MADRID
inter.noise 2019
June 16 - 19
NOISE CONTROL FOR A BETTER ENVIRONMENT

Community Noise Quality Assessment (Annoyance) by means of a Virtual Audio Environment

Misdariis, Nicolas[1]
STMS Ircam-CNRS-SU
1, place Igor Stravinsky – F- 75004 Paris, France

Dianoux, Camille[2]
STMS Ircam-CNRS-SU
1, place Igor Stravinsky – F- 75004 Paris, France

Marignier, Lucie[3]
STMS Ircam-CNRS-SU
1, place Igor Stravinsky – F- 75004 Paris, France

Leiba, Raphaël[4]
Sorbonne Université, CNRS, Institut Jean Le Rond d'Alembert
4, place Jussieu – F- 75005 Paris, France
STMS Ircam-CNRS-SU
1, place Igor Stravinsky – F- 75004 Paris, France

ABSTRACT

This study deals with community noise quality assessment (annoyance) of urban environments. It is based on a previous work that resulted in: a) a road-traffic vehicle audio signals extraction tool, allowing to estimate different sonic properties; b) a modeling tool that estimates the perceived annoyance level by adapting a current multiclass psychoacoustic model, given in the literature. The present work follows these outputs and aims at validating experimentally the annoyance estimations. A listening test is designed in order to be able to collect perceived annoyance in a virtual environment. Sound scenes made of urban soundscapes (background) and vehicles pass-by (foreground) are built and encoded in a 2D to multichannel algorithm, in order to create a listening experience as ecological as possible. Then, a consistent experimental protocol is designed in order to measure the perceived annoyance caused by each synthesized sound scene. The focus is put on the spatial audio environment that is used for the perceptual experiment. Especially, urban sound scene synthesis methods and tools are presented. Results from the perceived naturalness and immersion collected during the test are presented and analyzed. The global issue addressed by the use of virtual reality in laboratory experiments is finally discussed.

Keywords: Noise, Environment, Annoyance
I-INCE Classification of Subject Number: 63 ; 79

---

[1] nicolas.misdariis@ircam.fr
[2] camille.dianoux@ircam.fr
[3] lucie.marignier@gmail.com
[4] raphael.leiba@sorbonne-universite.fr

1. INTRODUCTION

The present work is undertaken in the large frame of acoustic ecology, and more precisely concerns urban soundscapes and community noise assessments. Within a global research program that dealt with life quality in urban soundscape (Mouvie, 2014-17), we initially postulate that – either for both air quality or noise exposure –, one of the main sources of urban pollution is related to mobility, that is to say largely consists in road vehicles that daily populate more or less all the cities worldwide.

Since the seminal works of R.M. Schafer and colleagues within the World Soundscape Project [1], a lot of scientific studies and applied works have been done to investigate, qualify, understand, and sometimes improve urban sonic environments, and indirectly their inhabitants' daily life. In that scope, since the last decades and among much others, we can cite the 2002/49/CE European directive that aimed at establishing strategic soundmaps – expressed in sound level L_{den} –, that force the politics to address these issues. As an alternative to this '*normative*' answer to noise pollution, a more '*sensitive*' way of considering the problem [2] came some years ago from the European COST Action "Soundscape of European Cities and Landscapes" (TD0804, 2013-17)) which claimed to consider environmental sounds as "a 'resource' rather than a 'waste'" [3] arguing that "reducing sound level does not necessarily lead to improved quality of life in urban / rural areas" (see [4], for more details). More recently, a collaborative research project, EUREQUA (2012-17), aimed at taking into account these two different approaches by objectifying and assessing the environmental quality at neighborhood scale through the identification of relevant criteria related to: *i)* the characterization of the physical environment (acoustics, but also air quality or climate); *ii)* the evaluation of the quality of life in the neighborhood by residents and users [5]. By means of an interdisciplinary and participatory dimensions, this kind of projects tends to make the (missing) link between what is measured and what is perceived in terms of community noise exposure.

In the same point of view, a PhD work has been recently defended (2017), that dealt with the development of a tool able to diagnose the perceived noise annoyance by inhabitants in a urban context [6]. This work globally relies on 3 main phases: *i)* the measurement of urban sound scenes, to determine, isolate and characterize single sound sources – in that case – of passing-by road vehicles; *ii)* the classification of sound events forming these scenes into perceptually coherent categories and represent them in a more synthetic manner; *iii)* the modeling of the noise annoyance generated by these sound scenes, to estimate the perceptual sensations felt by their neighborhood.

The first phase mainly relies on a multi-channel acquisition device (*Megamicros*) used with a beamforming processing. It allows to parse and extract isolated sound sources from a complex sound scene. *Megamicros* fits in acoustic imaging systems but is developed upon an original digital MEMS (Micro Electronic and Mechanical System) microphones array that can manage devices from a few tens to more than a thousand of sensors, scattered over a large area (typically several meters). Moreover, a video tracking algorithm allows to localize moving sources over an extended sound scene in order to extract their contributions to the overall emission with beamforming in frequency domain. In the time of the PhD work, this technical apparatus has been validated with outdoor implementations both at the vehicle scale (controlled single sources) and the street scale (flow of uncontrolled multiple sources). In particular, the validation at the vehicle scale was operated during a measurement campaign in a professional automotive testing environment (La Ferté-Vidame, exploited by PSA Peugeot-Citroën) and produced a database of controlled passing-by for different vehicles representative to the urban traffic (see [7] for more details on this phase).

The second phase consists in classifying the extracted audio signal over perceptual categories thanks to machine learning. This work is done upon the results of Morel *et al.*'s perceptual study of vehicle sound sources [8]. They outlined a perceptual representation of the soundscape related to mobility (vehicles) explained by two factors: vehicle types (light, heavy, 2-wheels) and driving conditions (acceleration, deceleration, constant speed). The work led to 7 categories: Cat.1 (2-wheels, constant), Cat.2 (2-wheels, accel.), Cat.3 (light+heavy, constant), Cat.4 (2-wheels, decel.), Cat.5 (light+heavy, decel.), Cat.6 (light, accel.), Cat.7 (heavy, accel.). From that, a Support-Vector Machine (SVM) using MFCC (Mel-Frequency Cepstrum Coeff.) description of the source signals – but also the driving conditions (extracted from trajectory tracking from the video analysis) – allows to build an automatic classifier able to assign an input recorded data in one of the seven Morel *et al.*'s categories, with an overall estimation error around 15% (see [9] for more details on this phase).

The third phase is the core of the present paper. It is done upon the results of Morel and colleagues perceptual modeling of vehicle sound sources [10]. In fact, for each of the perceptual category of vehicles (see above), they defined a model based on acoustic and psychoacoustic features, computed on the input signal, and able to deliver the estimation of a specific annoyance related to the given category. During his PhD, Leiba implemented and refined this model, especially with regards to computation of some psychoacoustic features (roughness, and fluctuation strength). The last step of this third phase then consists in closing the loop of the process with a quantified evaluation of the model, *i.e.* a compared analysis of computed *vs.* measured annoyance data.

We implement a laboratory protocol based on a virtual audio environment, inspired by literature and a previous work done for sound design purpose [11]. This environment is used to render audio realism, in order to generate valuable perceptual judgment on semi-virtual sound scenes, additionally controlled in terms of composition (type of category) and sound level.

In this paper, we will, first, briefly present the perceptual noise annoyance model that has been implemented (Sec. 2). This model will be challenged in Sec. 3. In details, the experimental approach will be developed (Sec. 3.1) and finally some comparisons between modeled and measured annoyance data will be presented (Sec. 3.2). A special focus will be given on the constitution of the virtual audio environment by addressing these three main questions : From what and how the stimuli have been designed ? With which hardware / software architecture the sound scene realism has been rendered in laboratory conditions ? How the realism has been felt and assessed by the participants ?

2. PERCEPTUAL NOISE ANNOYANCE MODEL

The present perceptual noise annoyance model belongs to a category that tries to explain annoyance by multi-linear combinations of acoustic and psychoacoustic features computed from the signal – in that case, a vehicle noise. This approach does definitely not include a non negligible part of perceived annoyance that involves other factors than physical or physiological – for instance social, demographic or cultural (see Berglund e*t al.* for a general development on that point [12] or de Coensel *et al.* for an example of such a multi-factorial model [13l]). Nevertheless, the present work assumes a rather low level initial approach that was scientifically built and validated few years ago.

2.1 Foundations

The model starts from Morel *et al.* works that established relationships between several acoustic / psychoacoustic features and specific annoyance for each perceptual vehicle categories [10] – that they also stated in a previous work [8] (see Introduction).

Resulting from this work, and in total, the features used to model annoyance in one – or several – categories are the followings:

- N : overall loudness (time averaged),
- $N_{15\text{-}18}$: specific loudness for Barks 15 to 18 (time averaged),
- R_{max} : roughness (maximum value)
- F, F_{max} : fluctuation strength (average and maximum value)
- $\triangle N^{-}$: loudness decreasing rate (after the passing of the vehicle)
- L_{MF} : A-weighted level in the medium frequency range (315 – 1250 Hz, 1/3 octave)

The precise analytical expressions (extracted from Morel *et al.*'s outcomes [10]) describe the specific annoyances (A_n), in the form of linear regression models that have been, in addition, perceptually validated (R^2 scores between 0.91 and 0.97, $p < 0.05$):

$A_1 = 1.03.N + 0.18$ $\quad$ $A_2 = 16.99.N_{15\text{–}18} + 0.10.F + 1.45$

$A_3 = 1.32.N - 0.32.\triangle N^{-} - 0.36$ $\quad$ $A_4 = 0.89.N + 0.02.R_{max} + 0.33$

$A_5 = 1.07.N + 0.08.F_{max} - 1$ $\quad$ $A_6 = 0.29.L_{MF} - 8.5$

$A_7 = 0.95.N + 0.10.F - 0.5$

2.2 Implementation

In more details, and regarding the most sensible features, it appears that: *i)* loudness N is used in all categories except Cat.6 (light vehicles in acceleration), and the loudness decreasing rate $\triangle N^{-}$ is used in Cat.3 (light/heavy vehicles at constant speed); *ii)* roughness R is used in Cat.4 (2-wheel vehicle in deceleration); *iii)* fluctuation strength F is used in Cat.2-5-7 (respectively, 2-wheel in acceleration, light/heavy vehicle in deceleration and heavy vehicle in acceleration).

For our purpose, loudness is computed with Fastl and Zwicker's instationnary loudness model [14], implemented by Genesis (*Loudness Toolbox*). Roughness is adapted from Daniel and Weber's model [15], inspired by Garcia's implementation [16], and improved in order to better fit the theoretical curves established by Fastl and Zwicker [14]. The optimization is done in a parametric way, by adjusting 5 structural parameters of the roughness model and minimizing a cost function that represents the distance between the digitalized theoretical curves and the computed data from the combinatory model. A similar protocol is used to adapt and optimize the fluctuation strength model, implemented by Osses Vechhi *et al.* [17], on the basis of the theoretical expression established by Fastl and Zwicker [14].

These modified seven models have finally been implemented and tested with real data (sound signals) acquired during two different measurement campaign on either a predefined traffic with controlled outdoor conditions (La Ferté-Vidame) and an undefined traffic with uncontrolled outdoor conditions (St-Bernard Quay, Paris). In brief, the results are congruent with a more classical measurement ($L_{A,eq}$) but offer an interesting reduction of data variability (see [6] for more details).

3. EXPERIMENTAL EVALUATION OF THE PERCEPTUAL MODEL

The experimental validation of the noise annoyance perceptual model presented in Sec. 2 forms the core of the present work. It aims at validating the whole research process by comparing numerically computed data (from the model) with perceptually measured data (from the experiment). In this section, we will especially pay attention to the environment implemented for conducting the experiment in an ecological way, *i.e.* the way to place the participants in realistic conditions for noise annoyance assessments. This being, the *Results* sub-section (Sec3.2), will first focus on the assessment of the context (realism, immersion) before delivering main results on the percept (noise annoyance) – that is more developed in a complementary paper [18].

3.1 Evaluation environment

In the light of experimental realism and immersion, the evaluation environment gets mainly three components answering three main questions: what do participants hear (stimuli) ? How are they acoustically contextualized (sound rendering) ? How are they mentally contextualized (experimental procedure) ?

3.1.1 Stimuli

On the basis of the generic description of a sound sequence by Nelken and de Cheveigne (2013) [19]: "a skeleton of events on a bed of texture", the stimuli of this experiment result from the sound design of sequences based on two main components, edited and mixed in a DAW – Digital Audio Workstation (Logic Pro):

- a background noise – the "bed of texture" – coming from a database of field recording sounds in outdoor spaces (Paris streets and avenues) initially realized with two different devices: an ORTF couple (Schoeps MSTC 64U) and an order-1 Ambisonic microphone (Soundfield ST250). This campaign was part of a previous work that aimed at developing a urban environment audio simulation for contextual evaluation of quiet vehicles' sound design [11]. Those recordings were realized by two professional sound engineers and were made at different places in order to have a sample of sound atmospheres from calm zones (narrow lanes, quiet zones, etc.) to very noisy areas (large avenues, with a high traffic flow, etc.).
- a set of specific vehicle sounds – the "skeleton of events" – coming from a database of sound recordings (and processings) obtained within the frame of Leiba's PhD [6] with the help of automotive infrastructures provided by a french car manufacturer (PSA Peugeot Citroen). This campaign consisted in recording separately a large set of different vehicles while controlling either types (light/heavy/2-wheel vehicles), engine specifications (number of cylinders and/or engine size), dynamics (acceleration, deceleration, constant speed) and recording context (a quiet, isolated place without noise pollution). The sound recordings were realized with three different devices: a reference monophonic microphone (B&K 4190), a binaural device (HEAD acoustics Artificial Head) and a 2D-microphonic array (256-channel MEMS microphones). The initial aim of this set up was to record calibrated samples for each perceptual vehicle category, as defined by Morel *et al.* [8].

On this basis, for the design step of the sound sequences – and after several rounds of expert listening –, we finally chose the stereo ORTF recordings for the soundscape source and the monophonic recordings for vehicle sources, as being the best compromise between sound quality and immersive capacities. The soundscape was edited in order to create a neutral – or amorphous as defined by Maffiolo (1999) [20], and used by Dubois *et al.* (2006) [21] – sequence, without any noticeable event to be heard. Vehicle sounds are then selected for each of the 7 perceptual categories and if needed edited to make a given category clearly recognizable – for instance, by manually reshaping the increasing or decreasing level profile of respectively acceleration or deceleration – and also to give them spatial properties – basically, left/right trajectories.

The output of this first technical step resulted in: *i)* a background sequence able to be looped without detecting any audio artifact; *ii)* 7 sets of vehicle samples (among Morel *et al.*'s categories) able to be randomly played at any time on the foreground of the sequence. This output will be used hereafter for building, in realtime during the experiment, the mixed urban sequence to be assessed according to annoyance.

3.1.2 Sound rendering apparatus

On the basis of seminal works done – among others – by Guastavino and Katz (2004) according to perceptual evaluation of multi-channel spatial audio reproduction architectures [22], we choose to use a 5.1 broadcast system is used as a compromise between ecological and immersive aspects, especially with regards to the binaural technology. Indeed, despite its great immersive quality, the binaural implementation may prevent from moving freely without being embarrassed by headphones and may also require a sensitive calibration procedure.

The set-up is installed in a mastering professional studio (Ircam-Studio 8 – surface area = 47 m^2, TR = 200 ms.) by means of a 5.1 PSI monitoring system (model A25-M + A225-M subwoofer) conventionally distributed in the room (0°, +30°, -30°, +110°, -110°), within a 2,4-meter diameter circle (Figure 1). Moreover, note that no specific equalization of the system is done for the current experiment.

Figure 1. the 5.1 set-up installed in the mastering studio 8 at Ircam (© LCA/Culturebox)

The digital software used with this hardware set-up is the Spatialisateur (Spat$^{©}$), a virtual acoustic sound rendering engine internally developed since few decades [23], and especially its most recent released module, Panoramix [24]. In our case study, this processing environment allows to encode the two groups of stimuli (background and events) in the 5.1 format, targeting both a diffused spatialization for the background and controlled localization trajectories for the events (see Figure 2 for a block diagram of the implemented audio process)

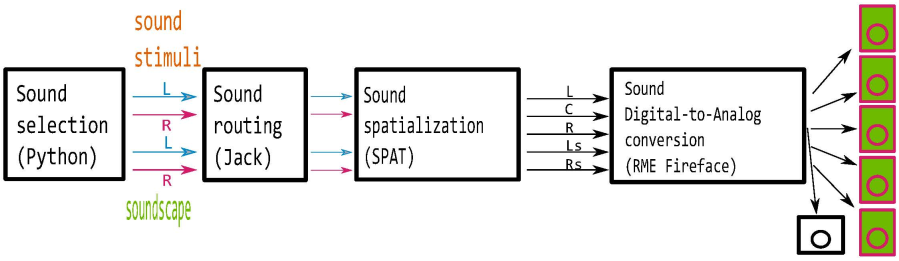

Figure 2. Block diagram of the audio architecture implemented for the experiment: from the sources selection composing the sound sequence (left), to the 5.1 diffusion system installed in the studio (right)

The background component is set by adjusting the aperture of the stereo image, encoded in the 5.1 format by means of a VBAP algorithm [25] – perceptually adjusted by the co-authors expert team in order to maximize the immersive impression. With the same perceptual/empirical approach, the event components are virtually placed in front of the audio set-up at a realistic distance with regards to an outdoor daily experience.

3.1.3 Experimental procedure

The experimental procedure is mainly inspired by two previous studies dealing with noise annoyance measurement of industrial and environmental sounds.

First, Morel *et al.* (2016) conducted a laboratory study where participants were asked to assess their perceived annoyance while listening to pass-by sound sequences (representing perceptual categories that the same team determined previously [8]). The context and the task were simply presented to the participants by the following sentence: “Imagine yourself at home while relaxing (*e.g.* reading, watching television, discussing, gardening or doing other relaxing activities you are used to). While you are relaxing, you hear this road traffic noise. How much would this noise annoy you?’’. Answers could be given in a continuous numerical scale with regularly spaced semantic labels from “not at all" to "extremely" [10].

Second, Brocolini *et al.* (2016) conducted a laboratory listening test to assess the sound annoyance in open-plan offices. The sequences were composed of 5 different sound conditions – 4 different sources and 1 control condition (printer sounds, intelligible / unintelligible human voices, phone rings, air-conditioner background noise). For each sequence, a task consisting in recalling words in a list put the participants in a cognitive context similar to a reading activity and provided a performance measure. After each sound condition (several sound sequences of the same type), a NASA-RTLX questionnaire is made to estimate participants’ mental workload. Then, at the very end of the experiment, an annoyance measurement was realized by answering the following question : “At which point did the sound sources affect you in accomplishing the task requested?”, by means of a continuous scale for each sound sources – that could also be recalled by playing back a 5-second excerpt of it [24].

On that bibliographic basis, we have built an experiment that reuses some of these components and introduces new ones. In fact, the experimental design considered the 7 vehicles categories already defined (Sec. 3.1.1) – plus a control condition where the background noise is played alone – presented at 4 different sound levels (from 50 to 62 dB(A) by 4-dB step), so that the sound corpus is made of 32 sound sequences. Each sequence contains 3 events of a given category randomly distributed in a 25-second sample of background noise set at 45 dB(A) – or the background noise alone, in the control condition (repeated 4 times, at the same level of 45 dB(A)). The presentation of this combination (categories *vs.* sound levels) is randomized for each participant.

Then, we divide the experiment in two distinct protocols with regards to dependent (measured) variables : a *direct* one and an *indirect* one. The *direct* protocol only consists in asking, after each of the 32 sequences, the perceived annoyance during the listening, assessed in a continuous numerical scale (0 to 10). The *indirect* protocol involves a cognitive task that consists in recalling words from a list visually showed (on a computer screen) during the playback of the sound sequence. The lists are made of words taken from Dubois and Poitou’s lexical lists [25]. Sixteen lists of twenty commonly cited words in general domains (like animals, trees, fruits, sports, professions, etc.) are selected. The recall performance is done by a verbal return. This protocol also includes a participant’s workload estimation by means of a NASA-RTLX questionnaire [26] that is asked at the end of each sound sequence together with the annoyance estimation (made in the same way than for the *direct* protocol). To sum up, the three dependent variables considered are: the (subjective) perceived annoyance, the (subjective) workload and the (objective) recall performance.

Moreover, general question are also asked at the very end of the experiment (after the 32 sequences). They mainly concern the fatigue felt after the test and, above all, the assessment of the experimental realism by means of a mark in a continuous 0-10 scale (0 = not realistic at all; 10 = very realistic), together with free comments.

A total of 36 participants (18 for *direct*, 18 for *indirect* protocol) took part to this experiment. They were aged from 19 to 48 years old (mean age = 26) quite equally distributed in terms of gender (19 women, 17 men). Moreover, they all declared not having neither serious hearing disorders (tinnitus, implants, etc.) nor neurological antecedents (epilepsy, etc.). For each of them, the experiment lasted approximately 1 hour including a break (*indirect* protocol), or 45 minutes for the *direct* protocol in which the 32 sound sequences are played twice in two consistent blocks separated by a break.

3.2 Results

Data collected are still in progress in terms of analysis. For this paper, and in this section, we focus first on experimental audio environment assessments (Sec. 3.2.1), and give some main results –further developed in [18] – about the global purpose of this work, *i.e.* concerning the perceptual annoyance model (Sec. 3.2.2).

3.2.1 Realism of the experimental audio environment

The realism of the audio set-up can be estimated with the judgment score and additional comments given by each participant at the very end of their test. They had to answer questions orally asked by the experimenter about the "realism of the sound environment", and if needed were asked to give additional comments about the "feeling of immersion during the test" and the "benefits and drawbacks of the experimental apparatus". The judgment on realism is done on a continuous numeric scale from 0 (not realistic) to 10 (very realistic) and the comments result from a free verbalization.

The numerical data (36 scores between 0 to 10, divided in two groups : *direct* or *indirect* protocol – see Sec. 3.1.3) show a good assessment of the apparatus (average = 8.4) and a low variability of these judgments (standard deviation = 0.9). Moreover, this global result tends to be consistent within the two groups of participants, according to the nature of the protocol (*direct* vs. *indirect*) – see Table 1 for details.

	Mean	**Stand. Deviation**
All	8.3889	0.9420
direct group	8.4722	0.9467
indirect group	8.3056	0.9570

Table 1. Analysis of numerical data on realism judgment (0: not realistic – 10: very realistic) collected from the 36 participants having passed either the 'direct' or 'indirect' experimental protocol (18 persons in each)

In addition, it is worth noticing that most of the free comments were rather positive about the realism and show a number of occurrences of the terms 'realistic', 'very realistic', 'realism ok' or similar around 20 (above 36) with sentences like: "very impressed by the sound rendering' (Subj. #31), "very realistic, especially the background noise" (Subj. #30), or "very realistic indeed, I thought that some sounds came from the outside of the studio" (Subj. #6).

Nevertheless some criticisms were formulated, especially with regards to the unrealistic low level of human voices ("people usually talk louder in the real life", Subj. #24), the level of some events from the heavy vehicle categories (played at the loudest value of 62 dB(A)) that seemed too loud for some participants, and a relative consensus on the fact that the deceleration scenario was felt to be the less realistic of whole set of sound scenes: "decelerations are not very realistic, the vehicles stop too frequently and too quickly" (Subj. #35), "deceleration not realistic, they stop suddenly" (Subj. #36), "too much vehicles that stop in one go, in the same time" (Subj. #33).

3.2.2 Global results on perceptual annoyance

This section looks into the perceptual data collected according to the experimental perceptual noise annoyance assessed in *direct* and *indirect* protocols. First of all, Figure 3 shows that, at a 1st order of analysis, these two sets of data can be considered similar in average. Indeed, the variance of these two sets of data is explained by a significant Pearson's correlation coefficient of 0.90 and their distribution fits well with the corresponding linear regression model.

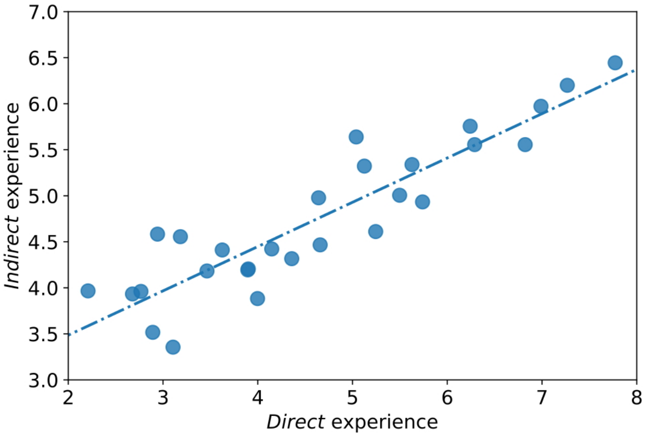

Figure 3. Linear regression between averaged data of annoyance assessed in the direct protocol (mean of block 1 and 2) and indirect protocol. Each point corresponds to a vehicle category (1 to 7) and a given sound level (50 to 62 dB). Pearson's coefficient gives a significant correlation of 0,90, (p <.001)

Secondly, a look into relationships between the main measured variable (annoyance) and one of the experimental factor (sound level), in Figure 4, leads to recover a common result stating that annoyance increase positively with noise exposure level – on which the current soundmap approaches are incidentally based. Currently, this result seems confirmed for all perceptual vehicle categories (nearly parallel curves) but also shows a rather high variability among values (offsets and large span of annoyance values, actually, between Cat.3 (light+heavy vehic. at constant speed) and Cat.7 (heavy vehic. in acceleration). In a way, this latter observation legitimates the perceptual approach which tends to refine this 1st level of modeling.

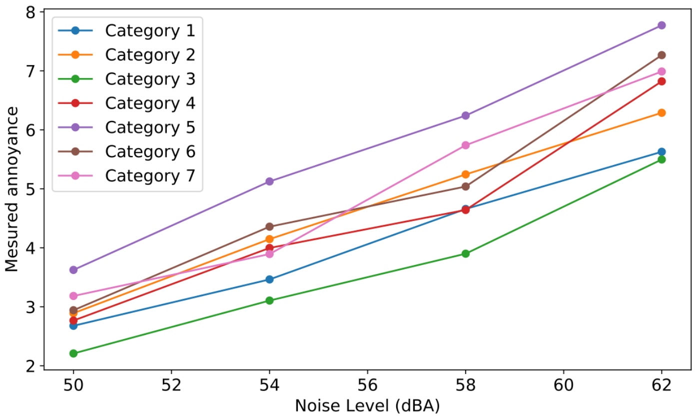

Figure 4. Influence of events noise level on perceived annoyance evaluation for each step of this factor (50 to 62 dB(A) by 4-dB step). Each line corresponds to a vehicle category (see Introduction).

Thirdly, the fitting between estimated (modeled) and measured (perceived) data of annoyance is investigated and depicted in Figure 5. It allows us to get an interesting result that confirms Morel *et al.*'s outcomes whereas it has been proposed in another scientific context and different experimental conditions. In fact, this result shows that the perceived annoyance of different vehicle categories and the corresponding modeled annoyance are well correlated and that the model explains a significative part of the variance between these two sets of data (R=0.89, R^2=0.79, p<.001). Figure 5 shows mainly three things: *i)* the points are well distributed along the regression line so that it validates the linear relation between measurement and modeling; *ii)* this being, some categories fit better than others; for instance Cat.2 tends to have a slightly different slope (2-wheel vehic. in acceleration – modeled by narrow band loudness, N_{15-18}, and fluctuation strength, F) or Cat.6 tends to be rather non linear (light vehic. in acceleration – modeled by equivalent sound level in medium frequencies, L_{MF}); *iii)* the estimated annoyance seems to be always twice as much the measured value. These issues are deeper investigated in [18].

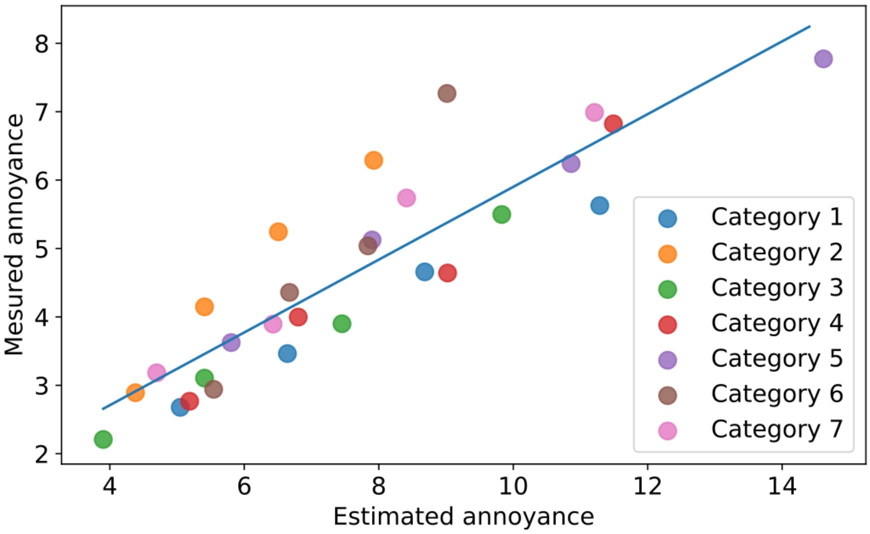

Figure 5. Linear regression between estimated annoyance and experimentally measured annoyance (direct protocol). Each point corresponds to combination of vehicle category (1 to 7) emitted at a specific noise level (50 to 62 dB(A)). Pearson's coefficient gives a significant correlation of 0.89, (p < 0,001)

4. CONCLUSIONS

This paper presents the whole synopsis of a large scale research project that investigates the characterization of urban soundscapes, with regards to road traffic sources [6]. It focuses mainly on the last part of this project that deals with experimental validation of a perceptual noise annoyance model, based on previous works [8, 10].

We detail the implementation of a laboratory set-up that relies on a virtual audio environment able to render as ecologically as possible urban sound scenes with different vehicle sources – considered as responsible for the perceived noise annoyance – in the foreground,. This percept is measured with an experimental protocol that tends to compare different cognitive listening situations (*direct* or *indirect* paradigm).

The results obtained show a global positive assessment of the realism of the audio scenarios, together with a rather good feeling of immersion inside the sound scenes. Moreover, the annoyance theoretical model seems to significantly correlate with the data experimentally obtained, and to be a more relevant answer to the noise annoyance estimation than the classical equivalent sound level ($L_{A,eq}$ or L_{den}).

On that basis, the perspectives of this work is two-fold *i)* to improve and consolidate the results already obtained by refining the analysis – this part is mainly done in a current complementary development [18] ; *ii)* to improve the model itself, either with regards to some psychoacoustic features that build the model, or in terms of global annoyance, leading to an important – an yet unresolved – issue that focuses on the summation process occurring in front of a complex mixture of sound events.

5. ACKNOWLEDGEMENTS

The authors are thankful to Prof. Bert De Coensel and Prof. Massimiliano Masullo for their invitation to participate in the Technical Session on " Virtual Reality for Noise Control and Sound Quality " of the Internoise 2019 conference.

This work is in line with the research made within the Chair "Mobilty and Quality of Life in Urban Surroundings" (2014-2017) led by The UPMC Foundation and sponsored by the Donors (PSA Peugeot-Citroën and Renault).

6. REFERENCES

1. Schafer, R. M. (1977). The tuning of the world . Alfred A. Knopf, 1st .dition.

2. Chetelat, J. (2009). La figuration cartographique de l'espace sonore. Images Re-vues. Histoire, anthropologie et théorie de l'art , (7).

3. Schulte-Fortkamp, B. (2013, June). Soundscape-focusing on resources. In Proceedings of Meetings on Acoustics ICA2013 (Vol. 19, No. 1, p. 040117). ASA.

4. Misdariis, N., Marchiano, R., Susini, P., Ollivier, F., Leiba, R., & Marchal, J. (2014, October). Mobility and life quality relationships–Measurement and perception of noise in urban context. In *Internoise-International Congress on Noise Control Engineering*

5. Gauvreau, B., Guillaume, G., Can, A., Gaudio, N., Lebras, J., Lemonsu, A., ... & Haoues-Jouve, S. (2016, September). Environmental quality at district scale: A transdisciplinary approach within the EUREQUA project. In *Ficup 2016, First International Conference on Urban Physics* (p. 13p).

6. Leiba, R. (2017). *Conception d'un outil de diagnostic de la gêne sonore en milieu urbain* (Doctoral dissertation, Université Pierre & Marie Curie-Paris 6).

7. Leiba, R., Marchiano, R., Misdariis, N., & Marchal, J. (2016, August). Urban acoustic imaging: from measurement to the soundscape perception evaluation. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 253, No. 2, pp. 6803-6809). Institute of Noise Control Engineering.

8. Morel, J., Marquis-Favre, C., Dubois, D., & Pierrette, M. (2012). Road traffic in urban areas: A perceptual and cognitive typology of pass-by noises. *Acta acustica united with acustica*, *98*(1), 166-178.

9. Leiba, R., Ollivier, F., Marchal, J., Misdariis, N., & Marchiano, R. (2017, December). Large array of microphones for the automatic recognition of acoustic sources in urban environment. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 255, No. 5, pp. 2662-2670). Institute of Noise Control Engineering.

10. Morel, J., Marquis-Favre, C., & Gille, L. A. (2016). Noise annoyance assessment of various urban road vehicle pass-by noises in isolation and combined with industrial noise: A laboratory study. *Applied Acoustics*, *101*, 47-57.

11. Misdariis, N., Gerber, J., & Aleonard, J. (2014, October). Urban environment audio simulation for contextural evaluation of qient vehicles' sound design. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 249, No. 2, pp. 4957-4966). Institute of Noise Control Engineering.

12. Berglund, B., & Nilsson, M. E. (1998, October). Total annoyance models for community noises explicated. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 1998, No. 4, pp. 1013-1016). INCE.

13. De Coensel, B., & Botteldooren, D. (2010). A model of saliency-based auditory attention to environmental sound. In *20th International Congress on Acoustics (ICA)*.

14. Fastl, H., & Zwicker, E. (2006). *Psychoacoustics: facts and models* (Vol. 22). Springer Science & Business Media.

15. Daniel, P., & Weber, R. (1997). Psychoacoustical roughness: Implementation of an optimized model. *Acta Acustica united with Acustica*, *83*(1), 113-123.

16. García R. (2015). "Modelling the sensation of fluctuation strength," Master thesis, Eindhoven University of Technology

17. Osses Vecchi, A., Garcia Leon, R., & Kohlrausch, A. (2016). Modelling the sensation of fluctuation strength. In *Proceedings of Meetings on Acoustics 22ICA* (Vol. 28, No. 1, p. 050005). ASA.

18. Misdariis, N., Marignier, L., Dianoux, C., Leiba, R. (2019) Experimental validation of a urban traffic noise annoyance model. Proceedings of the 26th International Congress on Sound and Vibration *(submission in progress)*

19. Nelken, I., & de Cheveigne, A. (2013). An ear for statistics. Nature neuroscience , 16 (4), 381.

20. Maffiolo, V. (1999). De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain . (Doctoral dissertation, Universit. du Maine, France).

21. Dubois, D., Guastavino, C., & Raimbault, M. (2006). A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. *Acta acustica united with acustica*, *92*(6), 865-874.

22. Guastavino C., Katz B. F. Perceptual evaluation of multi-dimensional spatial audio reproduction. The Journal of the Acoustical Society of America, 116(2), 2004.

23. Jot, J. M., & Warusfel, O. (1995). A real-time spatial sound processor for music and virtual reality applications. In ICMC: Intern. Computer Music Conference.

24. Carpentier, T. (2016). Panoramix: 3D mixing and post-production workstation. In *42nd International Computer Music Conference (ICMC)*.

25. Pulkki, V., & Karjalainen, M. (2008). Multichannel audio rendering using amplitude panning [DSP applications]. *IEEE Signal Processing Magazine*, *25*(3), 118-122.

24. Brocolini, L., Parizet, E., Chevret, P. (2016). Effect of masking noise on cognitive performance and annoyance in open plan offices. *Applied Acoustics*, *114*, 44-55.

25. Dubois, D., & Poitou, J. (2002, Mars). « Normes catégorielles » et listes lexicales. *Cahiers du LCPE*, *5*, 31-104.

26. Hart, S. G. (2006, October). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 9, pp. 904-908). Sage CA: Los Angeles, CA: Sage Publications.