Buenos Aires – 5 to 9 September, 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

Soundscape, Psychoacoustics and Urban Environment: Paper ICA2016-267

Influence of experimental conditions on sound pleasantness evaluations

Pierre Aumond[(a)(b),*], Florent Masson[(c)], Leandro Beron[(c)], Arnaud Can[(a)], Bert De Coensel[(d)], Dick Botteldooren[(d)], Carlos Ribeiro[(e)], Catherine Lavandier[(b)]

[(a)] Ifsttar, Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux, Nantes, France, pierre.aumond@gmail.com / arnaud.can@ifsttar.fr
[(b)] Laboratoire Mobilité, Réseaux, Territoires et Environnement, Université de Cergy Pontoise, Cergy-Pontoise, France, catherine.lavandier@u-cergy.fr
[(c)] Universidad Nacional de Tres de Febrero, Buenos Aires, Argentina, fmasson@untref.edu.ar
[(d)] Waves Research Group, Department of Information Technology, Ghent University, iGent Technologiepark-Zwijnaarde 15, 9052 Ghent, Belgium, bert.decoensel@intec.ugent.be / dick.botteldooren@intec.ugent.be
[(e)] Bruitparif, Paris, France / Carlos.Ribeiro@bruitparif.fr

Abstract

Being able to characterize and estimate the urban sound perception is a key point to improve the city dwellers environmental quality. In the past decade, various studies have focused on collecting perceived global sound pleasantness at specific locations. Some of them were carried out on field in order to evaluate the soundscape perception of the participants directly in their context. Other studies were realized in laboratory to better control the stimuli and to increase the number of participants who were subjected to the same sound environment. Most of the laboratory experiments are done in large or semi-anechoic chamber with calibrated and highly realistic audio reproduction in order to respect the ecological validity of the experiment. On one hand, even with a high immersive level, the laboratory context is not as rich as the field context and the two types of experiment could lead to different results. On the other hand, few studies exist showing the influence of decreasing ecological validity for the same experience. This work presents a short statistical analysis of perceptive evaluations of ten urban locations under 4 different test conditions. First, evaluations are carried out directly in-situ in the city of Paris. Then audio-visual recordings of these locations are evaluated in three different experimental conditions: (i) in a well-controlled acoustic laboratory in Paris region with French people, (ii) in an acoustic laboratory in Buenos Aires with Argentinean participants and lowest immersive conditions, (iii) in a habitational room with Argentinean participants and subjective calibration. The study reveals that both the "country" factor and the experimental conditions in laboratory do not show any significant impact on the perceived sound pleasantness and perceived loudness assessments.

Keywords: Sound quality; Ecological validity; Experimental conditions

Influence of experimental conditions on sound pleasantness evaluations

1 Introduction

The increasing urbanization results in the deterioration of urban sound environments. Noise is regularly cited by city dwellers as the principal cause of annoyance. To fight against this nuisance, the European Directive 2002/49/CE imposes to cities of more than 100 000 inhabitants to develop and diffuse strategic noise maps [1]. The role of the produced maps is to help decision makers in the development of noise mitigation plans and to inform city dwellers about their exposure.

However, since they are classically based on a modeling chain that rests upon an identified corpus of sound sources, namely the road traffic, railway traffic, aircraft noise and the biggest industries, these maps hardly represent the complexity of urban sound environments. Indeed, the specificity of the noise pollution lies in its variety of sources, its high spatiotemporal variations, its rich spectral component, and the complexity of human hearing. As a consequence, a key point when evaluating the quality of urban sound environments is the selection of relevant descriptors, which should be easy to understand and handle, be sensitive to urban sound dynamics and correlated to perceptive dimensions [2]. The usual energy noise indicators fail to describe the complexity of urban noise dynamics, both physically and perceptively [2], [3].

Recent researches modified the approaches for characterizing urban sound environments, seeking descriptors or classification nearer of the city dwellers perception [4]. In this context, various experimental methodologies have been proposed in order to collect perception data from participants [5]. In Situ experiments have the benefit of achieving the evaluation of the soundscape perception of the participants directly in their context [2], [6]. In parallel, laboratory test enable to better control the stimuli and to increase the number of participants exposed to the same sound environment [7].

Before generalizing the conclusions on soundscape perception given by these contrasted methodologies, it is important to test if they offer similar results. The influence of some external factors on soundscape perception, such as the presence of visual information, has already been shown [7]. Further research is required to investigate the influence of other factors on sound pleasantness evaluation, such as the technology of the sound restitution system for example. Some studies put a lot of effort on increasing the participant's immersion or the ecological validity of their experiment, by means of immersion rooms such as fake living rooms or virtual reality system [7]–[9]. The same care is often put into the quality of the sound recordings or reproduction [10] or into classifying the participant on socio-cultural factors (age, country, etc.) [2], [11]. Thus, this is crucial to evaluate the influence of these variables on the perceived sound environment.

To provide new insights regarding the impact of these external factors, in this study a specific survey was proposed to different group of participants under different experimental conditions. The controlled variables investigated are the laboratory rooms, the sound reproduction system, the objective calibration, the country of residence. All the obtained results in laboratory are compared with in situ evaluations. The second section details the data collecting for the four experimental conditions. The section 3 shows a comparison of the results obtained by these different methodologies. The section 4 draws some conclusions concerning the generalization of perceptive tests.

2 Data Collection

The analyses presented are derived from a group of experiments developed within the GRAFIC Project (Fine and continued cartography of the sound environment), which aims to develop noise mapping methodologies close to the city dwellers perception. For this reason, the experiment was not fully designed for the purpose of this study only. For example, some locations are presented but are not analyzed.

2.1 In Situ Experiment

The in situ experiment consists of a perceptive test performed 4 times (referred further in the text as “runs” over a 2,1 km-long path, located in Paris (13rd district) as described in Figure 1. Perceptive data were collected during 3-to-5 minute stops over 19 points located along the path, resulting in a test duration of approximately 45 minutes. The points were chosen to contain a large variety of urban sound environments on both sides of soundscape transitions, resulting in an average distance of 115 meters between points.

Figure 1: Noise Map of the studied path, interpolated from mobile measurements (LAeq - dB).

Participants were divided into four runs, with about 10 participants per run on average. This group size can be considered small enough to not modify the surrounding sound environment keeping a sufficient number to perform statistical analysis. In total, 37 participants participated to the test (67% men, 33% women; 61% 10-29 y.o, 21% 30-39 y.o, 6% 40-49 y.o, 9% 50-59 y.o, 3% 60-75 y.o). The four perceptive tests were performed on Mondays 23/03/2015 and 30/03/2015, between 10-12 h and 14-16 h. The path, travelled each day, alternated from West to East (WE) and from East to West (EW). The participants received a small monetary compensation, and participated in only one of the four runs.

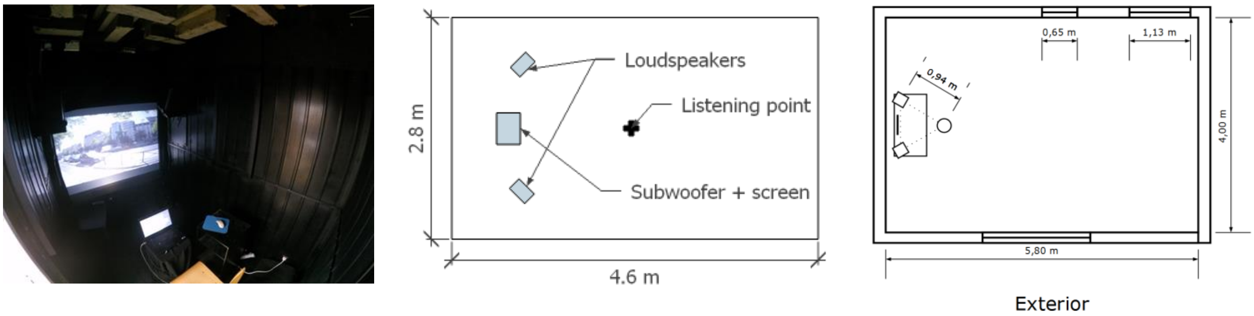
The perceptive test consisted of 16 questions on an 11-point bipolar semantic scale, administered at each of the 19 points. The questions covered two categories of perceptive parameters, most already investigated in previous studies [2], [12]: the Pleasantness, which describes the pleasantness of the sound environment for the stimulus or at the stopping point, from “unpleasant” to “pleasant”; the Loudness, which describes the perceived loudness of the same sound environment and the time ratio of presence for birds, voices and footsteps, which describes the perceived time of presence for these sound sources from “rarely present” to “continuously present”. Nine of the 19 points, representative of the corpus and presenting a high variety of sound environments, were picked to be evaluated in laboratory (see Table 1).

Table 1: Experiment stopping points soundscapes and approximate vehicle flow rates.

Point	Morphology	Flow Rate [Veh/h]	Perceived Sound pleasantness (median note for all situation)	Perceived Loudness (median note for all situation)
P5	Very quiet street with rare traffic pass-byes,	~10	9	3
P6	Animated street (Bars & Restaurants)	~50	7	8
P3	Quiet park	-	9	7
P11	Quiet street without traffic	-	7	6
P20	Street with high vehicle flow rates	~1000	2	10
P13	Street with very low traffic flow rates	~50	5	6
P16	Street with high vehicle flow rates	~2000	3	9
P17	Quiet pedestrian street located between two busy street	-	6	5
P19	Large Park	-	8.5	8.5
P2	Large two-lanes traffic street	~2 x 750	3	8

2.2 Laboratory experiments

For ten points, (Points 5, 6, 3, 11, 20, 13, 16, 17, 19, 2 and 20; see Figure 1), a video sequence of 1 minute was recorded with a ©GoPro camera some days before the in situ experiment. During the recording, a sweeping 270° angle was slowly done in order to integrate the visual context of the evaluated location. Similar sound environment conditions were chosen (recordings on Mondays, between 10 and 12 h, or between 14 and 16 h). All the recordings taken close to roads include at least both red and green traffic light phases. The audio files were recorded simultaneously to measurements, thanks to high quality binaural microphones. Prior to each recording, a calibration tone (1kHz/94dB) was recorded by the microphones for the need of the laboratory experiments.



(a) (b) (c)

Figure 2: Photography or sketch for the laboratory experiments (a) Experiment 1 (b) Experiment 2 and (c) Experiment 3.

Each laboratory evaluation was proceeded individually: one participant was seated in a chair in front of a computer showing the test instructions. Sound events were reproduced through loudspeakers, as observed in Figure 2. The participants were also naive with regard to the hypotheses under test, and received a small monetary compensation for participation. Participants were involved in only one experiment;

2.2.1 Experiment 1: UCP Laboratory

The listening test of the first experiment took place in a semi anechoic room in the University of Cergy-Pontoise (UCP), France; see Figure 2.a. The video sequence was projected on a very large screen located behind the computer. The sound sequences were reproduced through the transaural technique [13], using a 2.0 system composed by two loudspeakers (Tannoy) and a high quality sound card (RME Fireface 400) The listening point was located at ± 30º from loudspeakers. The transaural listening has the advantage to minimize the front/back confusions, which are known to appear with headphones listening, especially when individual HRTFs and head-tracking is not available, while preserving the perceptual characteristics of a diffused sound field. Also, the fact that participants are not using headphones improves the realism of the simulation technique. Thanks to the recorded calibration signal, the sound levels in the laboratory were set in order to have similar sound pressure levels than during the recordings. A group of 30 participants participated in the experiment, 18 women and 12 men with a mean age of 33 years old (SD = 14). They were all French citizens, and most urban inhabitants (>90%).

2.2.2 Experiment 2: UNTREF Laboratory

The listening test of the second experiment took place in a quiet room in the Universidad Nacional de Tres de Febrero (UNTREF), Argentina; see Figure 2.b. This laboratory has controlled conditions on background noise and reflections, but it is not a fully anechoic enclosure. The total volume of the room is of 25.5 m^3. The room reaches a NC 25 curve and has a mid-frequency reverberation time of 0.2 s. The video sequence was reproduced on a 21-inch screen. The sound sequences were reproduced through a stereo technique, using a 2.0 system composed by two Dynaudio BM6A loudspeakers and a Digidesign Mbox 2 sound card. A KRK subwoofer Rokit 10 was used to enhance low frequencies. The listening point was located at a 30º difference between loudspeakers. As the test room is not anechoic a calibration of the reproduction system has been done to ensure a relatively flat frequency response in the listening point. Again, thanks to the recorded calibration signal, the sound levels in the laboratory were set in order to have similar sound pressure levels than during the recordings. A group of 17 participants participated in the experiment, 4 women and 13 men with a mean age of 26.4 years old (SD = 3). They were all Argentinians citizens from Buenos Aires.

2.2.3 Experiment 3: Habitational room (ROOM)

The listening test of the third experiment took place in a large habitational room in Buenos Aires, Argentina; see Figure 2.c. The video sequences were reproduced on a 19-inch computer screen. The speaker setup complied with a traditional stereo technique, using a 2.0 system composed of two Tannoy 501a loudspeakers placed at equal distance from each other and from the listening point. Audio signal was reproduced through a Focusrite Pro 14 sound card. For this experiment, the sound levels were set without the calibration tone but following a subjective procedure. A first musical stimulus was presented to the participant. The participant had to adjust a sound level control (of 1dBFS steps) to a preferred listening level. A second stimulus was then presented. Using one of the stimuli selected for its relatively low sound intensity (point P11) the participant was asked to adjust again the sound level control to achieve the most immersive sensation on the soundscape. Thus, the sound control level remained at the same position while the ten-video sequence was presented to the participant in a random order as for experiments 1 and 2. This methodology turned out that firstly presenting a popular song assisted the participants to get familiar with the testing setup response before introducing any soundscape recording. Further investigations must be performed in order to determine if this is relevant for the sound levels adjustment.

A group of 10 participants participated in the experiment, all men with a mean age of 26.2 years old (SD = 2.4). They were all Argentinians citizens with at least 3 years of residency in Buenos Aires.

3 Results and discussion

Figure 3 shows the perceived sound pleasantness and loudness (median value and inter-quartiles) for each point, and for the different experimental conditions. For practical reason, the point 20 doesn't have In Situ assessment.

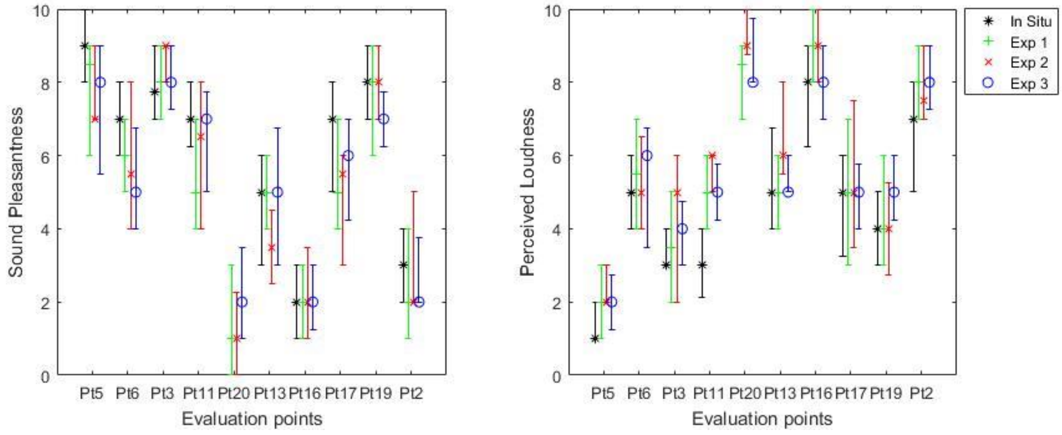


Figure 3: perceived sound pleasantness and loudness for the 10 selected points and the four experimental conditions (median and inter-quartiles values).

Interestingly, equivalent results are obtained for different points and experimental conditions. Slightly higher pleasantness values are observed for the in situ experiment, but the correlation coefficient between the results obtained by the four experimental conditions is rather high (r > 0.9) and significant (p < 0.01), which attests of a very coherent behavior (see table 2).

Table 2: Correlation coefficients between experiments for pleasantness and loudness.

Pleasantness	UCP	UNTREF	ROOM	Loudness	UCP	UNTREF	ROOM
In Situ	0,95**	0,91**	0,95**	In Situ	0,95**	0,88**	0,95**
UCP		0,94**	0,94**	UCP		0,96**	0,98**
UNTREF			0,95**	UNTREF			0,92**

** p < 0.01.

3.1 Laboratory conditions influence

A non-parametric test (Kruskall-Wallis) was proposed, in order to evidence the laboratory influence on subjective evaluations. No effect was observed for both the perceived pleasantness ($X^2(2,574) = 0.11$, p = 0.94) and the perceived loudness ($X^2(2,571) = 5.11$, p = 0.08). Thus, the hypothesis that the soundscape perception depends on the country of residence was not verified on this set of experiments. An objective sound level calibration in the listening room shows no significant influence. The others factors that differed within the laboratory experimental conditions, such as the sound reproduction system or the screen size, were not detected as influent on the evaluations.

Lots of the participants witnessed some discomfort due to the contrast between the very low noise levels within the room (25 to 30 dBA) and the high noise levels of the calibrated sequences (45 to 85 dBA) in UCP and UNTREF experiments. This discomfort however didn't have significant influence for the loudness evaluation.

3.2 Laboratory vs in situ results

As all the laboratory tests provided similar results, the laboratory notes were grouped and then compared to the in situ evaluations. A non-parametric test (Kruskall-Wallis) was used, which highlighted significant differences for both the pleasantness ($X^2(1{,}784) = 6.45$, $p < 0.05$) and the loudness ($X^2(1{,}781) = 8.2$, $p < 0.01$) evaluations.

As said before, the measurements used for the three experiments were not made at the same moment than for in situ evaluations. Thus, even if the same moment was kept constant (on Mondays, 10-12 h or 14-16 h), the sound events were not the same.

Nevertheless, a pairwise comparison was done using the same statistical test which highlights significant differences for only two points, the other ones being similarly evaluated over in situ and in laboratory conditions:

- The point "Pt6" presents a significant difference for pleasantness evaluation ($X^2(1{,}91) = 9.26$, $p < 0.01$). It is worthy to note that the test also shows significant differences on voice and footpath variables, respectively ($X^2(1{,}91) = 24.1$, $p < 0.01$) and ($X^2(1{,}91) = 9.6$, $p < 0.01$). Theses discrepancies can be explained because of the audio-visual stimuli, which were not sufficiently representative of the human presence in the environment. Indeed, this point is located in a street with a high presence of bars and restaurants. The participants of the in situ test may have been more influenced at the moment of the evaluation by the contextual factors than the participants of the laboratory tests. As a consequence, as the sound pleasantness is influenced by the human presence perception [2], its evaluation presents also significant discrepancy.
- The point "Pt11" presents significant difference for both pleasantness and loudness evaluations, respectively ($X^2(1{,}89) = 7.56$, $p < 0.01$) and ($X^2(1{,}89) = 16.6$, $p < 0.01$). Interestingly, significant differences are also present on presence of traffic and bird time ratio variables, respectively ($X^2(1{,}89) = 13.3$, $p < 0.01$) and ($X^2(1{,}89) = 45.5$, $p < 0.01$). Here again, the audio-visual stimulus was not enough representative of the traffic time presence. Indeed, this point was located in a very small pedestrian path nearby a street with traffic presence. Again, context factors may have influenced the evaluation, and in situ participants may have noted a very small traffic presence in the sound environment in comparison with the neighbor street.

The difference between in situ and laboratory experiments can also be explained when sound sources are differently present in the sound environment. An excerpt of 1 minute could be too short to represent the global context of an urban sound situation, with presence of all sound sources. Brocolini [12] proposes to analyze 10 minute recordings at least to characterize the sound environments. Such long stimuli cannot be used in a laboratory context because of the limited duration of perceptive tests. A duration of 3 minutes is sometimes preferred in laboratory context to study the sound perception of urban environments [14]. Even if for practical reason, it was not possible in this study to use such long stimuli, based on the precedent remarks; it could be definitively a better approach.

4 Conclusions

A short survey has been proposed in four very different experimental conditions. The results obtained by the different groups of participants provide these conclusions:

- The laboratory experimental conditions didn't have significant impact on the perceptive evaluations.
- The country of residence didn't have significant impact on the perceptive evaluations.
- An objective calibration of the stimuli sound level didn't have significant influence on the perceptive evaluations, compared with a subjective calibration.
- When presence of typical sources is important, the recording should respect the time ratios of sources otherwise the perceptive evaluations obtained over laboratory experiments may not be fully in agreement with the evaluations done in situ.
- If the length of the tested stimuli (1 min) was most of the time sufficient to reproduce the sound environment at a given location, longer stimuli (3 min) are recommended to avoid the over-representation of eventful sound sources, and integrate more representative sources in the sound ambiance.

It is worthy to note that the entire conclusions given in this paper are restrained to a survey including closed question, and focusing only on two variables, namely the perceived sound pleasantness and the perceived loudness. Further research is needed to validate the non-significant impact for each of the factors linked with the experimental conditions. Nevertheless, most of the perceptive experiment methodology is a compromise between resources, and a number of hypotheses to evaluate. If fully controlled laboratory or in situ experiments are needed to validate models or theoretical assumptions, this study tends to show that poor experimental conditions can also be a very useful tool to test original hypothesis. Indeed, a survey without the need of high ecological validity, or a high quality reproduction system, could be diffused to an increased and very large amount of participants.

Acknowledgments

This work has been carried out in the framework of the GRAFIC project, supported by the French Environment and Energy Management Agency (ADEME) under contract No. 1317C0028. The authors also acknowledge the students Ignacio Calderón, Felipe Lechuga and Lautaro Florentin who participated on the organisation of the UNTREF experiment.

References

[1] EC, "Directive 2002/49/EC of the European parliament and the Council of 25 June 2002 relating to the assessment and management of environmental noise," *Off. J. Eur. Communities*, vol. 189, no. 12, pp. 12–25, 2002.

[2] P. Ricciardi, P. Delaitre, C. Lavandier, F. Torchia, and P. Aumond, "Sound quality indicators for urban places in Paris cross-validated by Milan data," *J. Acoust. Soc. Am.*, vol. 138, no. 4, pp. 2337–2348, Oct. 2015.

[3] A. Can, L. Leclercq, J. Lelong, and J. Defrance, "Capturing urban traffic noise dynamics through relevant descriptors," *Appl. Acoust.*, vol. 69, no. 12, pp. 1270–1280, Dec. 2008.

[4] W. J. Davies, M. D. Adams, N. S. Bruce, R. Cain, A. Carlyle, P. Cusack, D. A. Hall, K. I. Hume, A. Irwin, P. Jennings, M. Marselle, C. J. Plack, and J. Poxon, "Perception of soundscapes: An interdisciplinary approach," *Appl. Acoust.*, vol. 74, no. 2, pp. 224–231, Feb. 2013.

[5] S. R. Payne, W. J. Davies, and M. D. Adams, "Research into the Practical and Policy Applications of Soundscape Concepts and Techniques in Urban Areas," Jan. 2009.

[6] A. Maristany, M. R. López, and C. A. Rivera, "Soundscape quality analysis by fuzzy logic: A field study in Cordoba, Argentina," *Appl. Acoust.*, vol. 111, pp. 106–115, 2016.

[7] S. Viollon, C. Lavandier, and C. Drake, "Influence of visual setting on sound ratings in an urban environment," *Appl. Acoust.*, vol. 63, no. 5, pp. 493–511, May 2002.

[8] M. Lindquist, E. Lange, and J. Kang, "From 3D landscape visualization to environmental simulation: The contribution of sound to the perception of virtual environments," *Landsc. Urban Plan.*, vol. 148, pp. 216–231, 2016.

[9] L.-A. Gille, C. Marquis-Favre, and A. Klein, "Noise annoyance due to urban road traffic with powered-two-wheelers: Quiet periods, order and number of vehicles," *Acta Acust. united with Acust.*, vol. 102, no. 3, pp. 474 – 487, Jan. 2016.

[10] Z. Schärer and A. Lindau, "Evaluation of Equalization Methods for Binaural Signals," *AES 126th Conv.*, p. 17, 2009.

[11] C. Lavandier and P. Delaitre, "Individual and shared representations on 'zones calmes' ('quiet areas') among the French population in urban context," *Appl. Acoust.*, vol. 99, pp. 135–144, Dec. 2015.

[12] L. Brocolini, C. Lavandier, M. Quoy, and C. Ribeiro, "Measurements of acoustic environments for urban soundscapes: choice of homogeneous periods, optimization of durations, and selection of indicators.," *J. Acoust. Soc. Am.*, vol. 134, no. 1, pp. 813–821, 2013.

[13] J. L. Bauck and D. H. Cooper, "Generalized transaural stereo and applications," *J. Audio Eng. Soc.*, vol. 44, no. 9, pp. 683–705, 1996.

[14] A. Trollé, J. Terroir, C. Lavandier, C. Marquis-Favre, and M. Lavandier, "Impact of urban road traffic on sound unpleasantness: A comparison of traffic scenarios at crossroads," *Appl. Acoust.*, vol. 94, pp. 46–52, Jul. 2015.