EXPERIMENTAL VALIDATION OF A URBAN TRAFFIC NOISE ANNOYANCE MODEL

Nicolas Misdariis, Lucie Marignier and Camille Dianoux  
STMS Ircam-CNRS-SU, Paris, France  
email: nicolas.misdariis@ircam.fr, lucie.marignier@gmail.com, camille.dianoux@ircam.fr

Raphaël Leiba  
Sorbonne Université, CNRS, Institut Jean Le Rond d’Alembert, Paris, France  
e-mail: raphael.leiba@sorbonne-universite.fr

This study deals with noise annoyance in urban environments. It is based on a previous work that resulted in two main outputs. First, a measurement tool that identifies and extracts sonic properties of the main urban noises sources, i.e. the different types of vehicles. Second, a modeling tool that estimates the perceived annoyance level by adapting from the literature, and implementing, a current multi-class psychoacoustic model. The present work follows these outcomes and aims at validating the annoyance estimations given by the model with an experimental approach. A listening test is designed in order to be able to collect perceived annoyance in a virtual environment (in lab test). Sound scenes made of urban soundscapes (background) and vehicles passing-by (foreground) are built and encoded in a 2D-audio immersive format (2D to multichannel algorithm), in order to create a listening experience as ecological as possible. Within this frame, an important experimental protocol is designed in order to measure the perceived annoyance caused by each synthesized sound scene and to control the cognitive load (attention on a list of words to play back afterwards) of the participants. The first rounds of data analysis show consistent results, especially in terms of influence of sound level or source typicality on the measured annoyance. A deeper round of analysis is currently being processed in order to examine the fine correlation between measured and predicted annoyance values, with regards to psychoacoustic features of source signals. All the results will be presented and discussed in the course of the conference.

Keywords: urban soundscape, annoyance, measurement, perceptual model, assessment

1. Introduction

The present study comes within the broad scope of environmental acoustics and community noise research and management, especially referenced by the seminal work of R.M. Schafer and colleagues’ World Soundscape Project [1]. It is guided by the 2002/49/CE European directive [2] that aims at establishing strategic soundmaps and encourages – or forces – the politics to address issues concerning noise exposure and pollution. It stands in a large-scale research program (Mouvie, 2014-17) that deals with relationships between mobility and life quality in urban environments; and makes the assumption that road traffic is the 1st-order source of pollution in cities, either for air quality or noise exposure – we will focus here on the latter (acoustic) part of this working hypothesis [3].
This being, and quoting Chételat [4], the urban noise pollution issue could be addressed either in a “normative” way – based on objective measurements, sound level thresholds and operational solutions (noises barriers, soundproofing, etc) –, or in a “sensitive” way – based on subjective metrics, features related to psychoacoustical, cognitive or even social components enabling a more detailed description of the noise exposure and a better understanding of its underlying impacts. Actually, from the last decades, the need for developing new methodologies to qualify subjective approaches of environmental soundscapes has been formalized [5] and numerous attempts have been made for modeling such environments with perceptual parameters. Among others, Kang (2006) lists several factors of different natures (physical, psychological, social, etc.) that are involved in the soundscape assessment [6], and Brocolini et al. (2012) proposes new sound features to characterize the perceived quality of urban soundscapes [7, 8]. More recently, and still among others, the European project HARMONICA (2011-2015) proposes a new perceptual metric, the Harmonica Index, based on a combination of different energetic features, and the french collaborative research project EUREQUA (2012-2017) aims at identifying relevant criteria related to characterization of the physical environment (acoustics, but also air quality or climate) and evaluation of the quality of life in the neighborhood by residents and users [10].

Within the Mouvie program, a research has been conducted to design a tool able to diagnose noise annoyance perceived by citizens in a given context [11]. Basically, this work relies on 3 main phases:

1. measurement of urban sound scenes to characterize single sound sources of passing-by road vehicles. This is done with a multi-channel device (Megamicros), coupled with a video tracking system. This allows to localize moving sources and extract their acoustic contributions [12].
2. classification of sound events into 7 perceptually coherent categories according to: a) type of vehicles (light, heavy, 2-wheels), b) driving conditions (acceleration, deceleration, constant speed) [13]. This results in single source automatic labeling (overall estimation error 15%) [14].
3. modeling of noise annoyance, on the basis of Morel et al.’s perceptual model using acoustic / psychoacoustic features for each class-specific annoyance [15]. This model has been refined during the current work, especially with regards to some psychoacoustic features.

The last step of this 3rd phase concerns the evaluation of the model and forms the core of the present paper. It basically consists in comparing computed (model) and measured (experiment) data, in order to confirm the accuracy of the modeling, and if needed, to adjust the model to the new data set. For that, we implement a laboratory protocol based on a virtual audio environment, and conducted a perceptual experiment that allows lay participants to express their feeling about annoyance of semi-virtual sound scenes, controlled in terms of context (urban noise), composition (vehicle category) and sound level.

After briefly presenting the perceptual noise annoyance model inherited from the literature and its improvements (Sec. 2), the paper largely details the experimental approach (Sec. 3.1) and develops the main significant results obtained according to the evaluation challenge (Sec. 3.2), trying to resolve this fundamental issue: how can we predict annoyance felt by people exposed to a complex urban sound scene, on the basis of the fine knowledge on the content of the scene ? The conclusion then gives the main perspectives for future works in that direction.

2. Perceptual noise annoyance model

The current perceptual noise annoyance model is based on a multi-linear regression of acoustic and psychoacoustic features computed from the signal, i.e. a vehicle sound signature. At this point, it is admitted that this 1st-order approach doesn’t take into account non negligible components of perceived annoyance like, for instance, social, demographic or cultural inputs (see Berglund et al. for general notions [16] or de Coensel et al. for a specific instance of such a cognitive multi-factorial model [17]).
2.1 Initial state

The model comes from Morel et al. work that established relationships between acoustic features and annoyance for the perceptual vehicle categories [13]. For each of the categories (Cn), a specific annoyance (Aa), is computed by a multi-linear regression model that significantly explains (R² scores between .91 and .97, p<0.05) the variation of perceptual data obtained experimentally [15] (Table 1).

Table 1: Morel et al.’s perceptual road traffic categories (Cn) [13] and specific annoyance models (Aa) [15]

<table>
<thead>
<tr>
<th>Category</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1: (2-wheel vehicles, constant speed)</td>
<td>A1 = 1.03*N + 0.18</td>
</tr>
<tr>
<td>C2: (2-wheel vehicles, acceleration)</td>
<td>A2 = 16.99<em>N_{15-18} + 0.10</em>F + 1.45</td>
</tr>
<tr>
<td>C3: (light+heavy vehicles, constant speed)</td>
<td>A3 = 1.32<em>N – 0.32</em>AN^– – 0.36</td>
</tr>
<tr>
<td>C4: (2-wheel vehicles, deceleration)</td>
<td>A4 = 0.89<em>N + 0.02</em>R_{max} + 0.33</td>
</tr>
<tr>
<td>C5: (light+heavy vehicles, deceleration)</td>
<td>A5 = 1.07<em>N + 0.08</em>F_{max} – 1</td>
</tr>
<tr>
<td>C6: (light vehicles, acceleration)</td>
<td>A6 = 0.29*L_{MF} – 8.5</td>
</tr>
<tr>
<td>C7: (heavy vehicles, acceleration)</td>
<td>A7 = 0.95<em>N + 0.10</em>F – 0.5</td>
</tr>
</tbody>
</table>

where N: overall loudness (mean); N_{15-18}: specific loudness for Barks 15 to 18 (mean); R_{max}: roughness (max); F, F_{max}: fluctuation strength (mean, max); ΔN^–: loudness decreasing rate (after vehicle passing); L_{MF}: A-weighted level in the medium frequency range (315 – 1250 Hz, 1/3 octave).

2.2 Current refinement

The initial model brings out the recurrent contribution of important psychoacoustic features. Loudness N is used in all categories except C6 (light vehic. accel.), and loudness decreasing rate ΔN^– is also used in C3 (light/heavy vehicles const. speed). Roughness R is used in C4 (2-wheel vehic. decel.). Fluctuation strength F is used in C2, C5 and C7 (respectively, 2-wheel vehic. accel., light/heavy vehic. decel. and heavy vehic. accel.). Then, the current improvement of the model mainly concerns these three features that are usually difficult to derive for complex sounds and even somewhat ill-defined.

In their study, Morel et al. used dBSonic software (from 01dB Metravib) to compute all the audio features. But this tool is no longer available. It led us to base our work on open-access implementations of the different index models. In this study, loudness is computed with Fastl and Zwicker’s stationnary loudness model [18], implemented by Genesis (Loudness Toolbox). For roughness index, an optimized version of García’s implementation [20] of Daniel and Weber’s model [19] is used. The optimization helps in better fitting the theoretical curves established by Fastl and Zwicker [18]. The fluctuation strength model used is implemented from Osses Vechhi et al. [21]. Also for this index, an optimization of the implementation is performed to better assess this auditory percept.

3. Experimental evaluation of the perceptual noise annoyance model

The assessment of the noise annoyance perceptual model presented in Sec. 2 forms the core of the current work. It aims at giving a final evaluation of the 3-phase research process described in the Introduction. The present section first describes the experimental method used to implement this goal (Sec. 3.1), before delivering a series of results regarding successively the consistency of participants’ answer, an indirect way to estimate annoyance and, finally, the accuracy of the model (Sec. 3.2). Note that the methodological point related to ecological validity and realism of the experimental audio rendering setup has been left aside, since it is largely developed in a complementary paper [22].

3.1 Method

3.1.1 Stimuli

By analogy with soundscape structural composition [1], the experimental stimuli are designed as 2-component sequences mixing an urban background noise (texture) and vehicle sounds (events).
Background noise comes from a database of field recordings in Parisian outdoor spaces realized by two professional sound engineers for a previous work [23]. They cover a large scope of atmospheres, from calm zones (narrow lanes, quiet zones, etc.) to very noisy areas (large avenues, with a high traffic flow, etc.). The recording is made, among other devices, with an ORTF couple (Schoeps MSTC 64U).

Sound events come from a sound database realized in the Mouvie framework [11, 12]. They were made during a campaign that recorded separately different vehicles with controlled conditions: vehicle types (light/heavy/2-wheel), engine types (number of cylinders, engine size), dynamics (acceleration, deceleration, constant speed) and general context (an automotive professional test track). The recording is made, among other devices, with a monophonic microphone (B&K 4190).

On this basis, sound materials for the experiment are designed in 2 steps. Recordings of urban backgrounds are edited to create an amorphous sequence [24, 25]. Monophonic recordings of vehicle sounds are selected to illustrate the 7 perceptual categories (Table 1).

3.1.2 Apparatus

On the basis of seminal studies dealing with realism of multi-channel audio rendering devices, we choose to use a 5.1 broadcast system as a compromise between ecological and immersive properties [26]. The set-up is installed in a mastering professional studio by means of a 5.1 PSI monitoring system (see [22] for details). Participants were seated on a chair, respecting the 5.1 disposition of the studio and they faced a computer on which the instructions of the experiment were displayed.

The software Panoramix [27] allows to encode the two groups of stimuli (background and events) in the 5.1 format, ensuring diffused spatialization of the background and precise localization of the events.

3.1.3 Participants

Thirty-six people (17 women) participated in this experiment (Age: 19-48 years old, Mean: 26). All declared having normal hearing abilities, no neurological impairment and no attention deficit disorder. They were paid for their participation. They were separated into two groups: group (a) only assessed noise annoyance and group (b) assessed annoyance “indirectly” (i.e., they performed a cognitive task while being exposed to the stimuli and rated their workload in addition to the perceived annoyance). For both groups, the experiment lasted for about 50 minutes.

3.1.4 Procedure

The experimental procedure is mainly inspired by two previous studies dealing with noise annoyance measurement of industrial or environmental sounds [15, 28]. An experimental plan with repeated measurements has been designed with 3 independent variables (vehicle category, sound level and participant’s group) and 3 dependent variables (annoyance, memory performance and workload).

Twenty-height sound sequences lasted for about 25 seconds were designed for both groups, with 3 events of the same vehicle category and at the same sound level – among 4 different levels (from 50 to 62 dB(A) by 4-dB step). Events appeared in a random order within each sequence, as each sequence. In addition, 4 sequences with only background noise were inserted at specific moments of the experiment (#6, 13, 19 and 26) and served as control conditions. In both groups, the experiment began with a training test to ensure the correct understanding of the task.

Transposing the method adopted by Morel et al. (2016), participants of group (a) were firstly asked to imagine themselves doing an activity of their choice requiring concentration (e.g. reading, writing an e-mail, etc.) at a coffee terrace. Then, for each sound sequence, participants answered the following question: “At which point did the sound sources annoy you in accomplishing the task you imagined doing?” [15]. They answered on a continuous scale (from 0 to 10), marked by 5 labels from “not at all” to “extremely” annoyed. Group (a) assessed each sound sequence twice: a 5-minute break was taken between block 1 and 2 that contained each the whole set of 32 sequences. Participants could take as much time as they wanted to answer each question.
Group (b) assessed the annoyance of each sequence while performing a working memory task [28]. For each sequence, participants were instructed to memorize the 10 words that were displayed on the computer. As in Brocolini et al. (2016), we chose words from the lexical lists established by Dubois and Poitou (2002) [29]. We selected 16 lists (such as Sports, Trees, or Fruits), of 20 frequently used words. Lists and words in lists were randomized for each participant. Each list appeared twice (with 2 different 10-word sets), so that each participant had finally seen all words of all lists.

The experiment was conducted as follows: first, participants did a Stroop task that is used to measure their capacity to inhibit interference effect (here, the noise environment) while performing a task [30]. Afterwards, for each sequence, a fixation cross appeared in the middle of the computer screen during 1.5s (to guide the participant’s gaze), then the lexical list’s name was displayed during 2s, and again a fixation cross appeared during 1.5s. After that, the 10 lexical list’s words were displayed successively for 2s each. Participants were asked to wait 5s before speaking (i.e., to let the vehicle pass-by noises end). Then, they had 15s to return orally as many words memorized as possible, knowing that each word recalled correctly earned 1 point, and each word recalled wrongly removed 1 point.

Furthermore, they completed a French version of the NASA-RTLX (Raw Task Load Index) questionnaire [28], which measures the level of workload using 6 aspects (i.e., mental activity, provided effort, frustration, etc.). They also assessed noise annoyance through the following question: “At which point did the sound sources affect you in accomplishing the task requested?” The 6 NASA-RTLX questions and the noise annoyance question were mixed and randomized for each sound sequence. To answer each question, participants could take as much time as they wanted and they had to place the cursor of the computer mouse on the scales – as it has been explained about group (a), for annoyance scale.

Regarding noise environment, the background noise was emitted continuously at 45 dB(A) during the experiment. The 3 pass-by noises of each sequence appeared after the second fixation cross. During the participant’s restitution and completion of questionnaires, only the background noise was diffused.

### 3.2 Results

We introduce here the first outcomes of this experiment: annoyance variability (Sec. 3.2.1), relationship between annoyance and performance (Sec. 3.2.2), and model accuracy (Sec. 3.2.3).

#### 3.2.1 Noise annoyance evaluation – group comparison

The annoyance rates within the groups (a) and (b) are averaged over all the participants for each of the 28 combinations of vehicle category vs. noise level. Figure 1 shows the averaged annoyance rate of group (a) with regards to group (b). A linear relation between the two set of rates can be noticed. We obtain a linear regression with a correlation coefficient of 0.88 ($p<.001$) and a well-spread distribution along this line, meaning that the average of annoyance judgment are significantly similar.

![Figure 1: Linear regression between mean annoyance assessed in group (a) vs. group (b). Each point is a vehicle category (1 to 7) at a given sound level (50 to 62 dB). Pearson’s coefficient correlation is 0.88, ($p<.001$)](image-url)
These results address the issue of annoyance definition and its difference to unpleasantness. Indeed, whereas annoyance is commonly linked to an action [31], our data seem to contradict this point of view and show a similar percept, independent of the listener’s state (passive or active).

3.2.2 **Performance as a good annoyance index**

In group (b) the words recall are used to compute participants’ performances. Raw performance (number of good words minus bad words) is normalized by the average performance during the 4 sequences without vehicle events. Figure 2 shows perceived annoyance vs. normalized performance, averaged over the participants, for each of the 28 audio combinations. Data show a certain correlation between these two dimensions. The plain regression line is obtained for the entire dataset. It gives a correlation coefficient of -0.52, which is not entirely satisfying. We can notice that vehicles in acceleration (C2, 6, 7) passing-by at 62 dB(A) form a cluster of “outlier” data. The dotted regression line is obtained for the dataset without these combinations. It gives a better correlation coefficient of -0.71 ($p<0.05$).

![Figure 2: Linear regression between mean annoyance and performance data.](image)

3.2.3 **Estimated vs. measured annoyance**

Using Morel et al.’s implementation (Sect. 2), we confront its outcomes with our experimental data, (group (a)). Results are presented in Figure 3. They show a significant correlation ($p<10^{-3}$) between the model and the experiments with a coefficient $r = 0.89$. We can notice that the major part of the categories is well-spread around the regression line. After all, two of them (C2, 6) seem to deviate from it.

![Figure 3: Linear regression between estimated annoyance and experimentally measured annoyance (group (a)).](image)
The model could be adapted for these two categories by modifying the coefficient governing the dependence on loudness (for C2) and the sound level in medium frequency (for C6). Finally, a global multiplicative correction coefficient of 0.58 seems to be needed to re-scale Morel et al.’s model.

4. Conclusion – Perspectives

This paper presents the whole frame of a large scale research program regarding the characterization of urban soundscapes, according to road traffic sources. It focuses on the last part of the program that implements an experimental validation of a perceptual noise annoyance model. The experimental procedure is precisely described, in addition to a complementary paper that deeper addresses the audio rendering ecology issue [22]. The results obtained show a good consensus between two groups of participants: (a) passive assessment, (b) active assessment with a cognitive task during the listening test. This may lead to challenge the common distinction between annoyance and unpleasantness. This being, the results also find out that a cognitive task performance measurement (e.g., memory) can be used, in a certain range of sources, to predict the annoyance percept. Finally, and most importantly, the results validate, to some degree, the initial perceptual model of annoyance established by Morel et al. [15], and give some cues to improve it.

Perspectives of this work are the following: i) to consolidate the results with complementary analysis and/or additional experiments; ii) to improve the model, regarding either psychoacoustic features refinement or global annoyance implementation. The latter point leads to an important – an yet unresolved – issue regarding the summation process that occurs with a complex mixture of sound events.

5. Acknowledgements

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

REFERENCES


