2	Sidewalk café-disctrict in pedestrian area
7	1-3	Montsouris, Paris 15	Large park near road and train station
8	1-3	Notre Dame, Paris 1	Open square near major road

Table I. Continuation. Description of sound sources.

Sample	Traffic noise	Human sounds	Other sources
1	Distant traffic	Lots of people talking, footsteps	Tramway
2	Cars and buses moving	Children playing, footsteps	Train, birds
3	Car moving, faint traffic	Children playing	Plane
4	Motorcycle, horns, distant traffic noise	Faint voices	Music (mechanical organ), birds
5	Cars, motorcycles, buses	A few people walking	
6	Very faint traffic	People talking, glasses/dishes	Faint music (speakers), birds
7	Car moving, distant traffic noise	Children playing, footsteps	Train, birds, construction work
8	Distant traffic	Lots of people talking, footsteps	

Table I. Continuation. Description of acoustic parameters. The Leq describes the average sound pressure level over the duration of the recordings, as measured in the room at the position of the listener with a sound analyzer CEL 328, over 500ms time periods (L=linear unweighted levels, and A=A-weighted levels). LN95 is the sound level exceeded for 95% of the time, which estimates background ambient sound level.

Sample	Exp.	Leq (dB L)	Leq (dB A)	LN 95% A	Duration (min)	Mean response time (min)
1	2	75	67	64	5.05	5.05
2	2	75	66	65	5.85	5.85
3	2	77	72	70	5.75	5.75
4	2	73	67	61	5.2	5.2
	3	72	66	59	0.71	14.2
5	2	77	67	65	5.71	5.71
	3	78	68	66	1.28	16.3
6	2	74	65	63	4.6	4.6
	3	74	64	62	0.61	13.1
7	3	75	66	63	0.53	17.3
8	3	78	73	71	0.7	12.4
Average Exp. 2	2	75.2	67.3	64.6	5.3	5.3
Average Exp. 3	3	75.4	67.4	64.2	0.76	14.6

la ville"), considered as collective noise, where no specific events can be discriminated. As the category name suggests, *source events* were primarily described with reference to specific sources generating noise, such as *cars* ("*Voitures*"). The number of occurrences was made proportional to the number of participants, to enable a comparison with the other experiments. Descriptions referering to the source producing noise represented 76% of occurrences[2]. In the descriptions of the *background noise*, however, there are few references to the source itself (15%). The verbal descriptions refer either to the effects on the subject, as in the case of the source events, or to the physical properties of the acoustic signal. The verbal data were classified into semantic categories emerging from the answers, namely **subject-centered** and **object-centered** descriptions. **Subject-centered** descriptions (86%) refer to the subjects by means of deverbal adjectives (*comforting* "*Rassurant*") and complex phrases (*has a negative impact on me* ("*A un impact négatif sur moi*") constructed on verbs indicating psychological effects. **Object-centered** descriptions (60%) refer to the acoustic phenomenon itself. They include simple adjectives[3] and denominations

[2] One participant's response may give to several occurrences.

[3] Simple adjectives in French, many of which translate into English by means of constructed forms, e.g. "*sourd*" (muted or muffled) or "*grave*"

referring to the physical properties of the acoustics signal in terms of frequency (*low-pitched* "*Grave*"), temporal structure (*continuous, monotone* "*Continu, monotone*") or level (*loud*, *intensity* "*Fort, intensité*").

The linguistic exploration of the verbal data is in agreement with previous research investigating cognitive representations of urban soundscapes [11, 17, 9, 7, 5, 12]: the perception of *source events* is mediated by the objects producing the noise, whereas *background noise* is described both in terms of properties of the sound itself and of its effects on the subjects, rather than with reference to the source. The results of the linguistic exploration identified in the field survey serve as a reference when testing the ecological validity of the experimental settings in Experiments 2 and 3.

3. Experiment 2: Stereophonic Listening Test

3.1. Methods

3.1.1. Subjects

Twenty-nine subjects participated in the experiment without pay. They were students and academic staff, aged 22 to 59, who had all been living in a large city for at least five years.

3.1.2. Sound material

The sound samples were recorded during the field survey (Experiment 1) in locations identified as representative of Paris and Nantes in [16, 13] respectively. The recording setup consisted of two omnidirectional microphones (Sennheiser MKH20) with a spacing of 60 cm [12]. The six recordings, presented in Table I, were five minutes in duration. They were reproduced at their original level (mean 75.2 dB; 67.3 dB A) in an acoustically damped room. The room was acoustically isolated (floated construction) with internal dimensions 2.77 × 3.24 × 3.62 m. The room had a flat frequency response and a reverberation time of <0.15 seconds for frequencies above 200 Hz. Below 200 Hz the reverberation time increased gradually to 0.4 seconds at 40 Hz. The recordings were played on a DAT player (Sony PCMR300) and presented on a 2.1 format with two speakers (Studer A1) and one subwoofer (JBL 4546C, 100 Hz crossover), all visible to the subjects. The speakers were 1.2 m from the listener, with an angular separation of 60°. The use of the subwoofer was one factor of the test where only half of the test samples used it.

3.1.3. Procedure

Subject were instructed to imagine themselves in an actual urban outdoor environment. For each sound sample, they were asked to spend a few minutes acclimating themselves to the recreated acoustic experience, and then to answer the same open questionnaire (written version) as in Experiment 1.

(low-pitched). This observation suggests the need for more precise analyses for the diversity of languages and further theoretical elaborations about interactions between language and cognition.

3.2. Results

The first observation concerns the use of the subwoofer. Most subjects remarked on the quality of the low frequency reproduction even when the subwoofer was not actually in use, *i.e.* had no signal. The conclusion is that the visual setting affected their impression of low frequency content. Subsequently, the listening room was redesigned for the next experiment so as to remove the visual bias, hiding the loudspeakers from view.

The linguistic analysis of the descriptions of *source events* yielded results similar to the reference study. As in the previous experiment, *source events* were described in reference to the source producing the noise (150%[4]), but subjects used more more generic terms (*vehicles*) than in the field survey, where they named specific vehicles. The number of occurrences in each category were compared to the reference study with the statistical χ^2 test. No significant difference was observed ($\chi^2(2) = 0.87, p > 0.05$). It should be noted however, that fewer descriptions of *source events* were collected in the field survey as subjects' responses were generally shorter on-site than in the laboratory. The comparison of *source events* descriptions collected in Experiments 1 and 2 is presented in Figure 1a, where the number of occurrences has been made proportional to the number of participants in each study.

This confirms that the stereo set-up used in earlier studies [18, 12], along with instructions given to the subjects, is ecologically valid for source identification. However, regarding *background noise*, the results are quite different. As before, few references to the sources (12%) were collected, and the descriptions could be classified into object-centered (59%) and subject-centered (19%) descriptions. The number of occurrences in each category was compared to the reference study with the statistical χ^2 test, revealing a significant difference ($\chi^2(2) = 25.01, p < 0.001$). The comparison of descriptions of *background noise* in Experiments 1 and 2 is presented in Figure 1b. It can be seen that a majority of the descriptions collected in the field survey were subject-centered in nature. On the contrary, in the stereophonic listening test, the *background noise* was primarily described in terms of objective properties (physical parameters), rather than subjective effects, as in the field survey. Further linguistic analysis shows that 27% of the subjects explicitly attributed the *background noise* to technical considerations about the reproduction system (e.*g. background noise of the speakers*) rather than urban ambient noise. This observation confirms the need to hide the speakers from view for soundscape reproduction in laboratory conditions. Moreover, 30% of the subjects complained that they did not feel as if they were "there", and that the envelopment provided by the subwoofer was not consistent with the frontal sound of the two stereo speakers. Previous research on urban noise had already stressed the importance of spatial attributes in the perception of urban *background noise*, which is spontaneouly described as very surrounding in questionnaire

[4] One subject's answer may give rise to several occurrences.

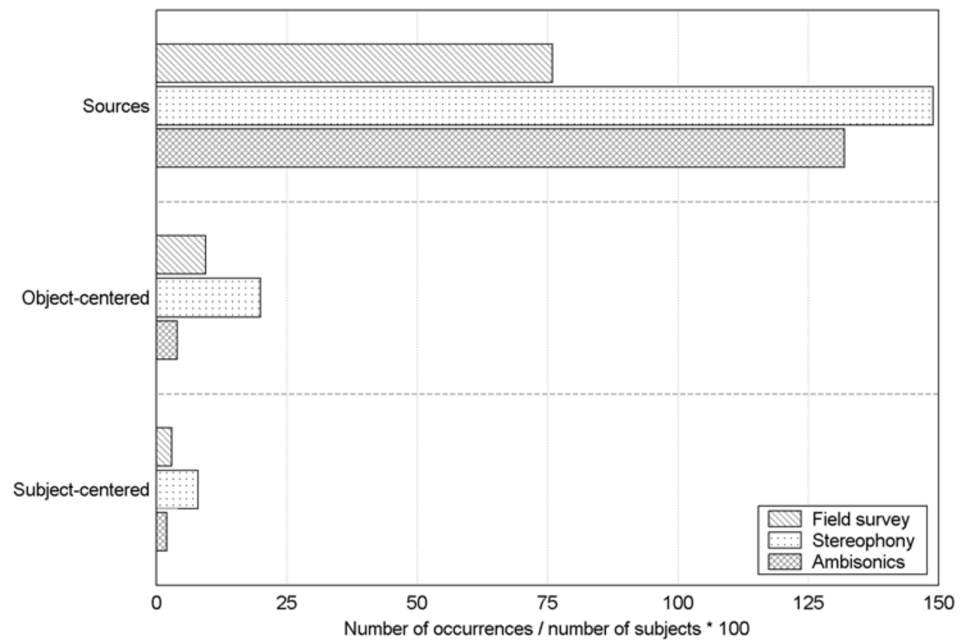

Figure 1. a. Comparison of spontaneous descriptions of *source events* classified into semantic categories emerging from the subject responses in the field survey (Experiment 1), the stereophonic listening test (Experiment 2) and the Ambisonics listening test (Experiment 3). No significant difference was observed between the three experiments.

studies [11, 19]. Together, these observations demonstrate that the frontal soundfield recreated by the stereophonic setup does not enable the subjects to process the *background noise* in an ecologically valid way. It furthers suggest that spatial immersion contributes to the cognitive representation of urban *background noise*.

Experiment 3 was designed to test this hypothesis by overcoming the limitations of the stereophony using multichannel reproduction and hiding the loudspeakers from view. If the verbal data collected in Experiment 3 are consistent with the reference study, it will entail that spatial immersion is a necessary condition to ensure the ecological validity of the experimental setting.

4. Experiment 3: Ambisonics Listening Test

Experiment 3 was designed to better understand the influence of spatial presentation of cognitive processes in laboratory conditions. Before answering the open questionnaire, subjects were presented with a reproduction of the same sound scene over four spatial reproduction systems and asked to choose which one(s) sounded the most like their everyday experiences. The subsequent questions were identical to Experiment 1 and 2.

4.1. Methods

4.1.1. Subjects

A new group of 27 subjects participated in the experiment without pay. They were students and academic staff, aged 23 to 57, who had all been living in a large city for at least five years.

4.1.2. Ambisonics

There are various approaches for recording and reproducing spatially distributed audio. The recording industry has developed a wide range of methods over the years starting from 2-channel stereo, to 4-channel quadraphonic, and the current trend of 6-channel 5.1. The aim of the present experiment is to investigate the subjective influence of spatial immersion on cognitive processes in laboratory condition. Ambisonics was chosen as the best suited method for this study as it provides a strong feeling of immersion in the auditory scene. Guastavino [20] conducted a subjective evaluation of various multichannel reproduction methods (including 5.1 and transaural), and Ambisonics was judged as the most enveloping and true-to nature for the recreation of outdoor environments[5].

Ambisonics is an approach to soundfield recording and reproduction that decomposes the spatial soundfield into spherical harmonics. A compact complex microphone is used, containing near-coincident elements, which produces direction dependent signals corresponding to the spherical harmonics. Currently available first order microphones provide a decomposition of the sound field correspond to the spherical harmonics up to and including first

[5] It should be noted that the subjective evaluations of the different methods highly depended upon the source material. Other reproduction systems were preferred for indoor environments and musical passages. The results highlight the fact that there is no single system optimal for all conditions (see [15] for a more detailed investigation).

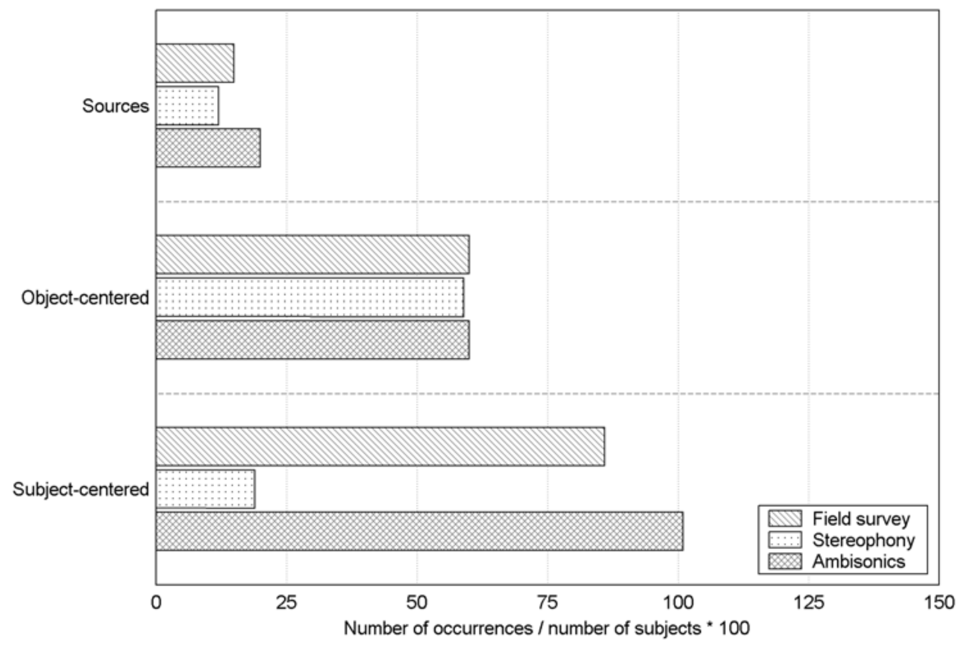

Figure 1. b. Comparison of spontaneous descriptions of *background noise* classified into semantic categories emerging from the subject responses in the field survey (Experiment 1), the stereophonic listening test (Experiment 2) and the Ambisonics listening test (Experiment 3). A significant difference was observed between the field survey (Experiment 1) and the stereophonic listening test (Experiment 2). No significant difference was observed between the field survey (Experiment 1) and the Ambisonics listening test (Experiment 3).

order: W (0th order omnidirectional) and XYZ (three first order components representing the Cartesian axis with figure of 8 directivity patterns). This output is in a 4-channel recording, termed B-format, that captures the spatial information of the sound field, resolved into a mono reference signal and left-right, front-back, and up-down information. Reproduction of the sound entails a decoding process from the B-format signal to the array of loudspeakers. The decoding process results in a signal to each loudspeaker being composed of a combination of the spherical harmonics, dependent upon the location of the speaker. There are various parameters in the decoding process whose discussion is beyond the scope of this paper (see [21, 22, 23] for more detailed discussions of this method).

4.1.3. Sound material

All recordings used were made with a B-format model Soundfield ST250 microphone, in the same locations as Experiments 1 and 2 (see Table I). The five recordings were 32 to 77 seconds long. They were reproduced at their original level (mean, 75.4 dB L; 67.4 dB A). The B-format files were decoded using the full in-phase decoding scheme without shelf filtering [23] on a 2-D and 3-D arrays of speakers (Studer A1) regularly spaced at 1.2 m from the listener. The omnidirectional channel W was simultaneously low-passed filtered at 100 Hz and then output through a subwoofer (JBL 4546C). In response to the apparent contribution of the visual aspects, the loudspeakers were hidden from view by means of acoustically transparent panels (for a detailed presentation of the setup, see [15]). The acoustics of the room were re-designed to be as dry as possible, given architectural limitations, in order to allow for the multichannel reproduction of outdoor sound scenes. The re-designed room had a flat frequency response and a reverberation time of < 0.05 seconds for frequencies above 200 Hz. Below 200 Hz the reverberation time increased gradually to 0.2 seconds at 40 Hz.

The reproduction system provided for 13 channels of discrete playback including a low frequency subwoofer for frequencies below 100 Hz. The 12 speakers (low frequency roll-off at $\sim$100 Hz) were suspended on a pipe grid that encircled the room and extended from floor to ceiling behind the visual screens. The subwoofer (flat response to 20 Hz) was placed in one corner of the room.

The test configurations were the 2-D and 3-D arrays with and without the subwoofer (x and x.1) following the familiar 5.1 convention. The 2-D configurations, 2-D and 2-D.1, consisted of six speakers located at seated listening level at the corners of the hexagon. The 3-D configurations, 3-D and 3-D.1, included six additional speakers, placed in two sets, three at ceiling level and three at floor level, corresponding to +/-49° on alternating sides of the hexagon (see Figure 2). This provided slightly reduced coverage for elevation sounds and full horizontal coverage in the listening plane.

The level of the speakers was carefully adjusted to achieve a flat frequency response across the crossover frequency of 100 Hz. The 12 speakers were time and level aligned at the center of the listener position. The low-frequency level was equalized at the center of the listening

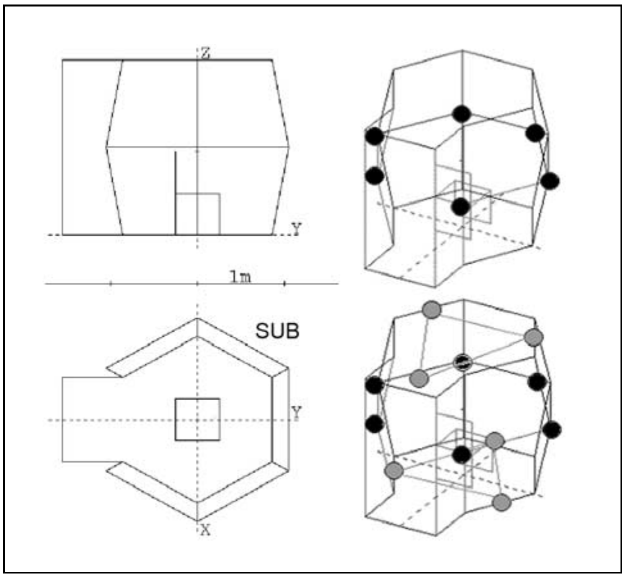

Figure 2. Listening room (adapted from [15]). Loudspeaker locations for the 2-D and 3-D arrays are represented by black dots. A chair is included for reference.

position using a reverberant room recording of white noise decoded over the speakers/subwoofer set-up. The subwoofer channel content was identical between 2-D.1 and 3-D.1 configurations and level matched to provide a flat frequency response over the crossover region. The recordings were played from a computer using a ADAT sound card (DIGI9636 Hammerfall) and external 24-bit DA converters (RME ADI-8 pro and Frontier Design Tango 24).

4.1.4. Procedure

Instructions were given to subjects to direct their response strategy towards everyday listening situations. The experiment consisted of five sessions corresponding the five soundscapes presented in Table I. In each session, subjects used a graphical interface to play the same soundscape reproduced on four different speaker configurations, randomly ordered. The nature of the test and the details of the reproduction system used were not disclosed to the subject prior to the test. The first task was to listen to the four version as many times as desired and select the one(s) that sounded the most like their everyday experiences. Following this comparison task, subjects answered the open questionnaire used in Experiment 2 and 3. They could play the selected versions of the soundscapes as many times as needed to acclimate themselves to the recreated sonic environment.

4.2. Results

4.2.1. Comparison

The results of the comparison test showed a strong preference for the 2-D configurations over other methods. Total results for the naturalness selection for the four reproduction setups were 62 (2D), 45 (2-D.1), 42 (3-D), and 20 (3-D.1). These results suggest potentially negative effects linked to providing too much irrelevant spatial information. These findings were replicated in a more detailed perceptual evaluation of spatial reproduction systems using a wider ranger on sound samples beyond the scope of this paper (see [15]): 2-D configurations were preferred for the recreation of outdoor environments, whereas 3-D configurations appeared to be more preferred for the recreation of indoor environments.

4.2.2. Verbal data

The linguistic analysis of the free responses yielded the following results. As regards *source events*, results were similar to the two previous experiments reported here: 132%[6] of descriptions were given in terms of sources. The number of occurrences were compared to the reference study and no significant difference was observed ($\chi^2(2) = 5.73$, $p > 0.05$) (see Figure 1a). As regards *background noise*, few references to the sources were collected as before (20%). The remaining verbal descriptions were classified into object-centered (60%) and subject-centered (101%) descriptions. The numbers of occurrences in each category were compared to the reference study test and no significant difference was observed ($\chi^2(2) = 0.75$, $p > 0.05$). The comparison of spontaneous descriptions is presented in Figure 1b. It can be seen that the proportion of object-centered and subject-centered descriptions do not differ significantly between the field survey and in the Ambisonics listening test. In both cases, the *background noise* was primarily described in terms of psychological effects on the subjects and to a lesser extent in terms of physical properties of the signal, in contrast to the stereophonic listening test in Experiment 2.

These results confirm the fact that spatial immersion contributes to the cognitive representation of *background noise*. Hence, it can be concluded that the Ambisonics multichannel reproduction in a neutral visual environment (speakers hidden from view) enabled the participants to process the sound samples as if they were in actual outdoor environments. In other words, the experimental setting could be considered as ecologically valid for the study of urban *background noise*. It should be noted, however, that the procedure slightly differed between the two laboratory conditions since subjects could to play the sounds samples several times in Experiment 3 but only once in Experiment 2.

It can further be seen in Figure 1b that the same proportion of object-centered descriptions was obtained in all experimental conditions. This suggests that cognitive processing of the *background noise* as an abstract sound **object** is not affected by the listening context. It is only the subjective **effects** that differ among experimental conditions, and more specifically the meaning attributed to the acoustic phenomena. Similarly, the proportion of descriptions in terms of sources is similar in all three conditions, confirming the ecological validity of both laboratory conditions for source identification.

[6] The number of occurrences has been made proportional to the number of participants in each study.

5. Discussion

Previous research on the perception of environmental sounds indicates that the cognitive processes operate on the basis of source identification ([24, 25, 26, 8, 5, 7], see [20] for a review). Semantic features linked to the appraisal of the object generating the sound often determine the judgment of the auditory events [13, 20, 27]. However, it has been shown that in the presence of numerous sound sources hindering the process of source identification, urban soundscapes are processed as a whole rather than as independent sound events [12, 13, 11, 19]. They are then classified as *background noise*, considered as collective noise characteristic of urban areas. In this case, physical properties of the acoustic signal (object-centered descriptions) as well as semantic features and psychological effects (subject-centered descriptions) are spontaneously evoked by subjects, suggesting the importance of both perceptual and conceptual features.

In the present research, spontaneous descriptions of urban soundscapes were collected and contrasted in three listening contexts: a field survey and two listening tests using spatially different reproduction schemes. To measure the ecological validity of the reproduction system in experimental settings, the verbal data collected during the two listening tests were compared to those collected in the field survey, which served as the reference context. Similar descriptions of *source events* were obtained independent of the context, with reference to the object generating the acoustic phenomena. Hence, both laboratory conditions can be considered ecologically valid in terms of source identification. However, the descriptions of the *background noise* differed significantly in the first laboratory condition, suggesting that different cognitive processes were involved. In the stereophonic listening test with visible speakers, the *background noise* was attributed to technical considerations of the reproduction system rather than ambient urban noise. It was described in terms of physical properties as an acoustic signal rather than meaningful collective noise characteristic of city life. Furthermore, it has been shown that differences between listening tests and everyday listening situations could not be imputed only to the artificial laboratory conditions and the required processes of abstraction, but also to the "quality" of the recreated listening situation. Indeed, descriptions similar to the field survey were collected in the second listening test, which used an immersive multichannel reproduction method (Ambisonics) with no visual reference to the speakers. Hence, it is necessary to provide both a neutral visual environement and a good sense of spatial immersion in the recreated soundfield to ensure ecological validity of an experimental protocol to test psychological effects of urban *background noise*. In a similar vein, Guastavino and Katz [15] conducted perceptual evaluations of spatial audio reproduction systems using a wider range of sound material (musical passages, indoor and outdoor environments). A strong correlation between auditory scenes and reproduction methods was observed, highlighting the fact that there is no single system that is optimal for arbitrary source material. Together, these findings underline the dependence of different sound quality criteria upon source material.

6. Conclusion

The present results show that the same acoustic phenomenon can give rise to two different cognitive representations which integrate properties of mental representations (here, spatial immersion) into physical descriptions of the stimuli. More generally, the reproduction system of listening tests must be adapted to the purpose of the study, to allow the subjects to treat the test samples as potentially familiar experiences through cognitive processes elaborated in actual situations. Specific properties of mental representation must be taken into consideration when designing experimental settings to ensure ecological validity. It has been showed that these properties can be inferred by analyzing spontaneous verbal descriptions collected in a reference study, using linguistic discourse analysis techniques, in close relation to theoretical elaborations of the relations between language and cognition. The verbal data are then classified into semantic categories emerging from the spontaneous descriptions (here, reference to the source, object-centered and subject-centered descriptions). Occurrences within each category can be compared to those obtained in the reference study with statistical tools. This justifies the linguistic analysis of verbal data as a reliable measure of the ecological validity of reproduction systems in experimental settings.

Acknowledgments

Experimental work performed while at the Laboratoire d'Acoustique Musicale, Université Paris 6-CNRS, 11 rue de Lourmel, F75015 Paris, France. The authors would like to thank Pascale Cheminée for her valuable help in linguistics. This work was funded by the French Ministries of Environment (equipment) and Research (funding for C. Guastavino's doctoral work conducted at the Université Paris 6), and a Chateaubriand Fellowship (B. Katz), Office for Science and Technology of the Embassy of France in the United States. The writing of this paper was supported by a grant from VRQ to D. Levitin, and a post-doctoral fellowship grant to C. Guastavino from CIRMMT (Centre for Interdisciplinary Research in Music Media and Technology).

References

[1] E. Benveniste: Problème de linguistique générale (general linguistics). Vol. 1. Gallimard, Paris, 1966, 124–130.

[2] C. Normand: La quadrature du sens (the quadrature of meaning). Presses Universitaires de France, Paris, 1990.

[3] S. Hawakaya: Language in thought and action. Harcourt, NY, 1949.

[4] F. de Saussure: Cours de linguistique générale (general linguistics). 1955 ed. Payot, 1916.

[5] D. Dubois: Categories as acts of meaning: The case of categories in olfaction and audition. Cognitive Science Quaterly **1** (2000) 35–68.

[6] B. Berlin, P. Kay: Basic color terms. New York, Academic Press, 1969.

[7] S. David: Représentations sensorielles et marques de la personne: contrastes entre olfaction et audition (sensory representations and reference to the speaker: contrasts between olfaction and audition). – In: Catégorisation et cognition. D. Dubois (ed.). Kimé, Paris, 1997, 211–242.

[8] F. Guyot, M. Castellengo, B. Fabre: Etude de la catégorisation d'un corpus de bruits domestiques (categoristion of domestic environmental sounds). – In: Catégorisation et cognition: de la perception au discours. D. Dubois (ed.). Kimé, Paris, 1997, 41–58.

[9] D. Dubois, C. Guastavino, V. Maffiolo, M. Raimbault: A cognitive approach to soundscape research. Journal of the Acoustical Society of America **115** (2004) 2592. Invited paper presented at the 147th Meeting of the Acoustical Society of America. New York: USA.

[10] D. Dubois, C. Guastavino, V. Maffiolo, M. Raimbault: Acoustic phenomena between noise(s) and sound(s). Strasbourg, France, März 2004, CFA/DAGA: 7th Congrès Français d'Acoustique, 30th Congress of the German Acoustical Society.

[11] C. Guastavino, P. Cheminée: Conceptualisations en langue, représentations cognitives et validité cologique: Une approche psycholinguistique de la perception des basses fréquences (language conceptualizations, cognitive representations and ecological validity: a psycholinguistic approach to low frequency perception). Psychologie Française **48** (2003) 91–101.

[12] V. Maffiolo: De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement sonore urbain (semantic and acoustic approaches to sound quality of urban environments). Dissertation. Université du Maine, Le Mans, 1999.

[13] M. Raimbault: Simulation des ambiances sonores urbaines: intégration des aspects qualitatifs (urban soundscapes simulation: integrating qualitative aspects). Dissertation. Université de Nantes, 2002.

[14] J. Gibson: The ecological approach to visual perception. 1986 ed. Houghton Mifflin, Boston, MA, 1979.

[15] C. Guastavino, B. Katz: Perceptual evaluation of multidimensional spatial audio reproduction. Journal of the Acoustical Society of America **116** (2004) 1105–1115.

[16] V. Maffiolo, C. Vogel, J.-D. Polack, M. Castellengo, D. Dubois: Ambiances sonores représentatives d'une ville: le cas de Paris (typical urban soundscapes: the case of Paris, France). Actes du 4^{e} Congrès Français d'Acoustique, Marseille, 1998, Teknea, 303–306.

[17] D. Dubois, C. Guastavino, V. Maffiolo: The meaning of city noises; investigating sound quality in Paris, France. Journal of the Acoustical Society of America **115** (2004) 2495. Invited paper presented at the 147th Meeting of the Acoustical Society of America. New York: USA.

[18] J.-D. Polack, M. Castellengo, V. Maffiolo, C. Vogel, G. Guastavino, B. Katz: Soundfield reproduction: the limits of the physical approach. Strasbourg, France, März 2004, CFA/DAGA: 7th Congrès Français d'Acoustique, 30th Congress of the German Acoustical Society.

[19] C. Guastavino, P. Cheminée: A psycholinguistic approach to the ecological validity of experimental settings: The case of low frequency perception. Food Quality and Preference **15** (2004) 884–886.

[20] C. Guastavino: Étude sémantique et acoustique de la perception des basses fréquences dans l'environement sonore urbain (semantic and acoustic approaches to low frequency perception). Dissertation. Université Paris 6, 2003.

[21] P. Fellgett: Ambisonics reproduction of directionality in surround sound systems. Nature **252** (1974) 534–538.

[22] M. Gerzon: Design of ambisonic decoders for multi-speaker surround sound. New York, 1977, 58th AES Convention.

[23] J. Daniel: Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia (acoustic field representation, application to the transmission and the reproduction of complex sound environments in a multimedia context). Dissertation. Université Paris VI, CNET, Sep. 2000.

[24] N. Vanderveer: Ecological acoustics: Human perception of environmental sounds. Vol. 40. Dissertation Abstract International, 1979. (University Microfilms number 8004001).

[25] W. Warren, R. Verbrugge: Auditory perception of breaking and boucing events: A case study in ecological acoustics. Journal of Experimental Psychology: Human Perception and Performance **10** (1984) 704–710.

[26] J. Ballas: Common factors in the identification of an assortment of brief everyday sounds. Journal of Experimental Psychology, Human Perception and Performance **19** (1993) 250–267.

[27] M. Mzali, D. Dubois, J.-D. Polack, F. Létourneaux, F. Poisson: Mental representation of auditory comfort inside trains: methodological and theoretical issues. The Hague, The Netherlands, 2001, Internoise.