Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Noise
Session 2aNSa: Transportation Noise

2aNSa4. Detectability study of warning signals in urban background noises: A first step for designing the sound of electric vehicles

Nicolas Misdariis*, Anais Gruson and Patrick Susini

***Corresponding author's address: STMS Ircam-CNRS-UPMC, 1, place Igor Stravinsky, Paris, F- 75004, Paris, France, misdarii@ircam.fr**

Electric vehicles, tends to become a growing category of today' s human means of transport. But, because these kind of vehicles are actually quiet, or even silent, the question of a dedicated sound design arise almost inevitably in order to make them more present – then secure – both for their proximity (pedestrians) and their users (driver). This being, current issues for a sound design research framework is then to exploit and explore sound properties that, first, will fix a goal of functionnality (emergence, recognition, acceptance) and, second, will define guidelines for the development of new aesthetics to be included in a general design approach. Thus, a first study focusing on detection of warning signals in urban environments was achieved. Based on the state-of-the-art, a corpus of elementary signals was built and characterized in a time / frequency domain for representing basic temporal and spectral properties (continuous, impulsive, harmonic, etc.). A corpus of representative urban environments was also recorded and realistic sequences were mixed with a dynamic approaching-source model. A reaction time experiment was conducted and leads to interesting observations: especially, specific properties promoting emergence. Moreover, a seemingly significant learning effect also rises from the data and should be further investigated.

Published by the Acoustical Society of America through the American Institute of Physics

Received 21 Jan 2013; published 2 Jun 2013

INTRODUCTION

Are hybrid cars too quiet ? This question, posed by Robart & Rosenblum some years ago [1], can define quite well the general framework of the present study. In fact, the main specificity of Quiet Vehicles (QV) – either electric (EV) or hybrid in electric mode (HEV) – is that they are rather completely zero emission with regards to sound, at least in a certain range of use. Thus, the increasing amount of this mode of propulsion in today's industrial panorama leads to two main issues: i) *integration*: how does a quiet car can cohabit with a noisy car, especially in some specific contexts where cars are not the only elements of the sonic environment (urban, peri-urban); ii) *transformation*: what will be the impact of this rather silent mode of propulsion and transportation in our daily soundscape, with reference to Schafer's terminology [2].

Then, the way to analyze the intrinsic quietness of QV and actions proposed to deal with the global issues mentioned above seem to point out some divergence. Particularly, the solution that will consist in equiping the vehicle with an additional warning sound gets its pros and cons. Among others, Sandberg & al. argued that tyre/road noise level is sufficient enough in many current traffic conditions for signifying the presence of the vehicle and moreover, that the difference between "normal" (i.e., with internal combustion engine – ICE) and electric vehicles is rather negligeable in urban area at the present time, mainly because of the pregnant efforts made to reduce ICE noise [3]. On the opposite side, Le Nindre & Guyader have compared the impact of both kinds of vehicle (EV *vs.* ICE) on a real usage scenario: the vehicle is parked on a parking lot, about to start and move into the traffic. By measuring the noise level, in dB(A), for three representative chronological states of this scenario (start, idle and acceleration), they have clearly shown that the acoutic cues related to ICE are delivered faster along the scenario and louder above the urban background noise; the EV acoustic cues being only relevant in the middle of the third state (acceleration), i.e. globally too late to warn the pedestrians that the car is in motion [4].

This being, it seems nowadays widely accepted that Quiet Vehicles don't provide enough auditory cues either for direct (drivers) or indirect (pedestrians, cyclist, etc.) users and that there is a real need for conceiving a well-grounded QV sound signature. Indeed, as the Quiet Road Transport Vehicles United Nations working group officially mentioned in one of its reports: "*vehicles propelled in whole or in part by electric means, present a danger to pedestrians"* [5]. At that point, this general framework can represent a good opportunity to work on a dedicated sound design for Quiet Vehicles in order to fulfill both functional and aesthetics requirements. Sustainable solutions must be relevant either in terms of safety to face at the risk QV can potentially represent (weak presence for its vicinity and lack of feedback for the driver) but also in terms of integrity with regards to environmental noise impact and pollution. The next challenge seems now to define which type of sounds could be optimum to solve this problem.

In that point of view, the present study tries to contribute to this challenge by focusing on the notion of detectability and exploring basic sound properties able to promote the detection of prototypic warning signals, moreover, taking into account the influence of the background noise. The paper presents the chronological components of the study: a selective bibliographic review – state of the art – that led to the definition of warning signals and background noises typologies; a collection (retrieved or recorded) of several examples of these differents types of sounds and a physical caracterization of these corpus; a 3-step experimental procedure based on a reaction time paradigm measuring time values for detection – and additionaly collecting qualitative judgments (preference, acceptability); an analysis of the results pointing out the infuence of different factors (signal properties, type of backgound noise, number of repetitions) and their mutual interactions; finally a conclusive discussion suggesting different tracks for the pursuance of the work.

STATE OF THE ART

Compared to other wider applications, the bibliographic ressources concerning the specific domain of Quiet Vehicle external sound is not very provided – except if one consider all the types of publication such as technical patents, official reports coming from institutional organizations (UN or national associations), unofficial notes from manufacturers' technological watching technological or the numerous press articles released on this topic. Nevertheless, from the very end of the last century and Otto's article on "electric vehicle sound quality" [6], there is a significant increasing number of scientific papers framing with this particular scope; all the more that it can be extended to a larger formalized framework – controlled insertion of a sonic interactive object in a given environment – that can be related to more general keywords as *relevance*, *emergence*, *acceptance* or *annoyance*, and then more general topics as alarm perception, interactive sonification or acoustic ecology.

Among all of this, the present study have only focused on a selective list of contributions able to inform two important points of the experimental approach : i) which type of sounds will be tested and which type of background noises will be considered ? ii) how does the laboratory experimental protocol will be conceived to render in a pseudo-realistic way some elements of context such as audibility, immersion or panoramization ?

Wogalter & al. [7] undertook a large survey (380 persons) to explore interests and concerns about electrically powered cars. The first general result is that 70% of the participants think that the lack of sound on a Quiet Vehicle is a source of danger for pedestrians because approximately the same amount of persons concede that sound is useful for detecting an approaching vehicle. More precisely – with oriented questions like "What type of sound do you recommend be implemeted (e.g., whistle, hum, engine noise, chimes, etc.)" – the study concludes in a collection of lexical suggestions and recommandations about the nature of sound to associate with QV: these were, first, traditional engine or hum sounds (38%), and secondly, music (14%), whistle (8%), beeps (5%) or horns (5%).

Nyeste & Wogalter [8] conducted a preference study to determine types of sounds that may be eligible auditory cues for Quiet Vehicles. The study built a sound corpus from Wogalter's lexical results established previously (cf. section above). A corpus of 18 sounds were selected on the basis of a 6-class typology: engine, hum, horn, siren, whistle and white noise. The soundfiles were played back with a video support (a commercialized EV model) and were evaluated on a semantic scale, from "not at all acceptable" to "extremely acceptable". Main results point out two significantly different groups ({engine, hum, white noise} *vs.* {horn, siren, whistle}) and show that engine and horn categories are respectively the most and the least preferred.

Menzel & al. [9] investigated the interaction between warning signal and background noise typologies within a parametric experiment. They studied the level of three potential warning sounds (car horn, gasoline engine at idle and bursts of band-filterd white noise) adjusted in four environmental background noises: a two-lane busy street in downtown, a two-lane road in residential area, a narrow road in a shopping area and a 6-lane heavy traffic. They were presented in a laboratory environment (sound-proof booth and headphones) in order to be adjusted by participants at two distinct levels : "just audible" and "clearly audible". Results estimate a 10 to 20 dB difference between these two thresholds and show that they both strongly depend on either backgrounds or warnings nature. Yamauchi & Menzel [10] transposed the study within an intercultural framework – Germany and Japan – and globally found any significative difference between these two populations

Kerber & Fastl [11] proposed a quantitative prediction of perceptibility (named "perception-distance") by comparing computed masked thresholds, with regards to a given background noise level, and measured vehicle's sound level along time – or vehicle's position. This latter weighting-function, physically measured on several cars and approximated by a logarithmic model, corresponds to the level of an approaching vehicle from 35 m. on the left to 0 m., i.e. when the car is passing in front of the listener ; it is assumed to be symmetric.

EXPERIMENTAL APPROACH

As mentionned in the introduction, the objective of the study was to make a first stage of exploration about the basic sound properties that best enhance the detection of Quiet Vehicles in a given background noise. The work was mainly experimental and consisted in several steps before its core experiment, a reaction time measurement (presented in section 4): 1.- forming and caracterizing sound corpus for both warning signals (warnings) and masking signals (backgrounds); 2.- defining the audibility threshold for all warnings in all backgrounds; 3.- operating a loudness equalization of all warnings.

1- Stimuli

Warning signals

For the constitution of the warning sound corpus, the six categories obtained by Nyeste & Wogalter [8] are firstly retained: engine, hum, whistle, horn, siren, white noise. This initial taxonomy is completed by a seventh one: clicking, on the basis of Wogalter's previous conclusions [7]. Then various instances of these seven categories are retrieved in commercial sound databases available at Ircam (SoundIdeas[1], Hollywood Edge[2], BlueBox[3]). Finally, 18 sounds are selected according to a rather balanced distribution (3 sounds per category, except for hum (2) and white noise (1)). Sounds are mainly chosen according to a criterion of causality: even if the categories are explicitly

[1] Sound Ideas General Series 6000 (Sound Ideas, Ontario, Canada – www.sound-ideas.com)
[2] Hollywood Edge Premiere Edition I, II and III (the Hollywood Edge, Hollywood, USA)
[3] Blue Box Audio Wav (Best Service GmbH, München, Germany)

labeled, a sound that doesn't clearly suggest the source or the production mechanism is generally prefered to a more causal one, so that its acoustic properties could be perceptually more predominant than its causality (see [12] for the distinction between acoustic and causal levels of perception of environmental sounds).

The initial corpus is then caracterized according to their acoustic properties. At a first level, these properties are roughly described by three gross categories in each dimension of the sound (time and frequency) : *continuous*, *modulated* or *burst* (serie of pulses) in the temporal domain; *harmonic*, *inharmonic* or *noise* in the spectral domain. At a second level, a precise modeling is achieved by means of computation of audio features extracted from the signal. The toolbox Ircamdescriptors[4] is used to compute a large number of temporal, spectral or spectro-temporal features [13], essentially based on FFT analysis and Fundamental Frequency (F0) computation which parameters (window size, hop size, etc.) are optimized for the present corpus.

After running this routine, two features are found to be consistent with the first level of modelisation mentionned above and sounds of the corpus: a *spectral flatness* and a *modulation rate*. These features are detailed below:

• *spectral flatness* (SFM) is a measure of the noisiness – or sinusoidality – of a spectrum (or a part of it); it is computed by the ratio of the geometric mean to arithmetic mean of the energy spectrum value (cf. Eq. 1) in each considered frequency band. Typically, SFM is close to 0 for tonal sounds and close to 1 for noisy signals.

$$SFM = \frac{(\Pi_{k=1}^{K} a(k))^{1/K}}{\frac{1}{K} \cdot \Sigma_{k=1}^{K} a(k)} \quad , \quad where \quad a(k) = amplitude \;\; in \;\; frequency \;\; band \;\; k \tag{1}$$

• *modulation rate* (m) is a measure of the signal modulation; it is figured out from Ircamdescriptors' energy envelop computation (i.e., evolution of the energy along the course of the signal). From that time serie, values of maximum and minimum energy are extracted and used to compute modulation rate that is the ratio between difference and sum of these extreme values of the energy envelop (cf. Eq. 2)

$$m = \frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}}, \quad where \quad E = \;\; energy \;\; envelop \tag{2}$$

These two features allow to put the initial sound corpus (18 sounds) in a two-dimensional space describing their temporal and spectral properties. On the basis of regular distribution of items in this space, a selection is then made among the 18 sounds. This representation also allows to point out some gaps in the description space and lead to synthesize new sounds as "intermediate sounds" between two existing items in order to consistently fill these gaps. Finally, the resulting corpus contains 10 sounds (9 extracted from the corpus and 1 obtained by synthesis), labeled as follows: *modulated white noise*, *white noise*, *burst of white noise*, *clicking*, *klaxon*, *electric hum*, *modulated electric hum*, *modulation*, *siren* and *whistle* (Figure 1). [Sounds are available at http://pds.ircam.fr/ica2013_sounds.html]

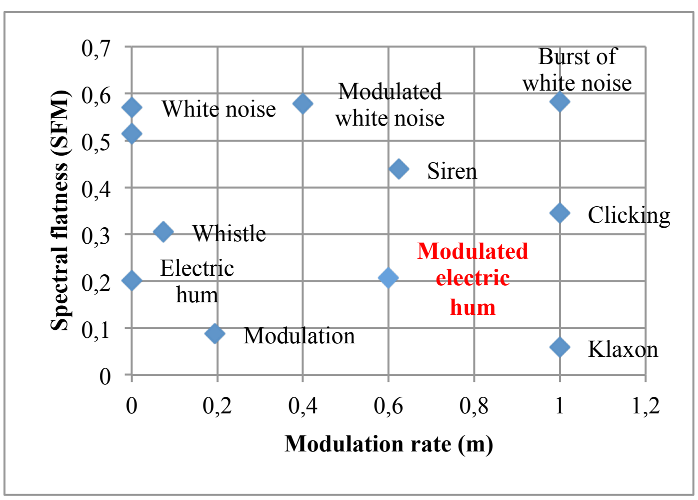

FIGURE 1. Distribution of the final corpus (10 sounds) in the description space: modulation rate *vs.* spectral flatness. Item marked in bold-red is obtained by synthesis.

[4] http://www.ircam.fr/1041.html?&L=1

Masking signals

Considering the masking signals, and on the basis of Menzel typologies [9], three representative background noises are retained for the present study: the two-lane busy street in downtown (named "*busy*" in the rest of the paper), the two-lane road in residential area (named "*residential*") and the narrow road in a shopping area (named "*shopping*"). The fourth Menzel's typology (6-lane heavy traffic) is rejected mainly because it seems rather unlikely to meet pedestrians or cyclists in such a urban configuration.

The corresponding soundscapes are recorded in situ by means of a portable digital recorder (Zoom H4N[5]) using the built-in stereo microphone in the X-Y configuration. The device – equipped with a windshield – is fixed on a tripod and placed at about 1,00 m. in height and 0,5 m. in distance with regards to the urban flux to be recorded. Each recording session is also associated to a sound level measurement by means of a sonometer (Bruel&Kjaer Mediator 2238[6]) that delivers a $L(A)_{eq}$ value, averaged along the course of the recording.

For that purpose, recording conditions are chosen according to Kerber's recommandations [11] saying that the recorded soundscape must not contain any strong emergent event (especially car passing, but also horns, etc.) in order to be used afterwards for a detection experiment. This condition can be related to the notion of "amorphous soundscape" defined by Maffiolo [14] and is quite essential in order to avoid important sound level variations and consequent masking effects. This condition has been strongly respected for two of the three selected backgrounds (residential and shopping), and as much as possible for the third one (busy).

In practice, three particular streets have been selected in the 1st and 4th district of Paris: Reynie street (residential), Sebastopol boulevard (busy) and Ferronnerie street (shopping). Several takes have been carried out for each of these spots and a selection has been made afterwards, according to sound quality and homogeneity criteria (as mentioned above). Finally, one sequence of 15 sec. has been kept, for each urban condition; respective sound levels are as follows (in dB(A)) : residential (65.1), shopping (56.1), busy (71.4).

Then, at the end of this first experimental step, the final corpus are as follows : 10 warning sounds (named "*warnings*" in the rest of the paper) representing two physical properties in the time and frequency domain (continuity, harmonicity), each of them caracterized by an acoustical feature computed on the audio signal; 3 background noises (named "*backgrounds*" in the rest of the paper) representing three distinctive typologies of urban soundscapes. [Sounds are available at http://pds.ircam.fr/ica2013_sounds.html]

2- Just Audibility Threshold

The aim of this first preliminary experiment was to define a "just audibility threshold" – as defined by Menzel & al. [9] – for a reference warning in each of the three backgrounds. This reference level will be adjusted afterwards to a clearly audible level and then used for the loudness equalization step (section 3). The reference chosen is the *white noise* warning according to its neutral properties compared to other items of the corpus.

The test took place in a double walled IAC soundisolation booth equipped with Yamaha MSP5 Studio loudspeakers, a RME Fireface 400 audio card and a Macintosh Mac Pro (OSX v10.5) workstation. It was run with PsiExp [15], a Graphical User Interface (GUI) developed in C/C++ and Java including stimulus control, data recording, and sound play-back. It has been performed by 11 persons (3 females, 8 males).

At the beginning of the procedure, participants were given written instructions presenting the experimental task but without detailing the overall context of the study (i.e. VE sound design application). The task was to adjust the "just audibility threshold" of the reference warning (*white noise*), mixed with a given background, by means of a potentiometer (horizontal cursor) related in real time to the reference sound level. The starting level of the reference was alternatively 0 (inaudible signal) and 1 (clearly audible signal) so that participants will achieved respectively a up and down process for the adjustment task. This procedure was repeated twice for each background. The experimental factors (background order, starting reference level) were selected randomly.

For all participants and for each background, means and standard deviations are computed for the 4 linear values of the adjustment coefficients (2 from down starting level, 2 from up starting level). Results are as follows: residential (C1 = 0.227), shopping (C2 = 0.287), busy (C3 = 0.433). An analys of variance (ANOVA) is also realized. It confirms the signicative difference of the 3 coefficients corresponding to the 3 backgrounds ($p<0.05$) and the fact that there is no significative effect with others experimental factors, neither repetition nor direction of the adjustment process.

[5] http://www.zoom.co.jp/english/products/h4n/
[6] http://www.bksv.fr/Products/handheld-instruments/sound-level-meters/sound-level-meters/type-2238.aspx

3- Loudness equalization

The aim of this second preliminary experiment was to realize a loudness equalization of all warnings of the corpus. For that, the linear coefficients (C1, C2, C3) found in the previous step was applied to the reference sound (*white noise*) for each background condition. Then, on the basis of Menzel's and Yamauchi's conclusions [9] [10], an extrapolated offset of +20 dB was applied to these three weighted references in order to make them "clearly audible" in their respective background. From that new clearly-audible references, the nine other signals of the corpus were perceptually equalized in loudness.

The test took place with the same hardware and software environment as the previous one (see section 2 above). It has been performed by 10 persons (4 females, 6 males).

At the beginning of the procedure, the participants were given written instructions presenting the experimental task but, as in step 2, without detailing the overall context. The protocol consisted in presenting sounds by pair. Participants had to adjust the loudness (i.e. the perceived level) of the second sound of the pair with regards to the level of the first one, being the reference sound. The three references defined previously were considered and succesively compared with each of the nine remaining sound of the corpus. Finally, 30 pairs of sounds were performed by each participant. The unique experimental factor (presentation order) was randomized for each participant. As in the previous experiment, the level of the second sound of the pair was adjusted in real time by participants with a potentiometer. This equalization procedure was done without considering any background noise.

For all participants and for each reference, means and standard deviations are computed on the data in order to give an averaged linear correction coefficient for each warning, in each background. These weighting values are applied to obtain 30 resulting equalized sounds (10 warnings x 3 backgrounds).

4- Reaction Time

The aim of the core experiment of this study was to measure the reaction time associated to the detection of the 10 equalized warnings in each of the 3 backgrounds. For that, a pseudo-realistic passing-by scenario was built, on the basis of Kerber's approximation of level course [11] in order to modulate in a congruent way the variation of the level with regards to the vehicle / listener distance. In this protocol, the listener was assumed to be static at a reference point and the vehicle was assumed to pass from left to right with its sound level (at the reference point) being modulated by a distance-dependant attenuation function.

Stimuli

This attenuation function was applied to all the warnings of the corpus, considering, for each of them, the clearly-audible equalized level – ouptuts of step 3 – as the maximum level (i.e. when the car is passing in front of the listener). Besides, at the constant speed considered (20 km/h), Doppler effect have been judged negligeable.

However, for experimental reasons mainly due to test duration and masking issues, the attenuation function has been rendered assymetrical: the left part of Kerber's profile (concerning the span -35 / 0 m.) has been only taken into account and has been extrapolated by a conventional $1/r^2$ decreasing function before -35 m. Thus, the attenuation function effectively applied to the warnings considers a -70 m. / 0 m. course assuming that the second part of the scenario (from 0 to +35 m.) is not very relevant when measuring reaction times that are expected to occur in the first part of the scenario, i.e. before the frontal position (cf. Figure 2). Moreover, because of a persistant side effect due to spectral masking, offsets of "clearly audibility" have been empirically adjusted in order to make all the warnings detectable in an acceptable range: neither too early (i.e. at the very beginning of the sequence) nor too late (i.e. not even at the end of the sequence). The correction values of this adjustment are the following : + 7 dB (residential), + 3dB (shopping), + 14 dB (busy). [Sounds are available at http://pds.ircam.fr/ica2013_sounds.html]

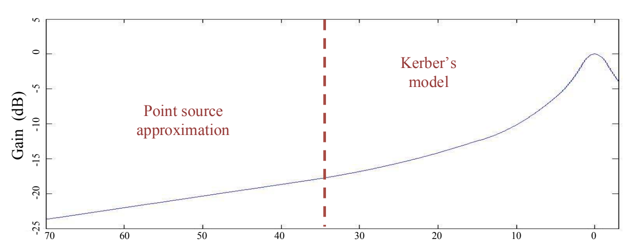

FIGURE 2. Attenuation function (in dB - normalized) applied in the sequence building scenario between -70 and 0 m.

Procedure

Six persons (2 females, 4 males) performed the experiment. The test took place with the same hardware environment as the prvious ones (section 2). It was run with Superlab Pro Edition[7] environment (v4.0), an experiment generator package including dedicated functionnalities and devices for reaction time procedures.

At the beginning of the procedure, participants were given written instructions presenting the task to be performed but, as previously, without detailing the overall context of the study. Participants had to react – by hitting the space bar of the computer keyboard – as soon as the warning inside a mixed sequence was detected. The protocol contained 4 separated blocks lasting about 1-hour each, and performed separately by participants with, at least, a 2-day gap between consecutive blocks. Each block contained 5 trials, and each trial contained the experimental corpus of 30 sequences (explicited above). Then, for each warning and background, 20 reaction time values were individually collected at the end of the test. The sequence order used to playback the 30 sequences in one trial was selected in a random way.

At the end, a questionnaire including test assessment and subjective judgements on the sounds was also filled by each participant.

RESULTS

Warnings / Backgrounds

On the basis of the data collected (600 reaction time values per participant), mean and standard deviation are first computed for each warning and background, i.e. averaging on participants and repetitions (Figure 3). This first step allows to observe noticeable differences on reaction times according either to warning or background types. To sum up and give a rough estimate for each configuration, min./max. values in seconds (and their associate labels) are given hereinafter: 3.74 (*clicking*) / 11.80 (*modulated electric hum*), for residential; 3.66 (*klaxon*) / 11.13 (*electric hum*), for shopping; 4.13 (*burst of white noise*) / 10.53 (*modulated electric hum*), for busy.

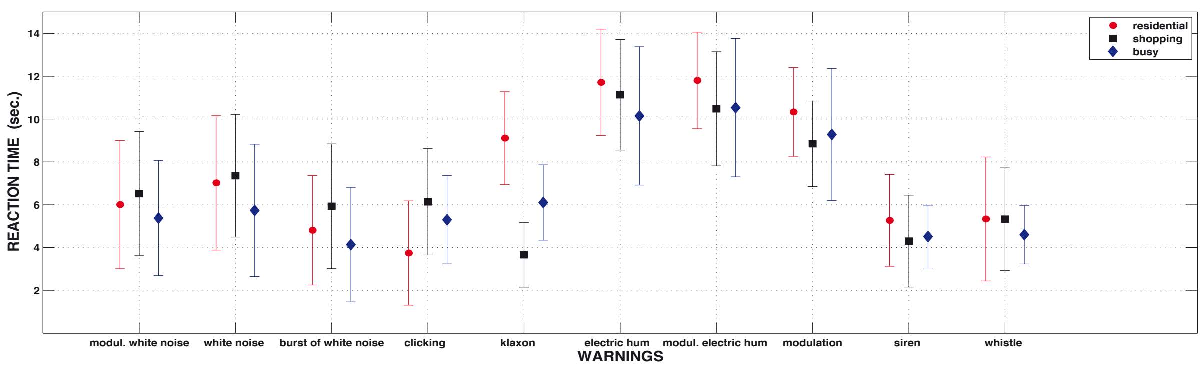

FIGURE 3. mean and standard deviation of reaction times (in sec.) for each warning and each background: residential (red circle), shopping (black square), busy (blue diamond).

To go further, an ANOVA with type of warnings (10), type of backgrounds (3), repetitions (4) as within-subjects factors is performed. A first look at the ANOVA results alllows to consider a significative influence of these single primary factors (warning / background types) (resp., $F(df=9)=370.31$, $p<0.05$; $F(df=2)=44.34$, $p<0.05$) but also a significative influence of the interaction between them ($F(df=18)=23.87$, $p<0.05$). At a first level, these results are coherent with Menzel's conclusions [9] who found, among others, a strong influence of the type of his three warning sounds (horn engine and noise) as well as his four background noises (busy, residential, shopping and heavy-traffic) on the level difference between "just audible" and "clearly audible" thresholds.

Besides, these results allow to rank the warnings with regards to their corresponding reaction times and to define, for each background, best ones in terms of detectability. Moreover, linked to the physical caracterization undertaken previously (section 1), this result can lead to draw regions of *best detectability* in the physical space of the corpus, in other words to determine – again for each background – ranges of parameters where the detectability is promoted (cf. Figure 4).

[7] http://www.superlab.com

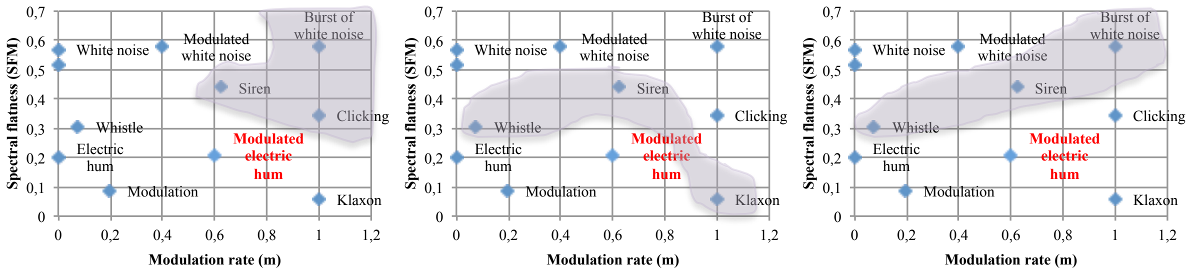

FIGURE 4. Depiction of *best detectability* regions in the physical space of the corpus with regards to reaction time averaged values ranking, for each warning and background: residential (left), shopping (center) and busy (right).

Conversely, even if further investigations and analysis must be done to reach more general conclusions – in case, independent to the background – these intermediate results can nevertheless lead to a first possible estimation of the worst parametric configuration for detectability with regards to acoustic properties. Actually, one can observe on Figure 4 that, whatever the background considered, a *critical zone* on the left/down quarter of the space can be defined. This area correponds to low values of the two physical parameters defined above (cf. Experimental approach / section 1) and is represented by typical warnings like *electric hum*, *modulation* or *modulated electric hum* that are rather harmonic and that contain a quite low temporal variation rate. This result – yet to be consolidated – can constitute a first element of answer to the initial question, by determining which kind of sounds is probably to avoid with regards to Quiet Vehicle's detectability performance.

Learning effect

Another important result pointed out by the ANOVA analysis is the significative influence of the third experimental factor, repetition, ($F(df=3)=80.67$, $p<0.05$) and the significative interaction between this factor and the type of warnings ($F(df=27)=2.27$, $p=0.0002$). This results can also be observed by the evolution of the mean reaction time values along the course of the experiment, i.e. considering separately the four experimental blocks. For the three backgrounds, the reaction times of most of the warnings are significantly decreasing. Figure 5 below depicts this evolution for the *shopping* background where different decreasing profiles can be observed: rapid (*modulated white noise*, *clicking*) or progressive (*modulation*, *whistle*).

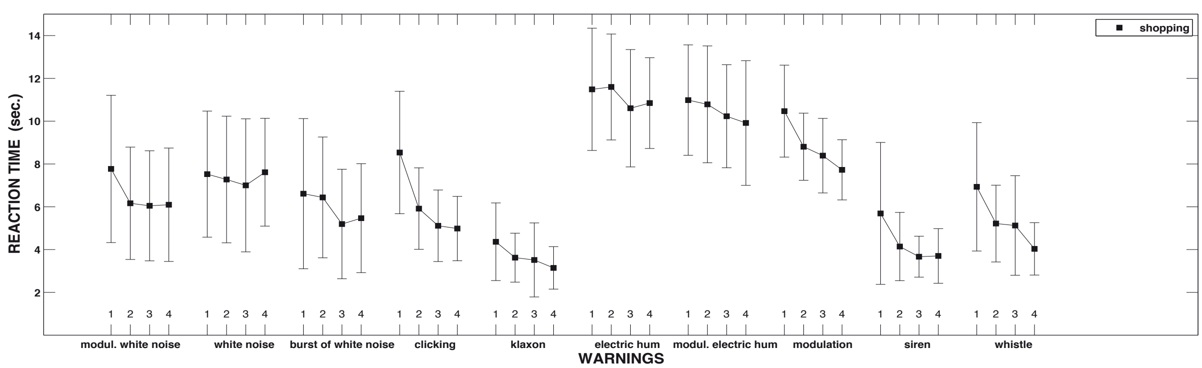

FIGURE 5. Evolution of the mean and standard reaction times along the course of the experiment (blocks 1 to 4), for each warning and for the *shopping* background (black square).

More precisely, a first round of analysis with the decomposed data (average on each block taken separately) related to each warnings on each background can globally lead to the following preliminary observations:

i) leaning effect seems to be significative for almost 75% of all configurations (warning+background).

ii) for some warnings (e.g. *clicking*, *whistle*), learning effect is significative for all backgrounds while for other warnings (e.g. *white noise*, *modulated white noise*) learning effect is only significative for one background.

iii) the significative aspect of learning effect appears mainly from the second block (65% – e.g. *clicking* in *shopping* on Figure 5) and, less frequently, from the third block (25%) or the fourth block (10% – e.g., *klaxon* in *shopping* on Figure 5).

CONCLUSION

In the emerging topic focusing on Quiet Vehicle, questions concerning relevance and nature of extra artificial sounds eventually added to the vehicle are asked. At that time, it undoubtedly seems that a dedicated audio warning is necessary to face at risks involved by this significative lack of presence either for the vicinity, especially the pedestrians, or for the drivers themselves. Then, in its entirety, this issue represents an emblematic application in the interactive sound design – or sonic interaction design – field, as it has to deal either with addressed functions to fulfill and new aesthetics to imagine, by exclusive means of the auditory modality.

Within that framework, a study has been undertaken to first investigate the specific question of detectability. The starting hypothesis was that sound properties of some signals may better signify the presence of a Quiet Vehicle inside a given background noise. A corpus of warning signals (*warnings*) and background noises (*backgrounds*) has been collected/recorded, physically characterized and mixed in 15-sec. sequences, after two preliminary steps of experiment that had led to the definition of "just" then "clearly" audible thresholds for a reference warning, and the loudness equalization of all the other warnings. These thirty 15 second-sequences (10 warnings x 3 backgrounds) were used in a reaction time experiment that measured the time for participants to detect the emergence of a simulated sonified vehicle immersed in a urban sonic environment, within an modelized approaching scenario at constant speed. This experimental protocol also included four different separated blocks of test over several days, allowing to observe the data evolution along time.

The main result shows that it seems to be a significative difference between warnings in terms of reaction time, according to a considered backgound; this leads to possible definition of a critical range in terms of physical parameters – or sound properties – corresponding to potential *bad warnings* for detectability. Concretely, signals that are rather harmonic and continuous seems to be less optimized candidates for efficient extra sounds in Quiet Vehicle, at least when detectability criterion is considered. Moreover, another type of results is highlighed in this study: the fact that reaction times are significantly and predominantly decreasing along the course of the experience (several days), meaning that a learning effect occurred for a large part of the warnings, either in all backgrounds or, at least, in one among three. This tendancy must be further investigated on the hypothesis that they can be associated to different learning mechanisms and strategies.

REFERENCES

1. R. L. Robart, L. D. Rosenblum, "Are hybrid car too quiet ?", J. Acoust. Soc. Am. **125**(4), 2744–2744 (2009).
2. R. M. Schafer, *The soundscape*, Destiny Books, October (1993).
3. U. Sandberg, L. Goubert., P. Mioduszewski, "Are vehicles driven in electric mode so quiet that they need acoustic warning signals", *Proceedings of 20th International Congress on Acoustics*, Sydney, Australia (2010).
4. B. Le Nindre, G. Guyader, "EV- sound quality: customer expectations and fears - crucial NVH stakes", *Proceedings of SIA conference (Société des Ingénieurs de l'Automobile)*, Saint-Ouen, France (2010).
5. QRTV-1, 2010, "Terms of Reference and Rules of Procedure for the GRB informal group on Quiet Road Transport Vehicles (QRTV)", Document QRTV-01-02-e, ECE/WP29/GRB, Geneva, Switzerland (2010).
6. N. C. Otto, R. Simpson, J. Wiederhold, "Electric vehicle sound quality", *Proceedings of the 1999 Noise and Vibration conference,* Vol.1, 361-363 (1999).
7. M. Wogalter, R. Ornan, R. Lim, R. Chipley, "On the risk of quiet vehicles to pedestrians and drivers", *Proceedings of the Human Factors and Ergonomics Society*, 45th annual meeting (2001).
8. P. Nyeste, M. Wogalter, "On adding sound to quiet vehicles", *Proceedings of the Human Factors and Ergonomics Society*, 52nd annual meeting (2008).
9. D. Menzel, K. Yamauchi, F. Völk, H. Fastl, "Psychoacoustic experiments on feasible sound levels of possible warning signals for quiet vehicles", *Proceedings of DAGA'11 conference*, Düsseldorf, Germany (2011).
10. K. Yamauchi, D. Menzel, H. Fastl, M., Takada, K. Nagahata, S. Iwamiya, "Cross-cultural study on feasible sound levels of possible warning sounds for quiet vehicles", *Proceedings of Internoise conference,* Osaka, Japan (2011).
11. S. Kerber, H. Fastl, "Prediction of perceptibility of vehicle exterior noise in background noise", *Proceedings of DAGA'08 conference*, Dresden, Germany (2008).
12. G. Lemaitre, O. Houix, N. Misdariis, P. Susini, "Listener Expertise and Sound Identification Influence the Categorization of Environmental Sounds", Journal of Experimental Psychology: Applied, vol. 16, n° 1, 16-32 (2010).
13. G. Peeters, B. Giordano, P. Susini, N. Misdariis, S. McAdams, "The Timbre Toolbox: Audio descriptors of musical signals", J. Acoust. Soc. Am. **130**(5), 2902-2916 (2011).
14. V. Maffiolo, "De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain", PhD thesis, Université du Mans (1999).
15. B. K. Smith, "PsiExp: an environment for psychoacoustic experimentation using the IRCAM musical workstation", *Proceedings of Society for music perception and cognition conference'95*, Univ. of Berkeley, CA, USA (1995).