



## Modeling task interference caused by traffic noise: results from a perception experiment

Jonathan Terroir<sup>1</sup>, Bert De Coensel<sup>2</sup>, Dick Botteldooren<sup>3</sup>, and Catherine Lavandier<sup>4</sup>

<sup>1,4</sup>Laboratoire Mobilités Réseaux Territoires et Environnements (LMRTE), Université de Cergy-Pontoise  
Neuville sur Oise, 95031 Cergy-Pontoise Cedex, France

<sup>2,3</sup>Acoustics group, Department of Information Technology (INTEC), Ghent University  
St.-Pietersnieuwstraat 41, B-9000 Ghent, Belgium

### ABSTRACT

This work compares experimental and simulation results on task interference, caused by exposure to combined traffic noise. On the one hand, a perception experiment was performed, conducted in a realistic setting resembling an at-home situation. Subjects were instructed to read, while being exposed to a combination of road and railway traffic noise. The number of train pass-by events, the distance to the railway track and the emergence of train events above the background noise was varied among subjects. After completion of the reading task, the subjects had to evaluate their perceived disturbance due to passing trains and to report how many trains they noticed. On the other hand, a computational model of auditory attention was used to determine the number of train pass-by events that subjects would notice, solely based on the acoustic stimuli used in the perception experiment. Using an optimized function that simulates the reading activity of the subjects, the model was able to replicate trends found in the empirical results, and estimated the number of noticed train events quite well.

Keywords: Train noise, Task interference, Auditory attention

### 1. INTRODUCTION

Selective auditory attention plays an important role in the perception of the complex acoustic environment to which humans are exposed [1]. In the long term, the sounds to which we pay attention will contribute to the creation of a mental image of our acoustic environment, and ultimately will shape our perception of its quality. When noise annoyance in an at-home context is considered, the interference of noise with the task or daily activity a person is performing, is essential [2]. There is clear evidence that the presence of irrelevant information (sounds) degrades selective attention, impairing the performance on the task at hand [3]. In particular, salient sound events may attract auditory attention, and therefore distract attention from the task at hand. In addition, traffic noise intruding one's dwelling usually has a negative connotation. Therefore, a good knowledge of the process of auditory attention is essential in order to understand the underlying mechanisms relating noise exposure, task performance and activity disturbance, noise annoyance and soundscape quality in an at-home context.

In the present work, results of a perception experiment on task performance under noisy conditions (exposure to road and railway traffic noise) are compared with simulation results of a model of auditory perception. In Section 2, the experimental methodology is explained. In Section 3, an overview is given of the computational auditory attention model. Finally, in Section 4, experimental results are compared with results of simulations using the computational model. The present work is the result of a collaborative effort between the University of Cergy-Pontoise and Ghent University.

---

<sup>1</sup> jonathan.terroir@u-cergy.fr

<sup>2</sup> bert.decoensel@intec.ugent.be

<sup>3</sup> dick.botteldooren@intec.ugent.be

<sup>4</sup> catherine.lavandier@u-cergy.fr

## 2. PERCEPTION EXPERIMENT

The perception experiment was carried out within the framework of a PREDIT/DEUFRAKO project on the influence of the temporal structure of transportation noise on annoyance, cognitive performance and sleep, initiated in 2006. A pilot study on the sound quality of single train pass-bys showed that temporal (psycho)acoustic dimensions such as the irregularity of the sound (rhythm during the pass-by) or the suddenness of the sound arrival (the slope of the level increase) were features allowing people to discriminate different current French train pass-bys, while the spectral dimension appeared of less importance [4]. On the one hand, the irregularity of the train sound was especially heard for recordings carried out next to the track (at a distance of 7.5m). However, few people in France live within such a short distance of a railway track. On the other hand, the suddenness characterized sound events recorded at 50m and 100m (a more common distance between dwelling and railway track). A new experiment was then carried out in a laboratory context, in order to investigate the influence of the distance to the track and the rate of train pass-bys on annoyance. A detailed description and results of this study can be found in [5]. Below, the experimental procedure is shortly described.

### 2.1 Sound stimuli

Stereo pass-by recordings were made for four different types of trains: TGV (high-speed trains), CORAIL (express trains), TER (regional trains) and FRET (freight trains). Subsequently, 8 one-hour stimuli were constructed, in which road traffic noise (a continuous background with no clear sound events) was mixed with a number of train pass-bys. Three parameters were varied among stimuli:

1. *The number of pass-bys.* To simulate a realistic multi-exposure situation, the average rate of train pass-bys in France between 6pm and 10pm was used (41 events for 4h). A regular number of 10 pass-bys per hour was considered (6 CORAIL, 2 FRET, 1 TGV, 1 TER), as well as a high number of 20 pass-bys per hour (12 CORAIL, 4 FRET, 2 TGV, 2 TER). The pass-bys were randomly distributed over time.
2. *The distance to the railway track.* Distances of 50m and 100m were considered.
3. *The dominance of the train pass-bys.* This parameter codes the emergence of the train sound events over the background. Two situations are considered: a dominant situation, in which the level of the railway sound is 6 dB(A) higher than the level of the road traffic noise, and a non-dominant situation, in which the level difference is only 3 dB(A).

As each parameter has only 2 states, in total, a number of  $2^3 = 8$  stimuli were considered. Finally, to avoid any effects of loudness, all stimuli were normalized to an  $L_{Aeq,1h}$  of about 52.5 dB(A). The difference in distance to the track resulted in differences in spectral composition of the stimuli, in rise time of the sound events, and in perceived duration of the pass-bys. Table 1 summarizes the acoustical properties of the stimuli. Note that, as the global sound level was equalized, the values of train level peaks vary between scenarios: the higher the number of train pass-bys, the lower the peak levels.

Table 1 – Scenario properties

| Scenario | Pass-bys | Distance | Dominance | $L_{Aeq,1h}$ [dB(A)] |        |         | $L_{max}$ [dB(A)] |
|----------|----------|----------|-----------|----------------------|--------|---------|-------------------|
|          |          |          |           | Total                | Trains | Backgr. |                   |
| 1        | 10       | 50m      | Dom       | 52.5                 | 51.5   | 45.5    | 72.0              |
| 2        | 20       | 50m      | Dom       | 52.2                 | 51.2   | 45.2    | 67.7              |
| 3        | 10       | 100m     | Dom       | 52.3                 | 51.3   | 45.3    | 69.6              |
| 4        | 20       | 100m     | Dom       | 52.4                 | 51.4   | 45.4    | 66.5              |
| 5        | 10       | 50m      | Non-Dom   | 52.4                 | 50.6   | 47.6    | 70.0              |
| 6        | 20       | 50m      | Non-Dom   | 52.7                 | 50.9   | 47.9    | 67.0              |
| 7        | 10       | 100m     | Non-Dom   | 52.7                 | 50.9   | 47.9    | 69.0              |
| 8        | 20       | 100m     | Non-Dom   | 52.8                 | 51.0   | 48.0    | 66.0              |

## 2.2 Apparatus

The stimuli were played back in a quasi sound-proof room, by use of a pair of loudspeakers and a subwoofer. In order to create a relatively realistic indoor situation, a picture of a half-opened window was projected in front of the subjects. Additionally, a filter was applied to the stimuli, in order to simulate the outdoor-to-indoor propagation of sound, including the effect of a half-opened window.

## 2.3 Subjects

Each stimulus was presented to 20 different subjects. The group of subjects was homogenous, mainly formed of students, with an average age of 22y. In order to avoid any learning effect, each subject was submitted to only one scenario. In total, 160 participants were involved in the experiment.

## 2.4 Experimental procedure

In order to recreate a common at-home situation during the evening, it was necessary to select an activity which is commonly performed by people at home, and which does not need an excessive amount of concentration, in order to avoid that subjects would be disjoined from the sound environment. The reading of comic books was finally selected. Subjects had to read comics for one hour, while a stimulus was played back. At the end of the hour, subjects were asked to rate their perceived noise annoyance, and to state how many trains they noticed. In this paper, only the perceived number of pass-bys is considered. The experimental design was finally investigated through an analysis of variance, in which the influence of each variable on its own was considered, as well as the interactions between variables.

## 3. COMPUTATIONAL ATTENTION MODEL

A computational model of auditory attention to environmental sound, developed at Ghent University [6], is applied to the experimental stimuli, in order to estimate the number of train pass-bys that would actually be noticed. This model simulates how listeners switch their attention over time between different auditory streams. The model is inspired by the structure of human auditory system, a detailed description of the model can be found in [6]; here only the most important ideas are outlined.

### 3.1 Saliency-based auditory attention

In a complex sound environment, humans base their perception on a detailed analysis of the auditory scene (ASA) [7]. Through this process, the mixture of incoming sounds is decomposed into several auditory streams, each associated to a perceived sound source. This decomposition is based on audio and visual cues. Subsequently, auditory attention is focused to a single stream, but the stream that receives attention may switch over time. Usually, ASA is considered to be an analysis-synthesis process. Firstly, the sound signal is decomposed into a collection of time-frequency segments; subsequently, auditory streams are created as a function of their likeness to arise from the same source.

Computational models for selective (auditory) attention are usually based on an interaction between two cues: bottom-up cues (based on the time-dependent saliency of each stream) and top-down cues (based on the amount of volitional focusing on particularly streams) [8]. The bottom-up mechanism selectively gates incoming audio information, enhancing responses to stimuli that are salient. This is a pre-attentive mechanism, independent from the activity or task of the person. Bottom-up saliency of auditory streams can be estimated through the calculation of a so-called saliency map, which provides a weighted representation of the sonic environment, emphasizing those time-frequency elements that are more conspicuous, and thus most likely detected [9]. The calculation of a saliency map depends on the use of a set of receptive filters applied to the spectrogram of the sound, and which quantify intensity, spectral contrast and temporal changes. The top-down mechanism focuses mental resources on the treatment of the most relevant information for the current goal-directed behavior of the listener. These mechanisms are guided by information already held in working memory through sensitivity control (activity or intention). A competitive winner-takes-all process finally determines which stream is selected for entry in working memory. An important factor in this dynamic process is also the inhibition-of-return (IOR) mechanism, which prevents attention from permanently focusing on the most salient auditory stream.

The above ideas have been implemented in a computational model, specifically targeted for environmental sound [6]. The model takes as input the sound signals of the different sound sources, and returns, on a statistical basis, the source of focus as a function of time.

### 3.2 The non-auditory function

An essential part of the model is that also a non-auditory stream has to be considered, as otherwise, attention would always be directed towards the auditory sense, which is not a realistic situation. This non-auditory stream simulates the distraction caused by non-auditory features (thought, visual events, activity etc.). It can be tailored to mimic a predefined activity pattern, or it can be implemented to be random. In this work, the non-auditory stream represents the amount of (bottom-up) attention that is required by the reading activity over time. It was modeled as the superposition of two random functions: a stationary level of bottom-up attention (slightly oscillating around its mean value), mimicking the continuous nature of the reading activity, combined with sudden peaks or drops in bottom-up attention, simulating positive or negative modulations of attention during reading. The latter were randomized for strength, duration and occurrence in time. Note that the proposed structure is linked to the activity itself; e.g. a task in which a calculation has to be done every five seconds would probably require a totally different structure, in order to achieve reasonable results. Finally, randomized values within given ranges were also chosen for the reference attention level due to the reading activity as well as for the rate of increases/drops in non-auditory attention, modeling the variance of experimental results due to inter-subject differences.

## 4. RESULTS

In this section, results of the laboratory experiment (in particular, the perceived number of train pass-bys) are compared with the number of noticed pass-bys, as estimated by the attention model. In accordance with the laboratory conditions, a simulation was run for every real participant in the experiment. As the model is only valid on a statistical basis, it is only meaningful to compare average trends (i.e. personal characteristics of the subjects are averaged out). Figures 1 to 4 show the averaged results of the simulations.

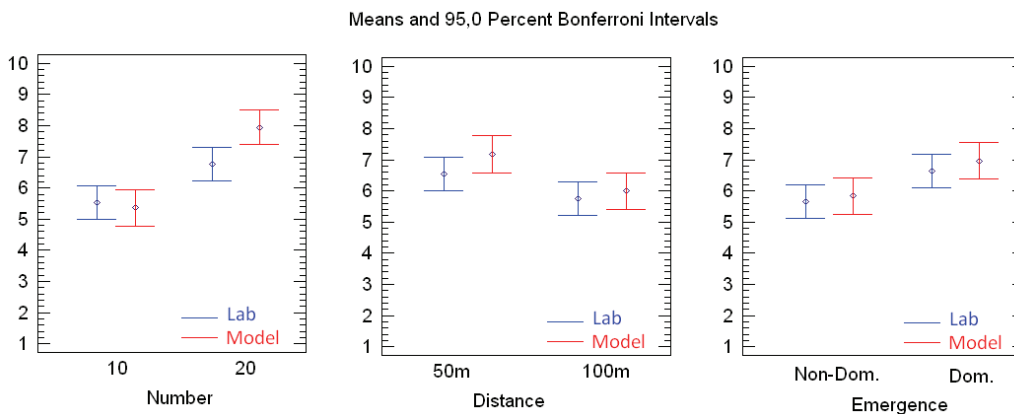


Figure 1 – Influence of (a) the number of train pass-bys, (b) the distance to the track, and (c) the dominance on the reported/simulated number of perceived train pass-bys.

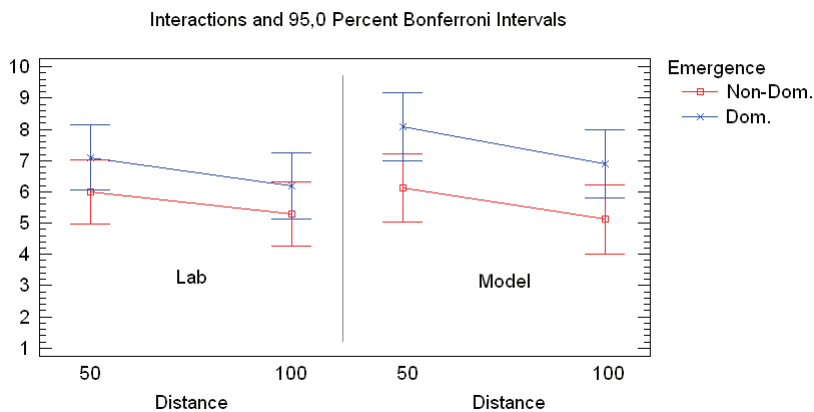


Figure 2 – Interaction between distance to the track and dominance on the reported/simulated number of perceived train pass-bys.

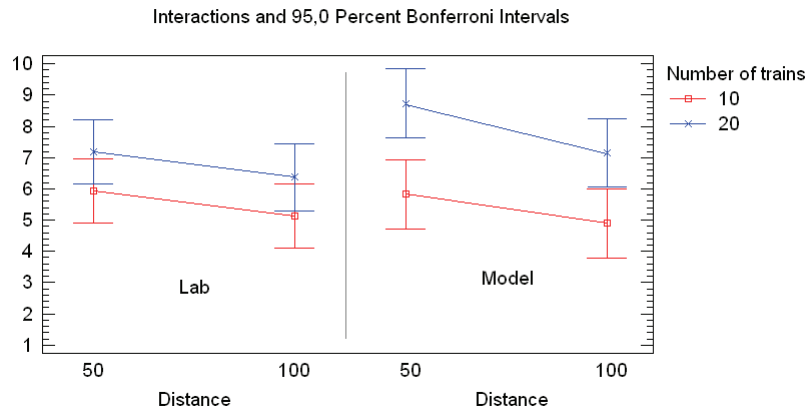


Figure 3 – Interaction between distance to the track and the number of train pass-bys on the reported/simulated number of perceived train pass-bys.

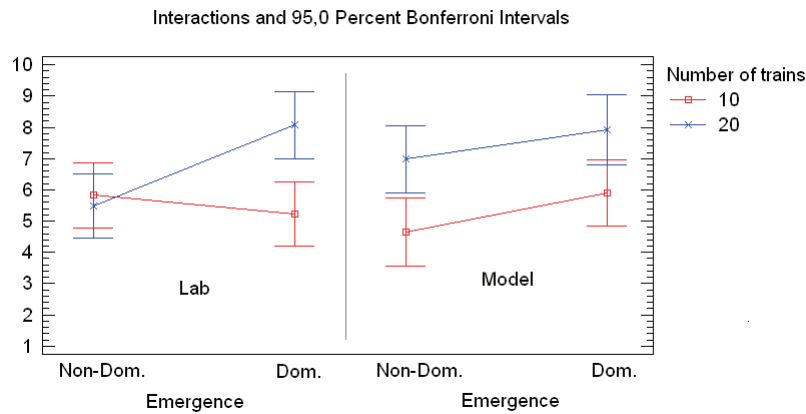


Figure 4 – Interaction between the dominance and the number of train pass-bys on the reported/simulated number of perceived train pass-bys.

As it can be observed in Figures 1 to 4, the influence of each variable (distance to track, number of train pass-bys and dominance) on the number of perceived train pass-bys is similar for the model and the laboratory experiment. The non-auditory function design appears to be adapted to the simulation of the reading activity, and with well chosen values for the model parameters, simulated results are in accordance with experimental ones. Additionally, the design of the non-auditory function (in particular the reference level range) results in comparable Bonferroni intervals, which are essential for the realism of the model results, by showing the dispersion of the subjects' answers, taking into account inter-subject differences.

An interaction between the dominance and the number of trains was experimentally observed but could not be found for any simulation (with any set of model parameters). After a detailed analysis of the experimental results, it appeared that some people reported 12 or 15 train pass-bys for stimuli 5 and 7 (see Table 1), although there were only 10 pass-bys in reality (a behavior which could not appear within the model). Thus, this interaction could be due to an overestimation of the number of trains. It can be expected that there is a difference between the reported number at the end of the experimental session, and the actually noticed number during the experimental session, because memory plays a role in this process. A "memory model" that would take into account this difference is not included in the model considered. It could be the case that the influence of this memory aspect increases with the number of events, as it becomes harder to remember the number of noticed pass-bys. Experimental and model results appear to be relatively close, such that, as a first approximation, it can be assumed that the number of over- and under-estimations cancelled out.

## 5. CONCLUSIONS

In this work, results of a perception experiment, in which subjects were instructed to read a comic book while being exposed to a combination of road and railway traffic noise, were compared to simulation results using a computational model of auditory attention. In particular, the reported and simulated number of perceived train pass-bys was compared. Using an optimized non-auditory attention function, which simulates the reading activity of the subjects, the model was able to replicate trends found in the empirical results, and estimated the number of noticed train events quite well. Moreover, the simulation results provide an explanation linked to the reading activity. As this work is still in progress, a more detailed statistical analysis on the relationships between annoyance, number of perceived train pass-bys and other stimuli properties will have to be carried out in the future.

## ACKNOWLEDGEMENTS

This collaborative study was performed within the framework of a Short Term Scientific Mission (STSM), supported by the European Cooperation in Science and Technology (COST) Action TD0804 “Soundscape of European Cities and Landscapes”. Bert De Coensel is a postdoctoral fellow of the Research Foundation–Flanders (FWO–Vlaanderen); the support of this organization is also gratefully acknowledged.

## REFERENCES

- [1] J.B. Fritz, M. Elhilali, S.V. David and S.A. Shamma, “Auditory attention – focusing the searchlight on sound,” *Curr. Opin. Neurobiol.* 17(4), 437-455 (2007).
- [2] K. Zimmer, J. Ghani and W. Ellermeier, “The role of task interference and exposure duration in judging noise annoyance,” *J. Sound Vib.* 311(3-5), 1039-1051 (2008).
- [3] S.P. Banbury, W.J. Macken, S. Tremblay and D.M. Jones, “Auditory distraction and short-term memory: phenomena and practical implications,” *Hum. Factors* 43(1), 12-29 (2001).
- [4] J. Terroir and C. Lavandier, “Qualité sonore des trains TGV: influence des facteurs perceptifs - analyse multidimensionnelle et préférentielle,” *Proc. Congrès Français d’Acoustique*, Lyon, France (2010).
- [5] C. Lavandier and J. Terroir, “Functional annoyance due to railway traffic noise in a laboratory context,” *Proc. Euronoise*, Edinburgh, UK (2009).
- [6] B. De Coensel and D. Botteldooren, “A model of saliency-based auditory attention to environmental sound,” *Proc. ICA*, Sydney, Australia (2010).
- [7] A.S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. The MIT Press, Cambridge, Massachusetts, USA (1994).
- [8] E.I. Knudsen, “Fundamental components of attention,” *Annu. Rev. Neurosci.* 30, 57-78 (2007).
- [9] C. Kayser, C. Petkov, M. Lippert and N.K. Logothetis, “Mechanisms for allocating auditory attention: an auditory saliency map,” *Curr. Biol.* 15(21), 1943-1947 (2005).