

---

# *alarm/will/sound*: Sound design, modelling, perception and composition cross-currents

---

ALEXANDER SIGMAN\* and NICOLAS MISDARIIS\*\*

\*International College of Liberal Arts (iCLA), Yamanashi Gakuin University, 2–7-17 Sakaori Kofu-shi, Yamanashi-ken, 400–0805 Japan,

\*\*Sound Perception and Design Team, STMS Ircam-CNRS-Sorbonne Université, 1, place Igor-Stravinsky, 75004 Paris, France

Emails: lx.sigman@gmail.com; nicolas.misdariis@ircam.fr

**An ongoing international arts-research-industry collaborative project focusing on the design and implementation of innovative car alarm systems, *alarm/will/sound* has a firm theoretical basis in theories of sound perception and classification of Pierre Schaeffer and the acousmatic tradition. In turn, the timbre perception, modelling and design components of this project have had a significant influence on a range of fixed media, electroacoustic and media installation works realised in parallel to the experimental research. An examination of the multiple points of contact and cross-influence between auditory warning research and artistic practice forms the backbone of this article, with an eye towards continued development in both the research and the artistic domains of the project.**

## 1. INTRODUCTION

*alarm/will/sound* is a tripartite collaboration between a composer (Alexander Sigman), a research fellow in the field of sound perception and design (Nicolas Misdariis) and a product designer/visual artist based in Stuttgart and London (Matthias Megyeri). This project began in January 2013 under the auspices of the Ircam-IRC (Interface Recherche/Création)<sup>1</sup> Artistic Research Residency programme, and is still in progress at the time of writing (April 2018).

While the design of most sound-emitting components of the automobile (e.g., the audio system, engine, horn, turn signal, door and even Vehicle Sound for Pedestrians (VSP) in the case of electric vehicles (Misdariis and Cera 2017)) have continuously undergone improvements and refinements in recent decades, the car alarm has changed very little. Unlike other domestic and industrial security systems, audible automotive alarms have proven to be ineffective as deterrents, and are often considered more as an anti-social nuisance than as a benefit to the public.<sup>2</sup>

Taking as a point of departure this proven ineffectiveness and (largely for that reason) the relative lack of research into and development of audible car alarm

design, we have sought to produce innovative modified alarm system prototypes, informed by musical, artistic, scientific and industry expertise, as well as sound perception research and acoustic modelling. In the larger scope, and to some degree, with reference to the work of Patterson and Datta (1999), *alarm/will/sound* can be seen as an attempt to extend the domain of auditory alarms, as it aims at developing an original framework and exploring new paradigms for proposing innovative morphologies and textures for warning signals. In addition, we consider the alarm system as a sort of new interface for musical expression (NIME), triggered by a user's position in space, trajectory and gestures. As such, a set of interaction design models have been proposed, in the interest of morphing the car alarm function from deterrence into engagement. These models are described in Misdariis and Sigman (2016).

In this article, we discuss how the legacy of the timbre perception and classification models attributed to Pierre Schaeffer (Schaeffer 1966/2017) and the acousmatic tradition have shaped our approaches to research in the area of auditory warnings. In turn, examples are provided as to how the sound taxonomy, modelling and geolocalisation dimensions of the project, as well as the results in two experiments on alarm sound causation identifiability and synthetic auditory warning attraction and repulsion levels (respectively) have been adapted to the composition of fixed media, electroacoustic and multimedia installation works. Finally, future artistic research directions for the project are proposed.

## 2. DESCRIPTION OF THE SONIC ENVIRONMENT

The first phase of the project entailed the elaboration and characterisation of a sound catalogue to apply to the modified car alarm prototypes. As is presented in Figure 1, the sound corpus taxonomy consists of three primary categories: Individual Sounds, Auditory Scenes, or sound complexes, and Real Car Alarm Sounds (i.e., the standard repertoire of six auditory warnings typical of audible car alarm systems). The Individual Sound category is divided into 1) Synthetic/Electroacoustic; 2) Vocal; 3) Film Danger Icons; and 4)

<sup>1</sup>[www.ircam.fr/creation/residence-en-recherche-artistique/](http://www.ircam.fr/creation/residence-en-recherche-artistique/)

<sup>2</sup>This nuisance factor is made evident by the numerous attempts to ban car alarms in New York City over the past 25 years or so. See [www.nytimes.com/2004/07/22/nyregion/city-council-bans-car-alarm-sales-but-a-mayoral-veto-looms.html](http://www.nytimes.com/2004/07/22/nyregion/city-council-bans-car-alarm-sales-but-a-mayoral-veto-looms.html) (accessed 7 April 2018).



**Figure 1.** Alarm sound catalogue taxonomy, constructed in January–February 2013. ‘Elec’ and ‘Mech’ (under INDUSTRIAL/MECHANICAL→electrically powered and mechanical sources, respectively). ‘Concat’ (under SYNTHETIC/ELECTROACOUSTIC→Grains) refers to the concatenative synthesis technique by which the entries in this sub-category were generated.

Industrial/Mechanical Sounds. Further subdivisions were made primarily along semantic/contextual lines. One may observe from the figure that the Industrial/Mechanical rubric consists of the greatest number of taxonomical divisions, and is the most populated with sonic entries among the Individual Sound subcategories.

Among the non-synthetic sounds in all three primary categories, the majority were mined from existing sound databases (e.g., SoundIdeas, Blue Box, Auditory Lab and Freesound).<sup>3</sup> Under the Auditory Scenes rubric, a series of field recordings of public spaces in Paris – streets, the Forum Les Halles shopping concourse, the

<sup>3</sup>[www.sound-ideas.com/](http://www.sound-ideas.com/), [www.bestservic.com/sound\\_effects.html](http://www.bestservic.com/sound_effects.html), [www.auditory-lab.com/](http://www.auditory-lab.com/) and <https://freesound.org/>, respectively.

Pompidou Centre, metro stations and train car interiors – were compiled in February 2013 by Alexander Sigman and Matthias Megyeri. It is intended that the collection of field recordings be expanded over time to include further site-specific entries.

### 3. HOW DO WE LISTEN TO OUR SONIC ENVIRONMENT?

#### 3.1. Different listening modes

A sonic environment can be defined, in a trivial way, by all the sounding objects or elements that belong to one’s daily surroundings. Grounded on a more

academic basis, this notion of sonic environment can be linked to the concept of *soundscape* developed by R. M. Schafer in his seminal book *The Tuning of the World* (Schafer 1977). In this publication, Schafer makes a clear distinction between *hi-fi* and *lo-fi* soundscapes, the former obtaining a ‘good’ signal-to-noise ratio (SNR) – that is, high enough to deliver an auditory perspective and make possible the distinction of several sources – and the latter being, on the contrary, a low SNR environment in which all sources seems to be mixed together and covered by background noise. Another way to define – and illustrate – the notion of sonic environment could also borrow from Nelken and de Cheveigné’s definition of a general sound scene, linked to the functioning of auditory system in the brain: ‘a skeleton of events on a bed of texture’ (Nelken and de Cheveigné 2013). This last proposition converges towards the understanding of a sonic environment as a composition of singular sound sources and extended noises perceived more or less easily by human beings.

That being said, it has been well established that the perception of a sound signal coming from any source can be processed in different ways and at different levels of meaning. Initiating the *musique concrète* current, Pierre Schaeffer redefined the notion of listening by formalising four modes, each of them corresponding to a singular perceptual attitude:

- *perceiving* (*ouïr*) means listening to the sound without questioning, like an elementary, raw and passive perceptual action;
- *hearing* (*entendre*) means applying reduced listening, that is, selecting in the sound signal some structural elements useful for the characterisation of the percept;
- *listening* (*écouter*) means looking for cues and indices in the sound that will inform of its possible origin, its source and the nature of the event that has caused it;
- *understanding* (*comprendre*) means arriving at the meaning of the sound by considering it as a sign within an entire language.

A similar terminology was proposed by W. Gaver (1993) with regards to environmental sounds as defined by Vanderveer (‘any possible audible acoustic event which is caused by motions in the ordinary human environment ... Besides 1) having real events as their sources ... 2) [they] are usually more “complex” than laboratory sinusoids ... 3) [they] are meaningful, in the sense that they specify events in the environment ... 4) the sounds to be considered are not part of a communication system, or communication sounds, they are taken in their literal rather than signal or symbolic interpretation’ (Vanderveer 1980)):

- *musical listening* focuses on sound properties;
- *everyday listening* focuses on the source that produced the sound.

The term *musical* is nevertheless possibly ambiguous because, as quoted by Tuuri, Mustonen and Pirhonen (2007), it may be possible to have a musical listening of a non-musical sound and, conversely, to focus on the sources (origins) of a musical signal – for instance, by trying to separate some instruments within the orchestra.

Combining these notions, in his work dealing with image/sound relationships, Michel Chion defines three listening modes, or auditory perception attitudes, which constitute a sort of synthesis of Schaeffer and Gaver’s approaches (Chion 1994):

- *reduced listening*, taken from Schaeffer’s terminology;
- *causal listening*, similar to Gaver’s ‘everyday listening’;
- *semantic listening*, related to the meaning of sound.

Since that time, some scientific studies have aimed at demonstrating – even partially – the mental structure associated with the perception of environmental sounds that has been established more or less empirically by their respective authors. In particular, within the framework of the European collaborative research project CLOSED (Closing the Loop of Sound Evaluation and Design; Susini, Misdariis, Lemaitre, Houix, Rocchesso and Polotti 2006), Lemaitre, Houix, Misdariis and Susini (2010) have collected and studied a sound corpus related to a specific daily environment – the kitchen – and have experimentally discovered two main listening levels, similar to those defined by Chion:

- *acoustic level* – analytic listening of the abstract properties of a sound;
- *causal level* – listening of the sound production phenomena.

Moreover, this study also showed the interaction of two main factors on these listening modes:

- *identifiability* (i.e., the ability of the sound to be identified by the listener) – the less the sound is identifiable, the more the acoustic listening mode is involved;
- *expertise* – the more the listener is inexperienced, the more the causal listening mode is involved.

### 3.2. Warning sounds

In the specific case of warning sounds, these various ways of interpreting the auditory perception mechanisms can be found again through different formal approaches and methodologies for designing auditory warnings. Simply sketched, on the one hand, Patterson (1982) proposed a structure for alarms based on a hierarchical decomposition of the signal into elementary components (from shortest to longest): pulses, bursts and warnings. A pulse is defined as an element

of noise or vowel-like signal of a few milliseconds in length, a burst consists of multiple pulses, and a warning of multiple bursts. In constructing this taxonomy, Gaver attempts (with limited success) to bridge the gap between the atomic level and the semantic level of an auditory warning. Despite its reductive nature, Gaver's formalisation within the framework of auditory icons (Gaver 1986) led to a recommendation to use environmental sounds for warnings (or at least auditory displays) able to be easily understood and associated with a particular object or context. A third way to consider the development of warning sets, in reference to Stanton and Edworthy's (1998) formalisation, rests upon a user-centred approach and the notion of auditory affordances.

This latter notion was borrowed by the authors from Gibson's theory of affordances, tied to the ecological approach that he advocates to explain perceptual mechanisms (Gibson 1979). This theory provides a precise specification of the objects of sensory stimulation that are available for the human organism. According to this theory, perception is active, linked to movement, and picks up the variations and structural stability of a flux of information received by the sensory organs. In contrast to top-down models of perception, the idea of an internal mental representation is alien to Gibson's, as information is already contained in the structure of the sensation (for instance, of light energy or sound timbre). The perceptual mechanisms are thus focused on the perceived object and its intrinsic data (*data-driven*), thereby obviating the need to process information coming from the sensory modalities for building a representation of the world.

To paraphrase Gibson, the notion of affordance can be defined as the potentially useful and relevant information for a person to take action (Gibson 1977).<sup>4</sup> Moreover, Stanton and Edworthy have extended and applied this notion to the perception of auditory warnings, by proposing a 'theory of auditory affordances', and consequently a methodology for designing auditory warnings, that assumes that 'people understand sounds in the world in terms of their *potential for action*' (Stanton and Edworthy 1998: 393–4). For instance, in an Intensive Treatment Unit (ITU) medical context, the sound from a syringe pump should confer the action of replacing the drug (*ibid.*: 392). More precisely, the associated methodology and the corresponding pilot study described by the authors examined and tested several options of sound design, ranging from *nomic* to *symbolic*, including *metaphoric* propositions for new auditory warnings in an ITU setting (for instance, heartbeats for cardiac monitoring, music for an infant warmer device or bubble sounds for syringe pump) (*ibid.*).

<sup>4</sup>Also interesting to note that this notion has been re-contextualised in the ergonomics field, as a characteristic of intuitiveness in the use of an object (Norman 1999).

Schalkwijk (2017) addresses associations between two classes of auditory warnings and their respective vehicle and aviation referents. The two classes of auditory warnings are identified as follows:

- *auditory icons* – iconic representations (or caricatures) of sounds occurring in everyday life situations and, again, transposed in virtual environments to embody objects they contain or actions they suggest (Gaver 1986);
- *abstract melodic alarm signals* (i.e., *earcons*) – arbitrary sounds linked to elementary actions in a virtual environment, and in some cases, using hierarchical rules for creating complex and composite messages (Blattner, Sumikawa and Greenberg 1989).

As such, the efficacy of music-element-based metaphorical alarms is systematically examined to a greater extent than has been the case in previous studies. The issue of prior associations/mental models of subjects for the presented abstract melodic alarm stimuli did emerge when conducting experiments. However, the demonstrated ability to form such associations that are not easily verbalised by subjects suggests their utility in improving non-verbal communication in human–computer interfaces (HCI).<sup>5</sup>

The formal approaches described above fit quite clearly into the scope of the *alarm/will/sound* project, especially by considering the way in which new synthetic auditory warnings may induce human actions/reactions against the sounding object (the vehicle, in this case).

### 3.3. Iconic vs abstract sounds: sound causation identifiability experiment (1)

The main objective of this first perceptual experiment was to determine levels of sound source identifiability among selected items of the sound catalogue (see Section 2 and Figure 1). Given the size and scope of the catalogue, only a subset of the Industrial/Mechanical category was chosen as stimuli. This category was preferentially selected because of its large number of sub-categories and the varying levels of source abstractness relative to the other Individual Sound categories. The expected result of this experiment was to assume the form of an abstractness-iconicity scale constructed across the corpus that would also help to further determine salient semantic and acoustical attributes of the sounds (see Section 4). Such a scale could be formally related to one of the the two

<sup>5</sup>However, this study does *not* investigate which musical parameters best map onto which respective referent features. That is to say, when designing abstract melodic auditory warnings, how to obtain the best results for learnability and differentiability among stimuli (Schalkwijk 2017).



fundamental concepts of *earcons* and *auditory icons*, as defined in Section 3.2.

The experimental framework was an online test in which subjects were asked to listen to each stimulus, provide a brief description of sound causality and indicate a confidence rating of causality identification on a standardised five-point Likert scale (see Figure 2 for a depiction of the experimental user interface).

Figure 3 presents the confidence ratings of the 39 tested stimuli, averaged over the 15 usable participant data sets and listed in increasing order (from low to high). After careful analysis of the (useable) data, it became evident that there were no instances of a high confidence level being correlated with an incorrect or vague sound identification, and vice versa. Given this finding, it was determined that, for the purposes of this experiment, confidence rating was a viable measure of sound iconicity. As such, a significant difference t-test

was performed to compare the response vector associated with each stimulus and those of the highest- and lowest-ranked sounds. This enabled the authors to localise the points along the confidence rating scale at which the difference between the upper-bound and *i*th sound, and symmetrically, the *i*th and lower-bound sound became significant. Fortunately, these points were identical. It was therefore possible to demarcate the threshold between iconic (strongly identifiable) and non-iconic sounds (represented by the thick vertical line in Figure 3).

An analytic view of Figure 3 can lead to the following assumptions: the sounds that caused the most confusion have either contextually obscure sources (wooden gears or milling machine), impulsive features (grenade blast, marbles in vase, or a ball-ricochet) or are steady-state and drone-like in nature, with motor sources (air conditioner, microwave oven, boat motor

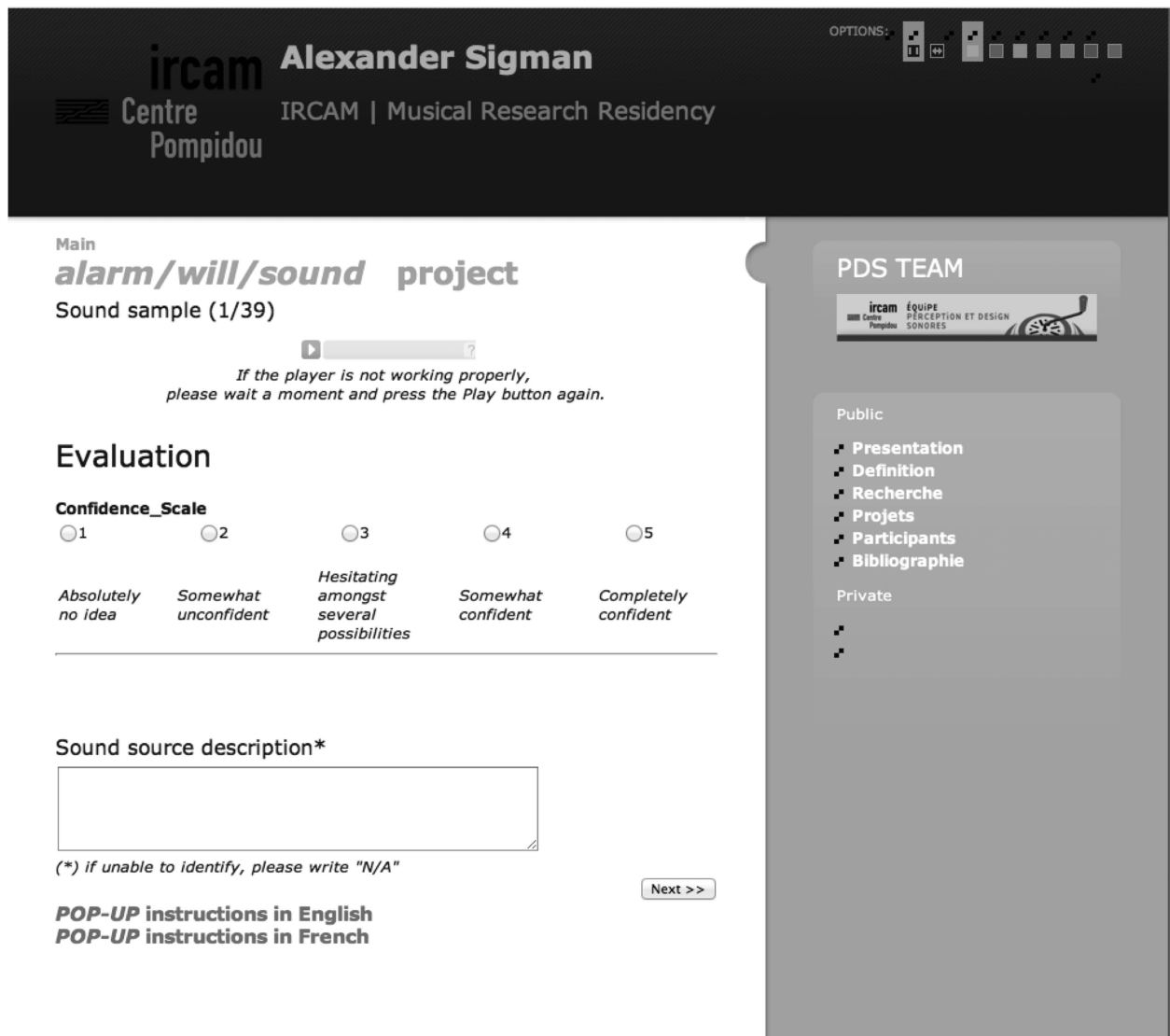
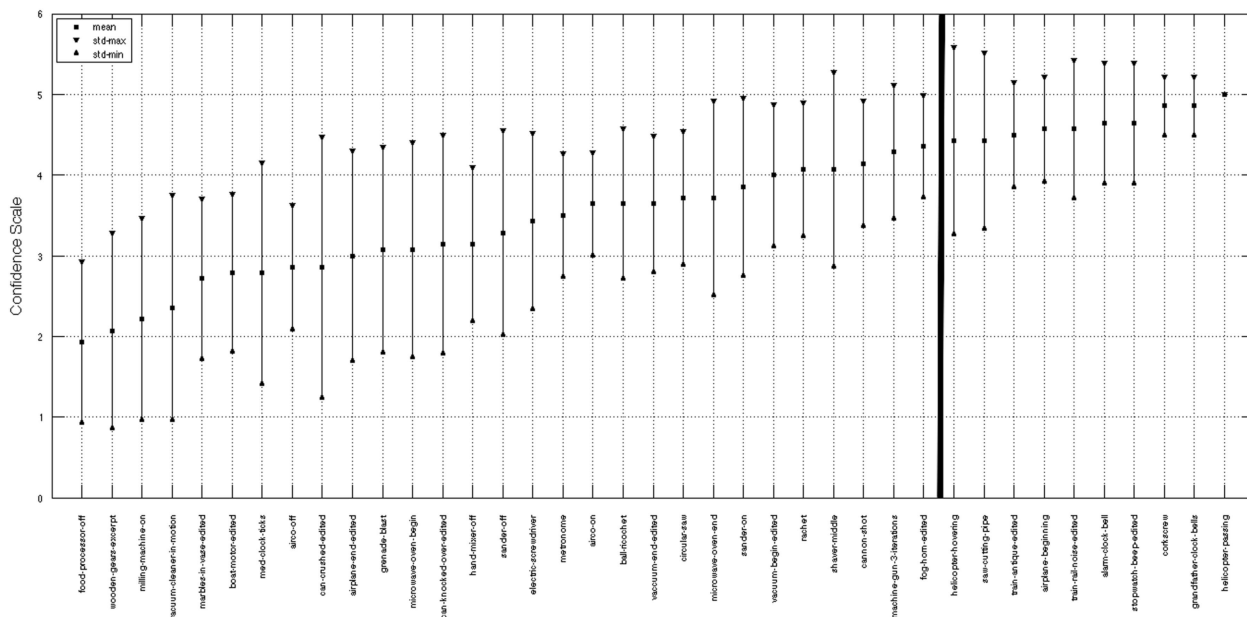


Figure 2. Experiment 1 user interface.



**Figure 3.** Experiment 1 confidence ratings for 39 stimuli. Mean scores, as well as maximum and minimum standard deviations are indicated. The thick black line indicates the iconicity threshold.

or food processor). By contrast, the iconic sounds are quite context-specific (grandfather clock bells, antique train or machine gun), time-varying (e.g., helicopter, airplane or train passing), or common (e.g., corkscrew opening a bottle or alarm clock sounding).

The results of experiment 1 in *alarm/will/sound* project allowed the authors – as expected – to situate the sound stimuli along an abstractness-iconicity gradient and to determine salient semantic characteristics of each sound of the corpus. It was then concluded that sounds falling to the right of the iconicity threshold (see Figure 3) would not fit the auditory warning context, being potentially too closely attached to specific sources that may exist in the vehicle's direct environment (e.g., airplane, train). The definition of this threshold also allowed the authors to define an operational sound corpus for the subsequent phases of the project.

#### 4. SOUND MODELLING AND DESIGN

The concept of sound modelling is closely connected to the development of tools and methods in the signal processing domain that allow users to analyse and describe the content of a sound signal. In some cases, these developments have been based on previous seminal works achieved, since the 1970s, in the domain of characterisation of musical timbre. Among the most cited, Grey (1977), Wessel (1979), Krumhansl (1989) and McAdam, Winsberg, Donnadieu, De Soete and Krimphoff (1995) have studied the subjective attributes underlying the perception of a specific musical sound corpus (synthesised or recorded, percussive or sustained,

etc.) and have tried to find acoustic correlates with some – or a combination of – objective parameters computed from the signal (see McAdams (2013) for an updated and comprehensive review on that question).

More recently, this fundamental approach has been transposed to environmental – that is, non-musical and non-speech – sounds in order to evaluate others types of sounds, upon the hypothesis that the notion of timbre could be applied to extended signals potentially encompassing a large frequency band, or containing very noisy components or a complex array of temporal profiles. Following on from the initial studies on musical timbre (mentioned above), and aiming at a larger 'universality' for a model of environmental sound perception, a meta-analysis was recently undertaken on several studies related to industrial sound stimuli (Misdariis, Minard, Susini, Lemaitre, McAdams and Parizet 2010). The meta-analysis revealed a two-level perceptual structure: on one level, a categorical clustering of sounds, plotted along continuous acoustic dimensions within each category; on the other, an inter-category plotting of sounds with respect to acoustic features shared across categorical boundaries. This confirms the hypothesis of common acoustic features for different kinds of sounds.

##### 4.1. Acoustic description of the sound corpus

On the basis of specific acoustic parameters that had been discovered in the Misdariis et al. (2010) study to be critical to an efficient description of heterogeneous sets of environmental sounds, the sound corpus used in the *alarm/will/sound* project – and more precisely, the subset of 31 Industrial/Mechanical sounds tested in

experiment 1 that fell below the iconicity threshold – has been characterised in a two-dimensional *physical space*, where each dimension, representing a specific acoustic feature, has been demonstrated to be significant in the perception of environmental sounds.

The two features defining this physical space are perceptual spectral centroid (PSC) and harmonic/noise energy ratio (HNR). These parameters were computed in the Timbre Toolbox framework by means of IrcamDescriptor v2.7 routines (Peeters 2004). PSC and HNR are defined as follows (see Misdariis et al. 2010 for details):

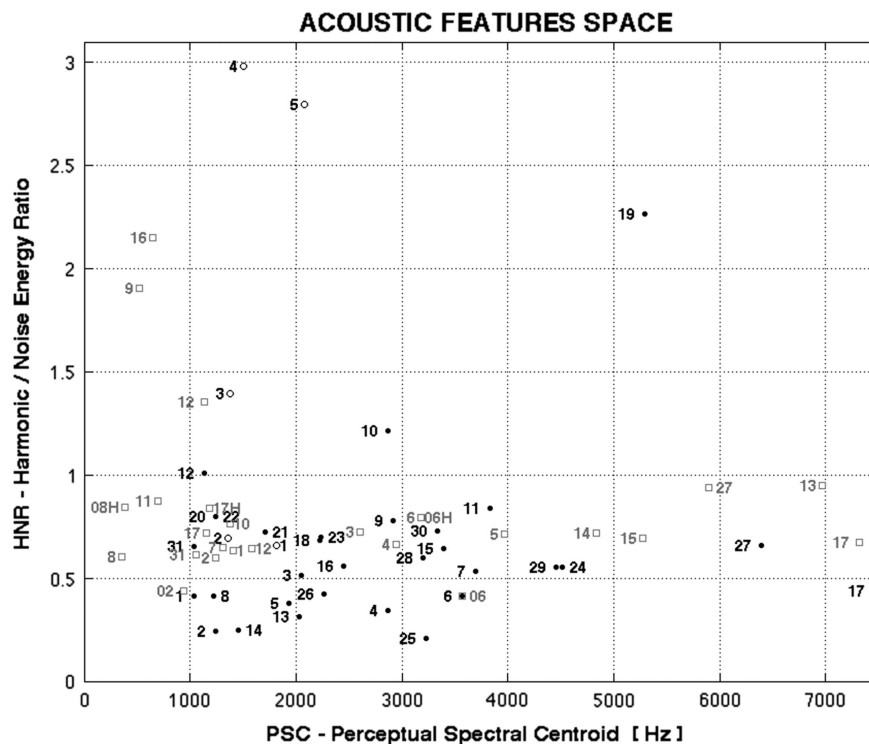
- *PSC (perceptual spectral centroid)* – amplitude-weighted arithmetic frequency average of the signal spectrum, computed on each channel of a gammatone filterbank based on the ERB (equivalent rectangular bandwidth)-rate scale (Slaney 1998).
- *HNR (harmonic/noise energy ratio)* – relative amount of harmonic energy and noise energy contained in the signal. Once both parts (harmonic/noise) of the signal are extracted, the feature simply consists of the ratio of their respective sound levels, that is, energy computed with regards to the ERB-rate scale, as above.

Ideally, a control experiment consisting of a dissimilarity listening test would have been useful to

verify the coherence between the a priori physical space and the actual perceptual space. In other words, this control test would assess the fact that the perceptual dimensions underlying this sound corpus can actually be explained by the chosen physical parameters PSC and HNR. At this stage of the process, the authors assume, on the basis of Misdariis et al.'s meta-analysis conclusions and the comparable nature of the sound stimuli used in the 2010 study, that these features are significantly relevant regarding the perception of the selected sound corpus.

#### 4.2. Populating the physical space: source, hybrid and reference sounds

After this computational stage, the subset of 31 stimuli, obtained by considering the iconicity threshold from experiment 1 on the initial 39-sound corpus, could be situated inside the PSC–HNR physical space, as depicted in Figure 4. In order to fill empty or under-represented areas of this physical space determined by the two acoustic features, hybrid sounds were constructed via applying constant cross-synthesis and LPC amplitude and phase source-filter synthesis algorithms to pairs of one-second samples of extant sounds



**Figure 4.** Acoustic space built upon two perceptually relevant acoustic features taken from (Misdariis et al. 2010): perceptual spectral centroid (PSC) and harmonic/noise energy ratio (HNR). Black numbers beside black dots = source sounds; black numbers beside empty circles = six existing car alarm sounds; grey numbers beside empty squares = hybrid sounds.

in AudioSculpt.<sup>6</sup> Decisions for each sample pair regarding which synthesis algorithm to employ, as well as constant cross- and source-filter analysis settings (e.g., LPC order, in the case of source-filter synthesis) were made based upon analyses of the combined signals, as well as a process of trial-and-error. PSC and HNR values for candidate hybrid out-signals were computed using the methodology described above. Those synthesised sounds whose PSC or HNR values fell roughly half-way between the values of the input signals from which they were generated were retained and integrated into the acoustic features space. Given that the objective of this procedure was merely to populate gaps in the descriptor space across the PSC range, and within the band of HNR  $\leq 1.5$  in which the vast majority of source sounds fell, it was neither necessary nor productive to apply stricter hybrid selection criteria.

Finally, for sake of homogeneity and curiosity, a subset of real car alarms collected during the initial step of constitution of the sound catalogue were also added to the physical space so as to be able to compare existing textures (timbres) used in current auditory warnings with the textures proposed and realised with the scope of the *alarm/will/sound* project. This comparison can be considered as the first step towards the innovative design of car alarms targeted by the project, and can be used as a first informal representation of the authors' ability to extend the current domain of auditory warnings.

### 4.3. From source/hybrid sounds (timbre) to auditory warnings (alarms)

After having preferentially focused in the previous step on the notion of texture (see above) via the timbre characterisation of a selected set of sounds within the initial corpus, the work was then continued by considering more precisely the temporal shape of the new warnings. This approach was centred around attending to two high-level sound parameters that could be considered *musical*, to a certain extent: 1) the temporal morphology of the energy and 2) the rhythm, in the form of a very specific attribute for alarms, the inter-onset interval (IOI). Finally, the combination of source texture, energy profile and rhythm/rate lead to the construction of an original corpus of new types of alarms which were evaluated and assessed in a second sound perception experiment (experiment 2: see below).

#### 4.3.1 Temporal morphologies of environmental sounds

Regarding the question of temporal morphologies, the authors' point of departure was Peeters and

Deruty's study of an (automatic) morphological description of sounds (Peeters and Deruty 2008), that has been thereafter perceptually reinforced and augmented thanks to an original protocol based on subjects' graphical production on a WACOM Intuos tablet (Minard, Misdariis, Houix and Susini 2010). In the latter study, 55 sound stimuli of varying duration were presented: 12 sounds ranged between 350 ms. – 1 second, 18 were between 1 and 3 seconds, and 25 between 3 and 5.15 seconds. These two studies formalise a taxonomy of six perceptually distinct environmental sound morphologies: stable, decreasing, increasing, single impulse, pulse-train and rolling, whose symbols – extracted from an 'average profile' of all subjects' drawings in the Minard et al. study – are depicted in Figure 5. In the context of *alarm/will/sound*, these six specific energy envelopes have then been modelled as BPFs (break-point functions) and have been used as amplitude modulation envelopes to create 186 ( $6 \times 31$ ) specimens of alarms combining different textures (timbre) and temporal profiles.

#### 4.3.2. Existing car alarm inter-onset intervals (IOIs)

The inter-onset interval (IOI) feature has been theorised, among others, by Patterson in the context of proposing guidelines for the creation of auditory warnings. In so doing, Patterson defines a standard structure of an alarm: a series of bursts made of a series of noise pulses (Patterson 1982). The time between the bursts in the series is identified as the inter-onset interval, a leading attribute of an auditory warning, responsible for creating a sense of perceived urgency (Sueid, Susini and McAdams 2008).

In order to explore this fundamental feature, the inter-onset intervals of six standard car alarms, derived from the Car Alarm Prototypes category of the sound corpus depicted in Figure 1, and situated in the acoustic features space (Figure 4), were calculated in

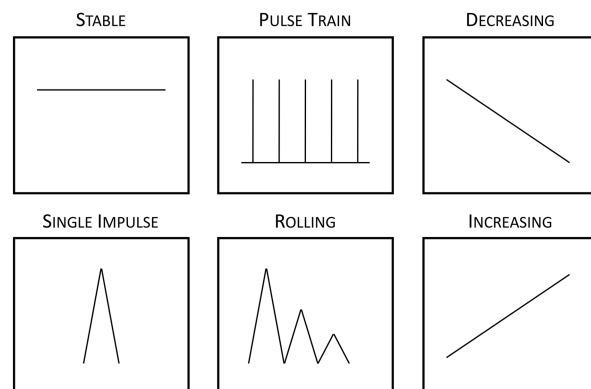
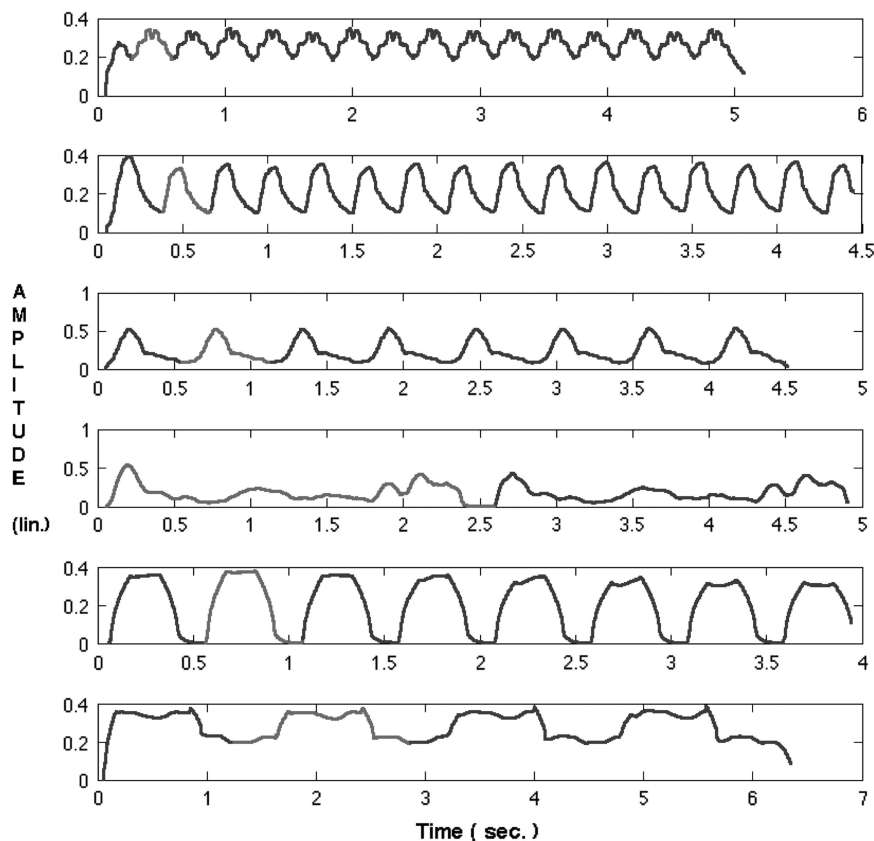


Figure 5. Taxonomy of temporal morphologies of environmental sounds (from Minard et al. 2010).

<sup>6</sup><http://anasyntn.ircam.fr/home/english/software/audiosculpt>





**Figure 6.** Six car alarm signals (1 to 6 from top to bottom). For each, the grey segment of the curve is the elementary component used for computing the IOI of the whole auditory warning signal.

the Matlab<sup>7</sup> environment. For each of them, an elementary component – the ‘burst’ in Patterson’s formulation – was manually isolated and used to detect the periodicity of the alarm that directly provides the IOI value (see Figure 6 for a depiction of this principle). Given the similarity in IOI of alarms 1 and 2, 3 and 5, and 4 and 6 (respectively) to each other, each pair of IOIs was averaged, producing three distinct IOI mean durations: 300 ms (named ‘short’), 536 ms (named ‘medium’) and 2,082 ms (named ‘long’). These three different IOI values were used to create 558 ( $3 \times 6 \times 31$ ) specimens of alarms combined with the previous combination set of different textures (timbre) and temporal profiles.

## 5. REPULSION-INDIFFERENCE-ATTRACTION TO SYNTHETIC AUDITORY WARNINGS (EXPERIMENT 2)

The objective of this experiment was to determine levels of repulsion, indifference and attraction to a set of synthetic auditory warnings, constructed from original and hybrid (cross-synthesised) sounds situated in the acoustic feature timbre space (Spectral Centroid/

Harmonic-to-Noise Ratio), deploying typical temporal morphologies of environmental sounds and IOI values of real car alarm sounds, as described above. This parametric approach would thereby make it possible to distinguish between synthetic stimuli – and signal parameters – that would be more suitable for serving a deterrence function, and those that would better lend themselves to providing engaging auditory feedback (e.g., in the context of exploring an alarm system interface).

In the interest of limiting the required completion time for the experiment to 15–20 minutes, 45 synthetic auditory warnings, rather than all possible permutations of sound sources, temporal morphologies and IOIs described above were employed as stimuli. These stimuli were selected on the basis of contrast of spectral and temporal characteristics. Ultimately, five sound sources – three original sounds and two hybrids – were chosen as follows (indices correspond to position in the timbre space depicted in Figure 4): 1) air conditioner being switched on (02); 2) hybrid between a toppled metal can and low frequency components of an electric screwdriver sample (06H); 3) hybrid between a microwave oven being activated and a foghorn (08H); 4) fog horn (12); and 5) stationary part of a shaver sample, between onset and offset (27). These five samples were

<sup>7</sup>www.mathworks.com

combined with three temporal morphologies (increasing, decreasing and single impulse) and the three inter-onset intervals.

Subjects were directed to the listening test URL,<sup>8</sup> and asked to read a set of instructions (in English or French) prior to beginning the trials. In the instructions, it was recommended that subjects, when categorising their reaction to each sound stimulus into one of three types ('repulsion', 'indifference' or 'attraction'), imagine hearing the presented stimuli in a realistic context/interactive situation. The concepts of repulsion/indifference/attraction to a given sound stimulus were illustrated via examples from the visual and olfactory domains (e.g., a blinding light or unbearable odour vs an enticing or intriguing light or odour). Additionally, it was explained that a 'repulsion' reaction would induce the subject to flee the source of the stimulus, while an 'attraction' rating would be predicated by approaching/expressing interest in the sound source. A sound with an 'indifference' rating would not have any major impact on one's current behaviour. Subjects were advised that it was not necessary to listen to each sound in its entirety prior to providing a response.

The 45 stimuli were presented to the subjects in random order. Playback was triggered via an embedded audio player (in a manner identical to experiment 1). Stimuli could be repeated any number of times, and stopped at any point. Advancing to the next trial was only possible after selecting one of the three reaction categories for the current stimulus. For each trial, an optional comments field was provided. In order to avoid auditory fatigue, the experiment was divided into three 15-trial stages.

A detailed report and discussion of the results obtained from experiment 2 are documented in Misdariis and Sigman (2016: 185–94). For the sake of brevity, a summary of the salient findings is provided here. In terms of experimental design, this second experiment measured a deterrence judgement that could be assigned three nominal modalities ('repulsion', 'indifference' and 'attraction'). The variation of this dependent variable was examined with respect to three independent controlled factors that could each be assigned to different discrete modalities: the nature of sound source (texture), the signal morphology and the duration of the IOI that structures the auditory warning. The raw data collected was arranged in an  $N \times m$  matrix,  $N$  being the number of sounds ( $N=45$ ) and  $m$  being the number of modalities (settings) of the measured variable ( $m=3$ ). This matrix was filled with the number of occurrences added up among all the participants, for each cell Sound/Variable. The sum of each line of this matrix corresponds then to the total number

of participants. Moreover, from this initial data set, three aggregated matrices were also compiled with respect to the three experimental factors: a five-by-three matrix gathering the occurrences of repulsion/indifference/attraction for the five sound sources, and similarly, two three-by-three matrices related to the two other factors (morphology and IOI duration, respectively).

If one examines the test results by experimental factors (source, morphology and IOI), it becomes clear from Figures 7–9 that while IOI duration seemed to have a direct impact on attraction/indifference/repulsion rating, the influences of morphology and sound source differences are less clear. With regards to morphology (Figure 8), the single impulse envelope correlated with slightly lower repulsion ratings than the increasing and decreasing envelopes. As for IOI duration (Figure 9), medium and short IOIs were associated with higher repulsion levels than long IOIs.

A deeper numerical analysis was made by means of a  $\chi^2$  statistical test. First, a general test of this kind made on the aggregated matrices of data conditioned

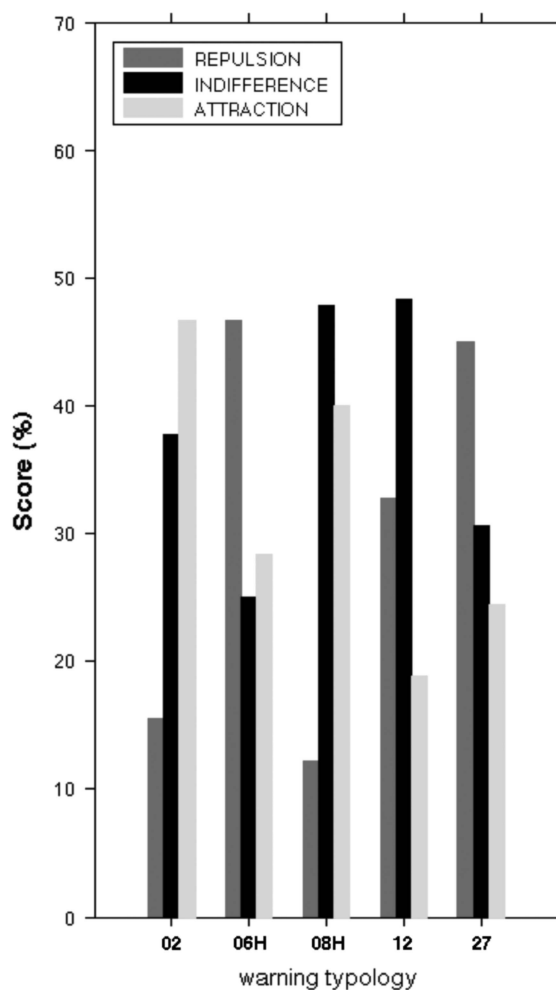
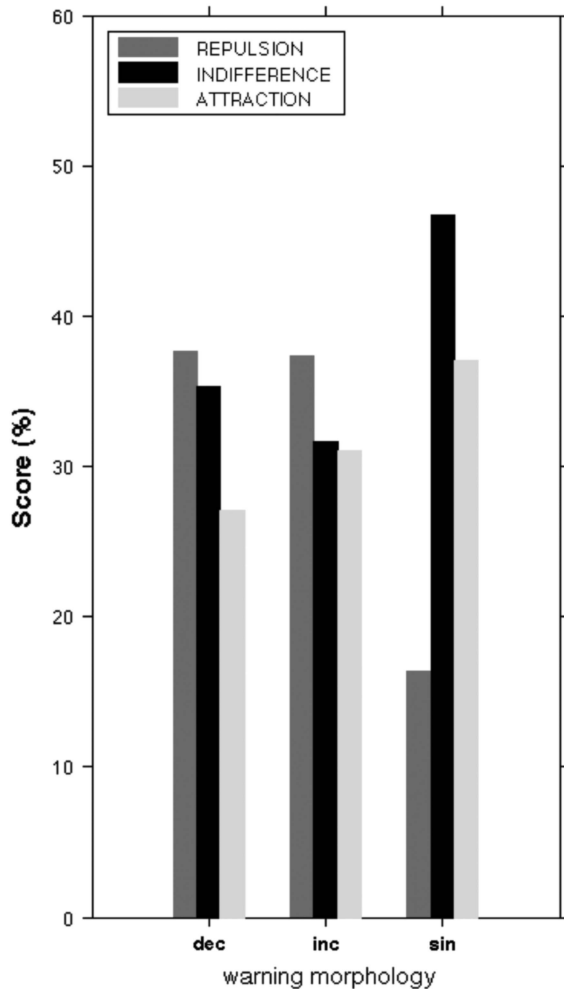


Figure 7. Experiment 2: repulsion/indifference/attraction scores (%) with respect to the five 'source' categories.

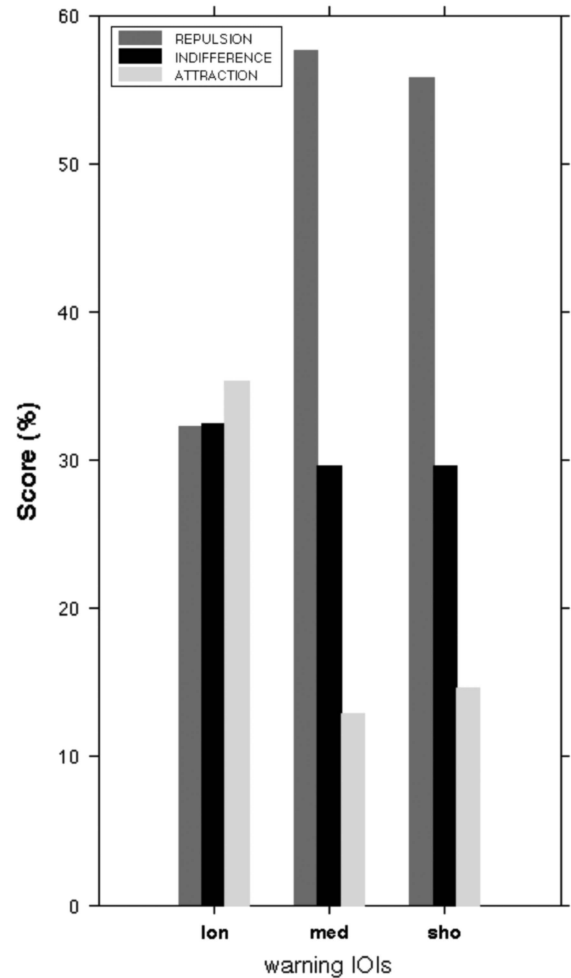
<sup>8</sup><http://recherche.ircam.fr/equipes/pds/projects/asigman/deterrence/src/EvaluationStart.php> (accessed 14 September 2018).



**Figure 8.** Experiment 2: repulsion/indifference/attraction scores with respect to three 'morphology' categories (decreasing ['dec'], increasing ['inc'] and single impulse ['sin']).

as mentioned above, allowed the rejection of a null hypothesis of independence between the measured variable (deterrence) and the experimental factors: source, morphology and IOI duration. In fact, for every configuration, the computed  $\chi^2$  value was significantly higher than the critical  $\chi^2$  value given by Pearson's Law<sup>9</sup> for a standard 5 per cent margin of error (and even for the lowest 0.5 per cent risk rate). A second stage of analysis, using a one-way test, allowed us to examine more precisely this same null hypothesis for each vector of observations, that is, each modality of the measured variable. The computed  $\chi^2$  values were once again higher than the critical one given by Pearson's Law, with the exception of two conditions corresponding to the conservation of the null hypothesis ( $H_0$ ): no significant relationships between either *Morphology* and

<sup>9</sup><http://sites.stat.psu.edu/~mga/401/tables/Chi-square-table.pdf>



**Figure 9.** Experiment 2: repulsion/indifference/attraction scores with respect to IOI durations (long ['lon'], medium ['med'] and short ['sho']).

*Attraction or IOI duration and Indifference* assessment could be inferred. (see Table 1 for exact values.)

As is illustrated in Figure 10, it became apparent that too few stimuli received sufficiently high attraction scores to determine a threshold between indifference and attraction. However, the repulsion levels and quantities of stimuli associated with elevated repulsion levels, combined with the verbal reporting of subjects, will enable the authors to construct a repulsion gradient (spanning 'neutral' to 'extremely repulsive'), and situate the stimuli along this gradient accordingly. While based on the data obtained it would be a speculative exercise to attempt to segment the synthetic auditory warnings into 'attractive' and 'repulsive' categories, it would be possible to filter and select the best deterrence candidates. On a deeper level, this study has confirmed the salient role that IOI duration plays in auditory warning design, and has enabled the authors to closely examine the relationships between the fine structure of auditory warnings and the repulsion effects that they produce.

**6. ALARMIWILLISOUND ARTISTIC ADAPTATION**

Stemming from the research undertaken over the course of the *alarm/will/sound* project, Alexander Sigmán has been creating works for ensemble, electronics, improvisation and media based upon the alarm prototype sound catalogue, as well as concepts pertaining

**Table 1.** Results of the Chi<sup>2</sup> statistical test: probability of the Chi<sup>2</sup> (p), Chi<sup>2</sup> value (χ<sup>2</sup>), degree of freedom (DoF) for the whole dataset (ALL) and the factorised data

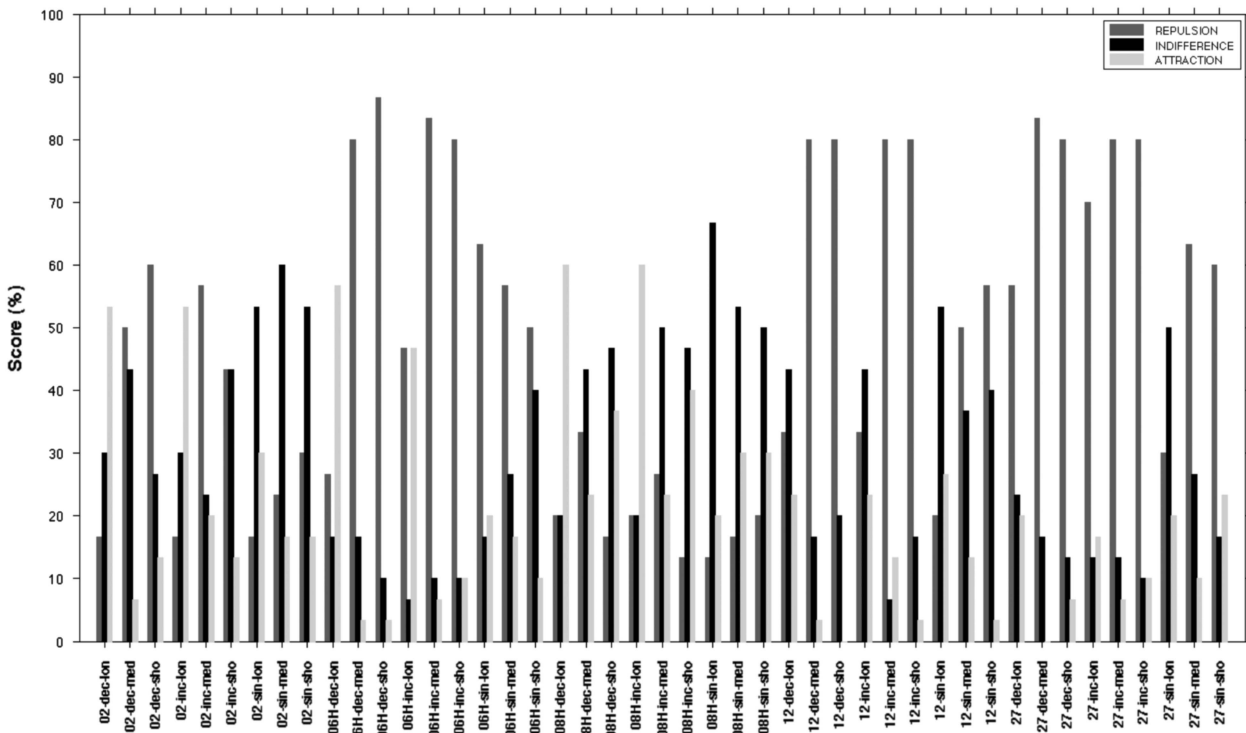
		p	χ <sup>2</sup>	DoF
ALL	Repulsion	< 10 <sup>-4</sup>	142.82 *	44
	Indifference	0.1741	52.65	
	Attraction	0.0011	78.31 *	
Source	Repulsion	< 10 <sup>-4</sup>	61.14 *	4
	Indifference	0.0004	20.27 *	
	Attraction	< 10 <sup>-4</sup>	29.61 *	
Morphology	Repulsion	< 10 <sup>-4</sup>	29.43 *	2
	Indifference	0.0079	9.68 *	
	Attraction	0.0907	4.80	
IOI duration	Repulsion	0.0003	16.21 *	2
	Indifference	0.8216	0.39	
	Attraction	<b>0.0001</b>	<b>18.08 *</b>	

\*statistically significant: H0 rejected.  
Source: Morphology and IOI duration.

to and data obtained from the two sound perception experiments associated with this project. Within the framework of the IRCAM Artistic Research Residency, it is intended that the artist-collaborator not only engage in research, but also apply this research in meaningful ways to their creative practice. The examples below illustrate a few such applications of *alarm/will/sound* concepts, materials and experimental research results to improvised, fixed media, electro-acoustic and interactive installation works realised over the past five years.

**6.1. detritus/reconstructions: interactive electronics improvisation**

The structured laptop improvisation *detritus/reconstructions* (2013) incorporates, analyses, processes and modulates samples derived from the Paris field recordings falling under the Recorded Auditory Scene sub-category of the sound catalogue (see Figure 1), as well as the Industrial/Mechanical Individual Sound sub-category. These processed samples are presented at varying levels of recognisability – ranging from iconic to thoroughly glitched. Using the insights and abstractness-iconicity gradient obtained from the experiment 1 results as a reference, frequency, speed and direction shifting, amplitude modulation and delay/feedback are applied to the selected source-sounds at varying levels, such that sound causation



**Figure 10.** Experiment 2: repulsion/indifference/attraction scores (%) for all 45 stimuli. Each stimulus is indicated along the x-axis as a ‘source-morphology-IOI duration’ triplet, using the abbreviation conventions applied to Figures 7–9.



(and context) identifiability is distorted and ultimately minimised. Through the layering of transformed field recording samples, synthetic auditory scenes are constructed, and identifiability of individual sources is further diminished. It is intended that, as the Auditory Scene taxonomy expands, samples will be selected on the basis of performance location.

## 6.2. *Future Creatures*: fixed media audiovisual work

A fixed audiovisual media work created using the Python-based concatenative synthesis program AudioGuide,<sup>10</sup> *Future Creatures* (2013) was produced in collaboration with Korean animation artist Eunjung Hwang (Movie example 1; Sound example 1). The sound segments and targets to which segments were mapped in the concatenative synthesis process were derived in part from the Industrial/Mechanical, Vocal Individual Sound (Human and Animal) and Car Alarm Prototype categories of the aforementioned catalogue. In so doing, both the (granulated) surfaces and (spectro-temporal) structure of the source sounds become the core material of the audio. These sound categories were selected in response to the characters, objects and events represented in the animation.

Given the garish nature of the colour palette of the visuals, and lack of ‘glitch’ elements therein, Industrial/Mechanical and Car Alarm Prototype sounds situated in the acoustic features space depicted in Figure 4 with relatively high PSC and HNR values were chosen.<sup>11</sup> The decisions as to the segment-size of the Industrial/Mechanical source sounds were determined based upon their respective positions along the aforementioned abstractness-iconicity gradient. That is to say, the more identifiable the source-sound, the smaller the segment-size and the larger the segment density. Echoing the frenetic and densely layered nature of the visuals,<sup>12</sup> these surface fragments and sound morphologies are situated in polyphonic textures, thereby creating synthetic auditory scenes, in which the sources stand in complex and unstable relationships with each other. The auditory events alternately reinforce and conflict with the sense of place conveyed by the visuals.

By dint of the sheer event-density of both the visual and the auditory domains, which saturate the senses, changes in the audio tend to trigger attentional shifts in the visual field, and vice versa. It is these cross-modal relationships on the part of the spectator that the composer wished to explore, as opposed to establishing a fixed translation between visual and audio elements.

<sup>10</sup>[www.benhackbarth.com/audioGuide/](http://www.benhackbarth.com/audioGuide/)

<sup>11</sup>Given the high level of contrast and relatively small number of entries currently populating the Vocal Individual Sound category, all of these sounds were utilised.

<sup>12</sup>The video may be viewed at <https://vimeo.com/70344676> (accessed 7 April 2018).

Thus far, no systematic study of the attraction/repulsion levels and internal properties of the audio elements that cause the greatest attentional shifts among viewers of the video piece has been conducted. However, as a supplement to experiment 2, such an investigation is planned for the near future. This type of empirical study would influence both further research into synthetic auditory warning attraction vs repulsion, and subsequent creative works in which sound sources of a similar nature to those found in *Future Creatures* are employed.

## 6.3. *fcremap* and *atrocitiy exhibition*: ensemble + live electronics electroacoustic compositions

The extraction and application of spectro-temporal features of synthetic alarm source-sounds in the creation of *Future Creatures* was taken a step further in *fcremap* (2014–16), a series of works for instruments, live electronics and visual media, based upon the video and audio of *Future Creatures*. With respect to raw material in the electronics, entries in the original sound catalogue and the hybrid timbres populating the acoustic feature space described above (and employed in experiment 2) were incorporated. Throughout the cycle, the timing and direction of physical actions of performers, which trigger routines in the electronics, are influenced by the morphologies and IOI durations of the synthetic alarms used as stimuli in experiment 2. These morphologies and IOIs were thus treated as abstractions, transposed onto a diverse array of instruments (e.g., various percussion, piano and mixed chamber ensemble), at multiple time-scales and degrees of polyphony.

In the cases of the pieces *fcremapercremaflvln* (for alto flute, violin, electronics and video)<sup>13</sup> and *fcremap'dB* (for ensemble of quasi-variable instrumentation, animated score and electronics, see Movie example 2, Sound examples 2 and 3),<sup>14</sup> these abstractions were employed as filters and linear (visual) elements imposed upon the still images comprising the video projection (in the former case) and the animated score (in the latter).

In *fcrcremap* (for improvising solo clarinetist and electronics) and *fcremap'dB*, the performer is presented with a graphic score, consisting of a series of modified images derived from *Future Creatures*, and a set of guidelines as to how to interpret the graphics on their respective instruments. The electronics thus function as a kind of ‘aural score’, which is to be reconciled with visual data. The application of six alarm IOIs to the electronics effectively creates a sequence of six temporal grids, which the performer

<sup>13</sup>The video projections and performance may be viewed/heard at <https://vimeo.com/167632597> (accessed 7 April 2018).

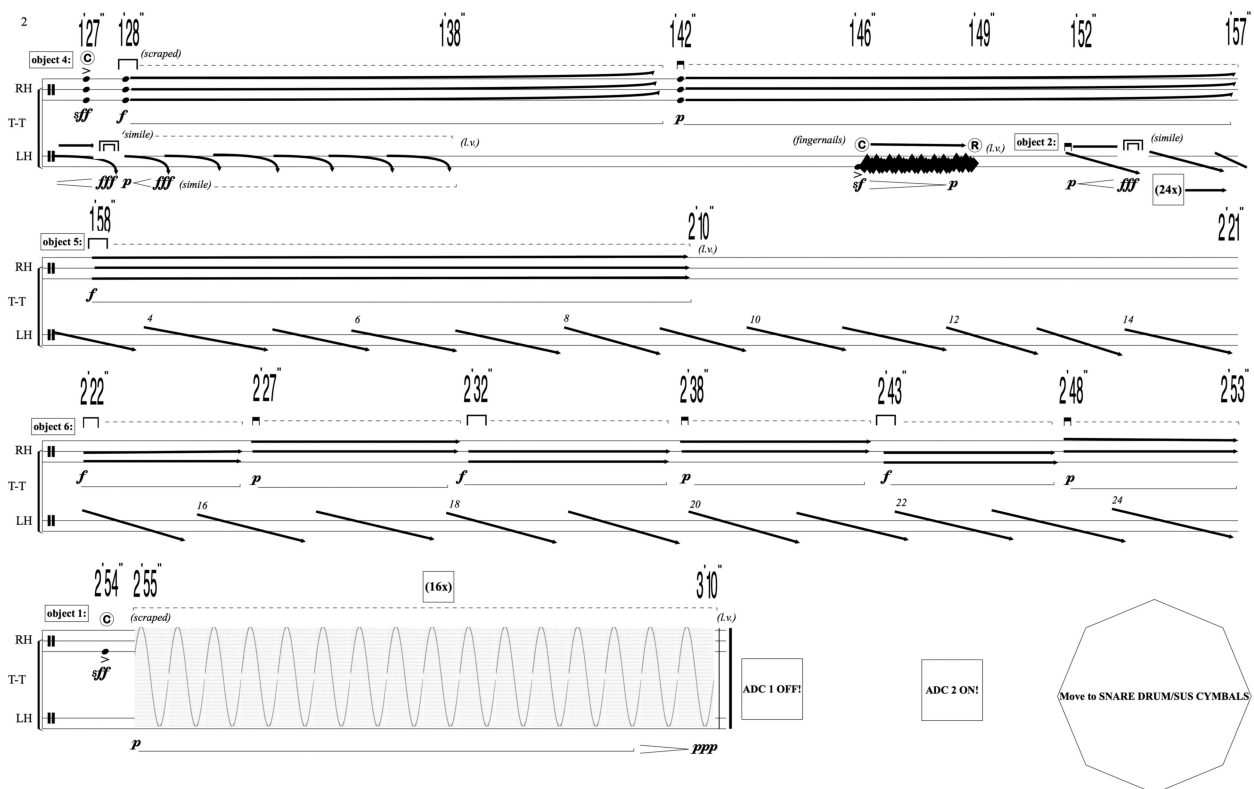
<sup>14</sup>Live performance, with projection of portions of animated score: <https://vimeo.com/197745730> (accessed 7 April 2018).

may resist, use as temporal reference points or (try to) ignore entirely. For *fcmap*, the clarinetist is also asked to approximate the timbral qualities of the audio segments. The surface and structural features of the alarms, coupled with relatively stable image-to-sound mappings between *Future Creatures* video elements and the instrumental and electronics material of the *fcmap* pieces, form a connective tissue for the composer and listener, binding the constituent pieces of the cycle together, despite contrasting performance scenarios and aesthetic characteristics (Figure 11 and 12).

Scored for horn, electric violin, electric guitar, bass guitar, prepared piano and live electronics, the basic material elements in the electronics and the instrumental parts of *atrocities exhibition* (2013) stem from Industrial/Mechanical, Animal Call (under the Vocal Individual Sound category) and Car Alarm Prototype entries in the sound corpus (Sound example 4). Both the sound morphologies and their respective spectral contents were employed to generate the rhythmic and harmonic structures, as well as playing techniques in the instruments, and the spectral content of the electronics. In the case of the piano, the sources of a few of the industrial/mechanical sounds (ratchet, glass shards, metal chain, metal rod and plastic plectrum) – each of which represents a unique physical material or mode of sound

production – are utilised on the piano strings. Furthermore, the electronics are projected through a pair of small transducers placed inside the piano (see Figures 13 and 14). As such, the boundaries between the production and projection of instrumental and electronic sound become blurred. This grey area between mechanical instrument and transducer activation also extends to the violin, electric guitar and bass guitar, whose output is distorted via effects pedals and amplified.

The range of influences and representations of the sound catalogue entries on the musical materials – as sources analysed for spectral and temporal content, as processed samples in the electronics and as models of physical causalities in the instruments – engenders an abstraction-iconicity gradient, along which these materials are situated. This investigation of levels of abstraction, iconicity and sound causality identification reflects the objectives of experiment 1 (outlined in Section 3.3). In constructing our sound catalogue taxonomy, the perceptual issue of listening modes (acoustic/causal/semantic) arose. Through presenting myriad manifestations across multiple musical dimensions of the original sound sources, and blending acoustic and electronic modes of sonic transmission, it was intended in composing *atrocities exhibition* to explore and reveal the complexities behind sound



**Figure 11.** An excerpt of the first movement of the *fcmapperc* score, scored for tam-tam, resonating metal sheet and live electronics. The pitch contours and IOIs of synthetic alarms utilised in experiment 2 are transposed onto the right-hand and left-hand actions of the performer.

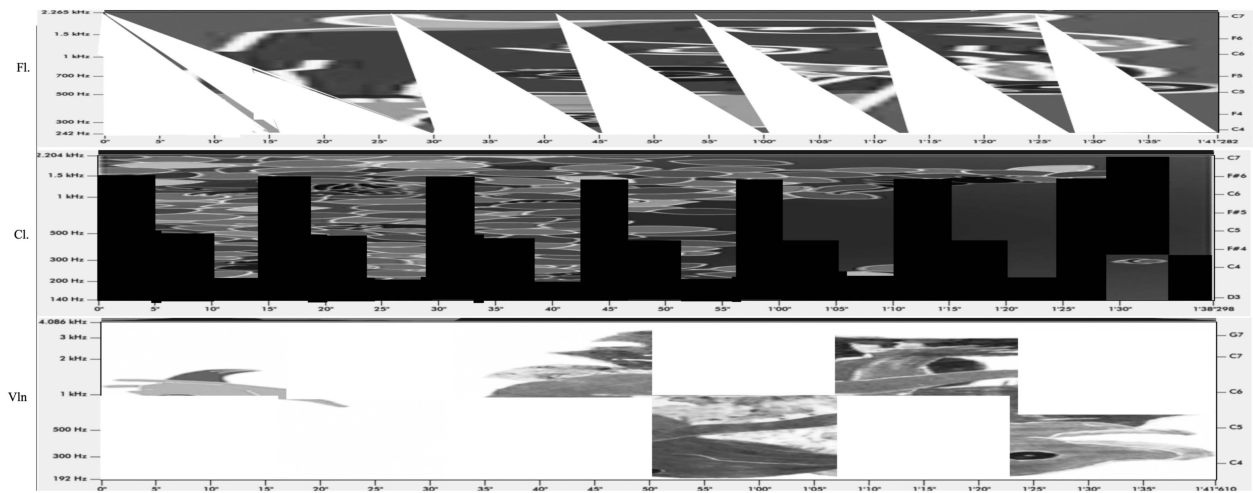


Figure 12. An excerpt of the *fcremap*'dB score, with synthetic alarm temporal patterns and pitch contours inscribed.



Figure 13. Small transducer (back) utilised in *atrocities exhibition* and *Minotaur*.

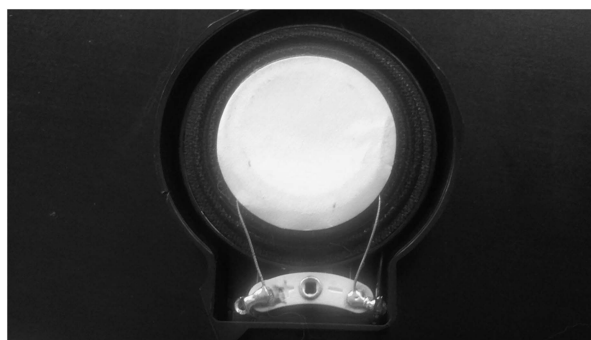


Figure 14. Small transducer (front).

source perception and identification, and the preservation, loss or transformation of sonic semantic associations.

#### 6.4. *Minotaur*: interactive media installation

An ongoing audio–visual–kinetic performance–installation collaboration *Minotaur* (2012–current)

brings together music/sound design, video, dance, sculpture, live painting and set design. With respect to *alarm/will/sound*, it transports the sound categories and features discussed above into the gestural domain, serving as a testing ground for a set of interaction design models that may be incorporated into alarm system prototypes emerging from our research and development activities.<sup>15</sup> Based upon twentieth-century re-castings of the myth of the Minotaur,<sup>16</sup> the locus of activity is a labyrinth constructed out of Plexiglas panels. Inside the labyrinth, dancers' movements are detected via microphones and sensors, which are processed in the electronics. In turn, audio output is transmitted to an array of transducers (similar to the model used in *atrocities exhibition*), affixed to the Plexiglas as well as to hollow sheet-metal sculptures at the centre of the labyrinth (see Figure 15). Consequently, the sculptures, which symbolise the primal desires and tendencies of the Minotaur, function as resonators, and the audio produces Plexiglas vibrations, which influence the dancers' actions. A feedback loop between sound generated inside the labyrinth and audio produced external to it is thus set in motion. Adaptable for gallery, concert and even telematic performance contexts,<sup>17</sup> sounds recorded directly outside of the venue during the performance/exhibition may be integrated into the audio, thereby localising the constructed labyrinth environment.

Audio events may be self-triggered at random intervals, manually activated by the composer or triggered in response to activity within the labyrinth. In all

<sup>15</sup>For a discussion of these proposed gesture-sound interaction design models, see Misdariis and Sigman 2016: 195–6.

<sup>16</sup>Borges's 'The House of Asterion', Swiss author Friedrich Dürrenmatt's *Minotaurus* and Russian novelist Victor Pelevin's *Helmet of Horror* in particular.

<sup>17</sup>Here is an example of a networked rendition of *Minotaur*, involving performers in New York and Oakland, CA: <https://vimeo.com/46066079> (accessed 15 August 2018).



**Figure 15.** Sheet-metal sculptures (with transducers attached), dancer, and Plexiglas projections, in one version of *Minotaur*.

cases, the temporal patterning of the synthetic alarm sounds used in experiment 2 has governed the number and rate of iterations of triggered sonic elements, with sounds intended to repel dancers from their current respective locations assigned medium and short-duration IOIs.<sup>18</sup> However, changes in physical parameters within the labyrinth cause shifts in the spectral and spatial properties of the triggered sound materials over time. Such a gesture-sound relationship introduces a degree of unpredictability to the electronics.

Given the feedback loops between labyrinth-internal and labyrinth-external audio, this unpredictability is compounded, and impacts not only the auditory, but also the tactile and kinetic dimensions of the constructed labyrinth environment. Consequently, the interaction model becomes multi-modal, multi-directional and time-varying. In the gallery context, audience members have the opportunity to enter the labyrinth and engage with the sensor/actuator system. To an extent, the installation version of *Minotaur* serves as a model for an eventual interactive public presentation of our alarm systems under development, as described below.

## 7. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

This paper was intended to outline how the *alarm/will/sound* project, predicated on the obsolescence of the audible car alarm, and informed by principles of the acousmatic tradition, has developed in multiple (and at times, unexpected) sound design, perception and modelling research and artistic directions. Although the creative works that the auditory warning design

<sup>18</sup>In the myth, as both a dangerous predator and the target of assassination for Theseus and his less-fortunate predecessors, the Minotaur both repels and attracts (at least in a goal-direction sense). As such, it is appropriate that sounds with both high and low repulsion levels are emitted within the labyrinth.

and research aspects of the project have influenced may bear few aesthetic parallels to those associated with musique concrète, the ecological sound classification models and listening mode categories emerging from research closely connected with the Schaeffer legacy, as well as the results of the sound perception experiments conducted under the auspices of *alarm/will/sound*, have borne a significant impact. As such, it is hoped this open-ended and multifaceted project serves as an innovative example of productive cross-pollination between research activities and artistic practices.

The various dimensions of *alarm/will/sound* bear a broad range of possible future artistic and research applications. Besides further studies in car alarm systems specifically, or auditory warnings/audio signals in public spaces, our approach to sound classification, research into sound abstraction/iconicity and creation of a customised (yet adaptable) acoustic descriptor space could inform burgeoning musical information retrieval areas such as searchable sound database construction and tagging. Along similar lines, the measurement of auditory warning attraction and repulsion levels (in experiment 2) and re-purposing of the car alarm as a localised audio communication device would intersect with the interests and objectives of noise pollution and ecological sound data collection, analysis and visualisation projects (such as NoiseTube, Citygram or SONYC).<sup>19</sup>

With respect to artistic extensions of the project, it would be of particular interest to explore auditory warning abstractness vs iconicity, geolocation and attraction/repulsion levels in greater detail. The former two could be addressed in the context of a crowdsourced interactive media installation, for instance. In this scenario, real-world and synthetic auditory warning recordings would first be collected from contributors from across the world and stored in a database. These source-sounds could then be presented in different combinations in objects distributed across a physical and/or virtual installation space. Listeners would be asked in this context to rate the levels of recognisability and attraction/repulsion. As an ecologically based sound perception experiment, this would enable the authors to examine the relationship between familiarity and attraction/repulsion dimensions of globally sourced stimuli. The sounds could then be tagged accordingly, and retrieved from the database via the appropriate search terms.

## REFERENCES

- Blattner, M. M., Sumikawa, D. A. and Greenberg, R. M. 1989. Earcons and Icons: Their Structure and Common Design Principles. *Human-Computer Interaction* 4(1): 11–44.

<sup>19</sup>[www.noisetube.net/index.html](http://www.noisetube.net/index.html), <https://steinhardt.nyu.edu/mar/research/citygram> and <https://wp.nyu.edu/sonyc/> respectively.



- Chion, M. 1994. *Audio-vision: Sound On Screen*. New York: Columbia University Press.
- Gaver, W. W. 1986. Auditory Icons: Using Sound in Computer Interfaces. *Human-Computer Interaction* 2(2): 167–77.
- Gaver, W. W. 1993. What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. *Ecological Psychology* 5(1): 1–29.
- Gibson, J. J. 1977. The Theory of Affordances. In R. Shaw and J. Branford (eds.) *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, NJ: Lawrence Erlbaum.
- Gibson, J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Grey, J. M. 1977. Multidimensional Perceptual Scaling of Musical Timbres. *Journal of the Acoustical Society of America* 61: 1270–7.
- Krumhansl, C. L. 1989. Why is Musical Timbre so Hard to Understand? In S. Nielzen and O. Olsson (eds.) *Structure and Perceptual of Electroacoustic Sound and Music*. Amsterdam: Excerpta Medica.
- Lemaitre, G., Houix, O., Misdariis, N. and Susini, P. 2010. Listener Expertise and Sound Identification Influence the Categorization of Environmental Sounds. *Journal of Experimental Psychology: Applied* 16(1): 16–32.
- McAdams, S. 2013. Musical Timbre Perception. In Deutsch, D. (ed.) *Psychology of Music*, 3rd edn. San Diego: Academic Press.
- McAdams, S., Winsberg, S., Donnadiou, S., De Soete, G. and Krimphoff, J. 1995. Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities, and Latent Subject Classes. *Psychological Research* 58: 177–92.
- Minard, A., Misdariis, N., Houix, O. and Susini, P. 2010. Catégorisation des sons environnementaux sur la base de profils morphologiques. In *10ème Congrès Français d'Acoustique*, Lyon, France.
- Misdariis, N. and Cera, A. 2017. Knowledge in Sound Design – The Silent Electric Vehicle: A Relevant Case Study. *Proceedings of the Conference on Design and Semantics of Form and Movement – Sense and Sensitivity, DeSForM 2017*, Eindhoven, The Netherlands.
- Misdariis, N. and Sigman, A. 2016. Innovative Sound Design of Car Alarms: A Case Study on Information Needs and Musical Creativity. In P. Kostagiolas, K. Martzoukou and C. Lavranos (eds.), *Trends in Music Information Seeking, Behavior, and Retrieval for Creativity*. Hershey, PA: IGI Global.
- Misdariis, N., Minard, A., Susini, P., Lemaitre, G., McAdams, S. and Parizet, E. 2010. Environmental Sound Perception: Metadescription and Modeling Based on Independent Primary Studies. *EURASIP Journal on Audio, Speech, and Music Processing*. <https://asmp-urasipjournals.springeropen.com/articles/10.1155/2010/362013>
- Nelken, I. and de Cheveigné, A. 2013. An Ear for Statistics. *Nature Neuroscience* 16(4): 381–2.
- Norman, D. A. 1999. Affordances, Conventions, and Design. *Interactions* 6(3): 38–43.
- Patterson, R. D. 1982. *Guidelines for Auditory Warning Systems on Civil Aircraft*. London: Civil Aviation Authority.
- Peeters, G. 2004. A Large Set of Audio Features for Sound Description (Similarity and Classification). CUIDADO Project Ircam Technical Report. [http://recherche.ircam.fr/anasy/peeters/ARTICLES/Peeters\\_2003\\_cuidadoaudiofeatures.pdf](http://recherche.ircam.fr/anasy/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf) (accessed 13 April 2018).
- Peeters, G. and Deruty, E. 2008. Automatic Morphological Description of Sounds. *Acoustics 08 Proceedings*, Paris.
- Schaeffer, P. 1966/2017. *Treatise on Musical Objects: An Essay Across Disciplines*. Oakland, CA: University of California Press (English edn pub. 2017 after French edn pub. 1966; trans. C. North and J. Dack).
- Schafer, R. M. 1977. *The Tuning of the World*. New York: Alfred A. Knopf.
- Schalkwijk, L. F. J. 2017. Alarm and Auditory-Interface Design: Learnability of Alarms and Auditory-feedback for Random and Meaningful Alarm Sounds Investigated in a Paired-associate Paradigm. Unpublished Masters thesis, Aalborg University, Aalborg, Denmark.
- Slaney, M. 1998. Auditory Toolbox. Interval Research Corporation, Tech. Rep, 10.
- Stanton, N. and Edworthy, J. 1998. Auditory Affordances in the Intensive Treatment Unit. *Applied Ergonomics* 29(5): 389–94.
- Suied, C., Susini, P. and McAdams, S. 2008. Evaluating Warning Sound Urgency with Reaction Times. *Journal of Experimental Psychology: Applied*, 14(3): 201–12.
- Susini, P., Misdariis, N., Lemaitre, G., Houix, O., Rocchesso, D. and Polotti, P. 2006. Closing the Loop of Sound Evaluation and Design. *2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems*, Berlin, Germany.
- Tuuri, K., Mustonen, M. S. and Pirhonen, A. 2007. Same Sound – Different Meanings: A Novel Scheme for Modes of Listening. *Proceedings of Audio Mostly*, Ilmenau, Germany.
- Vanderveer, N. J. 1980. *Ecological Acoustics: Human Perception of Environmental Sounds*. Doctoral dissertation, Cornell University, Ithaca, NY.
- Wessel, D. L. 1979. Timbre Space as a Musical Control Structure. *Computer Music Journal* 3: 45–52.