From acoustic descriptors to evoked quality of car door sounds

Marie-Céline Bezat[a)]
PSA Peugeot-Citroen, 2 route de Gisy 78943 Velizy-Villacoublay, France

Richard Kronland-Martinet
LMA, CNRS, UPR 7051, Aix-Marseille Univ, Centrale Marseille, F-13402 Marseille Cedex 20, France

Vincent Roussarie
PSA Peugeot-Citroen, 2 route de Gisy 78943 Velizy-Villacoublay, France

Sølvi Ystad
LMA, CNRS, UPR 7051, Aix-Marseille Univ, Centrale Marseille, F-13402 Marseille Cedex 20, France

(Received 28 November 2013; revised 28 March 2014; accepted 21 May 2014)

This article describes the first part of a study aiming at adapting the mechanical car door construction to the drivers' expectancies in terms of perceived quality of cars deduced from car door sounds. A perceptual cartography of car door sounds is obtained from various listening tests aiming at revealing both ecological and analytical properties linked to evoked car quality. In the first test naive listeners performed absolute evaluations of five ecological properties (i.e., solidity, quality, weight, closure energy, and success of closure). Then experts in the area of automobile doors categorized the sounds according to organic constituents (lock, joints, door panel), in particular whether or not the lock mechanism could be perceived. Further, a sensory panel of naive listeners identified sensory descriptors such as classical descriptors or onomatopoeia that characterize the sounds, hereby providing an analytic description of the sounds. Finally, acoustic descriptors were calculated after decomposition of the signal into a lock and a closure component by the Empirical Mode Decomposition (EMD) method. A statistical relationship between the acoustic descriptors and the perceptual evaluations of the car door sounds could then be obtained through linear regression analysis. © *2014 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4883364]

PACS number(s): 43.50.Pn, 43.50.Qp, 43.60.Wy, 43.66.Lj [BSF] Pages: 226–241

I. INTRODUCTION

Impact sounds associated with a human action have the distinct characteristic of transmitting to the listener, within a very short time frame, a wealth of information regarding the action and its context. Examples of such impact sounds abound: Sherlock Holmes tapping on walls to uncover hidden compartments (cited by Van den Doel and Pai, 1996) or potential car buyers tapping on the dashboard to verify the quality of a vehicle they would like to purchase. But how do listeners attribute meaning to these impact signals that mark their actions?

Here, we investigate the perception of one particular kind of impact sounds, namely, sounds produced when closing car doors (also referred to here as door-closure sounds). A customer in a car dealership tests the car door and perceives that the door is closed based on the sound of it closing. This sound also evokes for the client complex inferences regarding the quality and soundness of the vehicle, features to which the manufacturer pays special attention. Translating such complex cues into technical specifications is not an easy task. Characterizing a perceived quality on the basis of certain features extracted from the sound signal constitutes a first stage in the process which, if followed by the characterization of the organic sources and the acoustic transfer, could yield the desired technical specifications. This study focuses on this first stage, which aims at understanding what one perceives when hearing the door closure sound and extracting from the acoustic signal the underlying acoustic features that make it possible to predict the evocations induced by such sounds.

Investigations of door-closure sounds have already been undertaken from different view-points by several manufacturers. These studies have either focused on perceptual judgments, on acoustical phenomena or on the perceived impact of direct modifications on the physical sources of the car door. In the perceptual studies, the sounds are either characterized by their loudness (Hojberg, 1991; Fridrich, 1999), their pitch (Fridrich, 2007) or by multiple impacts linked to the lock and the door panel (Malen and Hancock, 1995; Champagne and Amman, 1995; Petniunas *et al.*, 1999; Blommer *et al.*, 2005; Hamilton, 1999). Kuwano *et al.* (2006) presented a study of subjective impressions of car door sounds using semantic differential with a predefined adjective scale showing coherent quality evaluations across subjects. Signal analyses have revealed the presence of impacts and vibrations, but very few specific acoustic parameters explaining the perceived phenomena are identified (Scholl and Amman, 1999). One study (Parizet *et al.*, 2008) used timbre parameters, frequency balance, and cleanness (only one temporal event is required) to predict preference

a)Author to whom correspondence should be addressed. Electronic mail: ystad@lma.cnrs-mrs.fr

 0001-4966/2014/136(1)/226/16/$30.00

ratings. This approach is based on paired comparison listening tests for similarity and quality ratings of 12 sounds and on the statistical links between signal parameters and perceptual dimensions. In a recent study, Shin *et al.* (2013) related the spectral kurtosis to door closure signals processed by an auditory filter model in order to objectively quantify the sound quality of door-closure sounds. Finally, by directly modifying the physical car door sources (joints, lock mechanisms, door panel), the weight of the car door was found to influence the perceived quality (Scholl and Amman, 1999).

By means of experiments or expert listening, these publications provide mechanical solutions to improve the sound quality, or signal indicators (loudness, frequency ratio) and signal analysis methods (wavelet transformation, percentile frequency method), which contribute to the understanding of signal morphologies that influence the sound quality. However, none of them propose a model that generally well predicts the sound quality ratings of a large number of door closure sounds. It appears that providing robust predictions of perceptual properties does not sufficiently account for why the sound is well perceived or otherwise. We assume that a better description of the percept of the door closure sound is necessary and that a first step should consist in being able to distinguish a good versus a bad door-closure sound before attempting to link sound quality to signal parameters. The perceptual properties reflected by a sound quality are best evaluated using a battery of perceptual tests. Following this, a link can be established between these properties and the evoking of a quality, on the one hand, and between the perceptual properties and the acoustic correlates, on the other hand, to identify *in fine*, the signal indicators that best characterize the percept of the door-closure sound.

When investigating the perceptual evaluations of sounds, it is important to be aware of the fact that listeners may focus on different aspects. Gaver (1993b,a) distinguished ecological or everyday listening from analytical or musical listening. Everyday listening is the experience of hearing events in the world rather than sounds *per se*. He suggested that, with ecological listening, a sound contains information regarding the event, the place, and the environment, and drawing from the ecological approach to perception, he went on to propose a new framework for describing sounds on the basis of the attributes related to the source.

In the case of ecological listening of a simple source, the listener pays attention to the sound producing object, such as its size (Lakatos *et al.*, 1997) and the material of which it is composed (McAdams *et al.*, 2004; Giordano and McAdams, 2006; Aramaki *et al.*, 2011). In the case of more complex situations reflecting for instance interactions between sound sources, the listener perceives properties related to the event as a whole. Warren and Verbrugge (1984) showed that objects that bounce and break can be distinguished by listeners with a high degree of accuracy, while Repp (1987) revealed that subjects were able to recognize their own recorded clapping and the hand position from recordings when someone else is clapping. More recently, Thoret *et al.* (2014) showed that subjects were able to recognize biological motions and certain shapes from friction sounds produced when a person is drawing on a paper.

The ability to identify physical sources also depends on the experience of the listener: A car mechanic can repair a mechanical problem using the sound of the motor, similarly a doctor can detect an abnormal sound in a patient's heartbeat (examples cited in McAdams and Bigand, 1993). Therefore, in an industrial context, to determine precisely the links between the percept and the mechanical aspect, it could prove helpful to draw on the experience of professionals who are expert listeners of the sources on which they work.

These examples illustrate the listeners' capacity, when presented with complex sounds, to perceive with ease their properties and the events related to them, in other words, to decode the origin of the sounds. This perception of ecological properties is, in all these examples, dependent upon the listener's need to make sense of their environment (and to survive in it).

In the case of analytical listening, the listener will focus on intrinsic sound properties linked, for instance, to loudness, pitch, and timbre rather than trying to identify the sound producing source and its properties. Electroacoustic composers have widely exploited this aspect by using so-called abstract sounds (for which the source cannot be easily recognized) to help listeners focus on intrinsic sound properties (Merer *et al.*, 2011; Schaeffer, 1966). The identification of analytical properties is also very interesting from a scientific point of view, since a bridge between these properties and the acoustic signal can more easily be formed than between ecological properties and the signal.

Sensory analysis is one method that can be used for evaluating the analytical properties of a group of sounds, and as such it serves as an alternative to dissimilarity evaluations and semantic differential rating. Its strength lies in its ability to identify well defined perceptual dimensions without the need for *a priori* knowledge of their continuous or categorical nature. In addition, sensory analysis permits the evaluation of a large number of sounds as opposed to dissimilarity testing which only allows for a reduced number of evaluations due to the paired comparison procedure that causes the number of evaluations to grow rapidly with the number of stimuli. Finally, sensory analysis carries out its analysis essentially by analytical listening.

The method of sensory analysis was introduced by the agri-food industry, and it still remains little used in psychoacoustics (Roussarie *et al.*, 2004; Poirson, 2006). It involves the identification of the sensory descriptors characterizing a group of sounds and of the evaluation of the sounds using each of these descriptors. Onomatopoeias are often recruited as descriptors in sensory analysis, for example, the sound of a motor could be described as "ON" "AN" (Roussarie *et al.*, 2004; Sciabica *et al.*, 2009). They provide a means of studying the analytical characterization of sounds. An example of this is the analogy between guitar sounds and vocal sounds that has proved fruitful in the synthesis of guitar sounds using vocal imitation (Traube and Depalle, 2004; Traube and D'Alessandro, 2005). Similarly, vocal imitations have recently been used to extract relevant features of kitchen sounds (Lemaitre *et al.*, 2011). It was revealed here that vocal imitations enabled the listener to recover most of the imitated sounds (except for liquid sounds) and that acoustic features responsible for the sound event recognition also could be identified and used to predict the categorization of the referent sounds.

As already mentioned, analytical sound properties give a closer description of the perceptually relevant parameters of the acoustic signal than ecological sound properties and is therefore essential when a link between perceptual evaluations and the acoustic signal is sought. These properties can also to some extent guide the signal analysis process like for instance when the evaluations are to be interpreted in terms of timbre, frequency, or loudness variations or when separate sources are identified and described by the listeners. In such cases, a decomposition of the signal can be useful before the acoustic descriptors are calculated and correlated with the perceptual evaluations. Knowledge of the auditory peripheral system can also be useful prior to calculating descriptors as seen in Freed (1990) who processed sounds of impacting pans with the Stautner (1983) model, which constitutes a filtering by critical bands as defined by Moore and Glasberg (1983). Giordano and McAdams (2006) simulated the average outer ear through filtering and the cochlea by gammatone filters, as defined by Patterson *et al.* (1995), to process signals of impacted materials before calculating the acoustic descriptors. In this particular study, we separated the signal in two parts according to the sensory analysis that revealed two important components of the door-closure sound, namely the lock mechanism and the resonance from the door itself. The Empirical Mode Decomposition method introduced by Huang *et al.* (1998) turned out to be a satisfactory method for such decomposition, possibly because it uses a constant $\Delta f/f$ filtering analogous to the functioning of the basilar membrane of the ear (Moore and Glasberg, 1983).

From an acoustical point of view, previous analyses on impact sounds have revealed two important physical aspects vis à vis the ear, i.e., the damping law, which characterizes the behavior of matter, and the distribution and number of frequency components that are most typical, given the geometric form of the source (Wildes and Richards, 1988; Lutfi and Oh, 1997; Klatzky and Pai, 2000; Aramaki and Kronland-Martinet, 2006; Aramaki *et al.*, 2007). The spectral distribution of energy depends upon the stimulation. The greater the amount of energy produced by an impact, the greater the number of modes stimulated. When modal density is high, individual components are not distinguishable and can be considered as a stochastic process with a given spectral distribution and a damping law characteristic of the material in question. Particular attention should therefore be given to these aspects when searching for a relation between the acoustic descriptors of the car door and the perceptual evaluations that enable the prediction of perceived quality.

In the first part of this paper, ecological properties of door closure sounds identified through a series of listening tests with naive and expert listeners are presented. In the second part analytical sound properties are investigated through sensory analysis. Links between the analytical and ecological properties are then established. Further, acoustic signal descriptors of each of these contributions are calculated in order to propose a prediction model of the perceptual properties based on linear combinations of the acoustic descriptors. A complete characterization of the perceptual properties and acoustic correlates concludes this article.

II. CHARACTERIZATION OF DOOR-CLOSURE SOUNDS BY ECOLOGICAL PROPERTIES

The perception of environmental sounds occurs essentially in a natural setting. Qualitative studies resemble most closely a natural context: Subjects are presented with an object and react to it in a natural manner without focusing on its sound. In a preliminary qualitative study, a description of how a door-closure sound is perceived was put forward by inspecting subjects' interaction with vehicles of different brands in a show-room (internal report). We gathered reports, recordings, verbatim of their thoughts regarding the most relevant intrinsic properties and associations to describe door-closure sounds. We found that listening to the sounds in a real context evoked impressions of solidity and, on a more general level, impressions of quality. In addition to these impressions, the sounds reflected the natural properties characterizing the source event, such as the efficacy of the gesture (e.g., the door is fully closed), specific information regarding the object (e.g., the weight of the door), and specific information regarding the action (e.g., the energy used in closing the door).

A. Ecological properties from naive listeners

Having identified the ecological properties in context, laboratory-based experiments were set up to study, in a detailed and quantitative manner, listeners' natural evaluation of a corpus of representative sounds.

1. Stimuli

To obtain a corpus of door-closure sounds that at best represents any car door sound, internal experts from the PSA company were asked to judge a large number of cars (spanning from high to low quality cars and from small to large sizes) to select the most varied sounds.

The door-closure sounds were measured outside the vehicle, with all windows closed and in a Semi-Anechoic Chamber (SAC). A HEAD Acoustics dummy head was used to record the door-closure sounds. The acoustic head was placed outside of the vehicle in the natural position of someone closing the door. It was mounted on a stand at a height of 1.6 m, at a distance of 0.7 m at the exterior of the door and 0.4 m behind it, and at an angle of 45°.

Sounds from both the front and rear doors of nine vehicles of different brands and segments (i.e., size of the vehicle) were hereby selected for the experiment.

2. Subjects

Two experiments were effectuated on two groups of subjects. In the first group, 120 subjects, were recruited (72 male and 48 female); 48 subjects were less than 35 yr, 48 were between 35 and 50 yr, and 24 were older than 50 yr. In the second group, 80 subjects were recruited (48 male and 32 female); 20 subjects were less than 35 yr, 40 were between 35 and 50 yr, and 20 were older than 50 yr. All owned vehicles bought within the last 4 yr. Half of the subjects were owners of a vehicle from the PSA Peugeot Citroen group, the other half owned other brands. Half owned a vehicle of segment B (vehicles of type Peugeot 208, Citroen C3), the other half owned a vehicle of segment M (vehicles of type Citroen C4 and C5, Peugeot 3008). The subjects were paid for their participation.

3. Procedure

All listening sessions were carried out in a room with low ambient noise-level using a Hewlett Packard Personal computer, Head Acoustics PEQ IV pre-amplifiers, and Head Acoustics HA II electrostatic headphones. The listening was binaural.

In these two experiments, the stimuli were evaluated on the basis of the five perceptual properties; closure energy, quality, solidity, weight of the door, and effectiveness of the door closing action. In the first experiment, the 120 subjects of the first group were asked to evaluate the sounds on the basis of the closure energy. The four remaining properties were evaluated on the same sounds by the 80 subjects of the second group. Note that the closure energy was evaluated by a separate group for practical reasons. In fact, this evaluation was made at an earlier stage during an investigation on the relation between the gesture and the closure velocity.

The closure energy, quality, solidity, and weight of the door were evaluated on continuous scales. The closure energy scale ranged from "gently" to "forcefully" (in response to the question "Listening to the sound of the door closing, do you think the door was closed gently or forcefully?"). The quality scale ranged from "very low quality" to "very high quality," the solidity scale ranged from "not at all solid" to "very solid," the door weight scale ranged from "very light" to "very heavy." The success of door closure was evaluated on a categorical scale of four modalities, i.e., "strongly disagree," "inclined to disagree," "inclined to agree," "strongly agree" (in response to the question "Do you think the door is fully closed?"). The graphical user interfaces (GUI) used for these evaluations were either designed in Labview or Fizz. Three representative examples of car door-closure sounds were first presented to familiarize the subjects with the stimuli before the test started. Then the stimuli were presented in random order using a Latin squares design.

4. Results

The judgments made using the continuous scale (ranging from 0 to 10) were analyzed using Ascendant Hierarchical Classification (AHC) while the judgment of the door-closure effectiveness, evaluated on a categorical scale, was analyzed by a χ^2 test.

The subjects' continuous evaluations (door closure energy, quality, solidity, and weight of the door) revealed a mutual agreement between subjects (Duncan comparison indicate that there is no significant differences between the judgments of the two subject classes identified as best by AHC). The mean evaluations and the associated Duncan classes are presented in Fig. 1. For all the properties

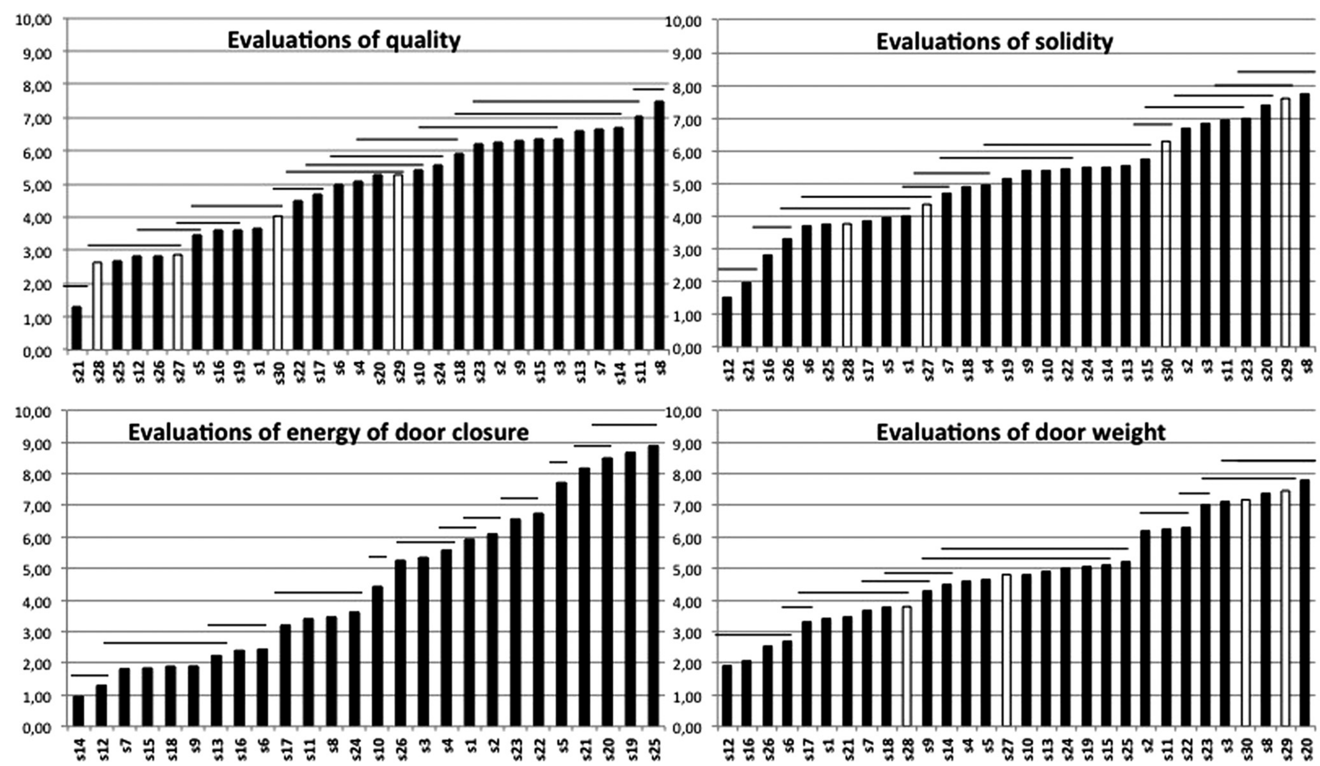

FIG. 1. Evaluations of quality, solidity, energy of closure and door weight. The horizontal bars group those stimuli whose discrimination does not reach significance according to the 5% Duncan test. Note that sounds s27 to s30 (white bars) are not considered in this study, since they were not tested with respect to the door closure energy.

considered, the stimuli were very well discriminated and distributed over the rating scales. The selected properties are thus meaningful in the characterization of our sound corpus. Pearson's correlations between the continuous properties reveal that solidity is strongly related to quality ($r = 0.75$) and to the door weight ($r = 0.90$), while it is not correlated to the closure energy ($r = 0.08$). There is a medium correlation between quality and door weight ($r = 0.46$), and a negative medium correlation between the quality and the closure energy ($r = -0.50$).

The door-closure effectiveness was evaluated on a scale of four modalities. The first two modalities consider the door as badly closed, and the second two consider it as fully closed. The subjects' responses, summed over modality pairs, were analyzed by a χ^2 test. Five door-closure sounds (s12, s16, s21, s24, and s25) did not evoke the association of a fully closed door and two of them (s12, s21) were clearly identified as originating from badly closed doors (despite being fully closed when originally measured).

By employing the perceptual properties used for the direct evaluation during the previously presented tests, we ensured that the door-closure sounds could be well discriminated and in a consensual manner. This also reassured us as to the relevance of these properties for describing the door-closure sound and their possible evaluation in a laboratory setting.

B. Ecological properties from expert listeners

Depending on the expertise of the listener, the properties perceived vary in the preciseness of their relevance to the source. In a qualitative study, a "naive" subject will listen to the global event, extracting the information useful for his/her needs, such as deciding if the door is fully closed or if it is heavy. An expert in the area of automobile openings has different requirements and will decompose the sound into its organic constituents (e.g., lock or joints). Therefore, we make the distinction between "naive" ecological listening and "expert" ecological listening.

1. Experts

Four experts on car door manufacturing and acoustics employed by the PSA Peugeot Citroen group (mean age 45 yr) participated in the experiment. All the participants were male and were not paid for their participation.

2. Procedure

The experts were asked to evaluate the 26 door-closure sounds used in the previous experiment by informal listening tests. The experiment took place in a room with low ambient noise-level using a Hewlett Packard Personal computer, Head Acoustics PEQ IV preamplifiers, and Head Acoustics HA II electrostatic headphones. The listening was binaural and they could listen to the sounds as often as they wished.

The experts were asked to use their own criteria (in terms of the lock mechanism, vibrations, joint noise, door rod noise, ringing tones, etc.) that they normally apply to characterize door-closure sounds *in situ* during a standard working situation. The sounds were presented randomly and the experts entered their descriptions directly in a text file on the computer. The experts finally took counsel together and managed without difficulty to deliver one global judgment of the stimuli.

3. Results

Two main aspects turned out to be important among the experts, i.e., the lock noise and the door panel contribution. According to these perceived elements, three categories of door-closure sounds resulted from this experiment:

(1) sounds representing mainly the sound of the lock ("lock sound": s1, s6, s7, s12, s16, s17, s18, s26);
(2) sounds representing a mixture of the sound of the lock and the impact of the door panel ("lock/closure sound": s2, s4, s5, s8, s9, s10, s11, s13, s14, s15, s21, s24);
(3) sounds in which the sound of the lock was not distinctly heard so that only the impact of the door panel was perceived ("closure sounds": s3, s9, s20, s22, s23, s25).

In addition, certain specific observations (specificities) regarding certain door-closure sounds were made. In two sounds the movement of the notches of the locks could be heard separately with ease (s12, s16). A sound related to the joints was heard in sound s24. Sounds s21 and s25 contained vibrations.

III. CHARACTERIZATION OF DOOR-CLOSURE SOUND BY ANALYTICAL PROPERTIES

In the previous section, we saw that evocations linked to natural properties of sounds (such as perceived quality or solidity) can be revealed through qualitative studies. To obtain a perceptual description of analytical sound properties that refer to more low-level perceptually relevant signal parameters, alternative experimental protocols are needed. Previous studies have used particular stimuli such as abstract sounds for which the source cannot be easily recognized in order to induce analytic listening. In a study on the attribution of meaning of sounds, abstract sounds were used in a word-sound priming experiment using the event-related brain potential (ERP) method. Results showed that a conceptual relation between words and sounds also takes place when subjects focus on analytical sound properties (Schön *et al.*, 2010). Abstract sounds were also used to assess abnormal perceptual experiments in schizophrenia (SCZ). Results highlighted ambivalence in familiarity and bizarreness in SCZ that might be due to the coactivation of ecologic and analytic listening in a conflicting way (Micoulaud-Franchi *et al.*, 2012). Finally, in a study on perceived motion of sounds, abstract sounds were used to identify motion categories and propose intuitive control parameters for synthesizing sounds evoking motion (Merer *et al.*, 2011; Merer *et al.*, 2013).

For other types of sounds such as the ones we are dealing with in this study, alternative approaches that incite the subjects to focus on intrinsic sound properties should be proposed. One such approach is the sensory metrology method that consists in training a group of people to identify a set of descriptors (acoustic or onomatopoeia) to characterize the sounds. In the present section, we first present this method and how analytical properties can be identified before we describe how these properties can be linked to the ecological properties found in the previous section.

A. Identifying the analytical properties

Sensory analysis consists in identifying a set of sensory descriptors that characterize a group of stimuli and further evaluate the stimuli with respect to each descriptor. In the present case where the stimuli are sounds, the sensory descriptors are either acoustic descriptors or onomatopoeias. Subjects first go through a training phase to improve their sensory acuity before they evaluate the descriptors.

1. Stimuli

The corpus used in this study consisted of 26 selected door-closure sounds (s1 to s26).

2. Sensory panel

The panel was composed of nine external subjects (the judges), nonautomobile experts, chosen on the basis of their sensory and logical aptitude and their motivation. Their audition was tested. All of the nine judges were female with an average age of 45 yr. They were paid for their participation.

3. Procedure

The sessions lasted for 4 weeks. The availability of equipment meant that the panel had to be divided into two groups of four and five subjects. Each group performed two sessions of 2 h per week, meaning that a total of eight sessions of 2 h were performed. The sensory profile was developed in four stages: A semantic training phase to establish a list of sensory descriptors characterizing the sound space, a sensory training phase to improve the subjects' sensory acuity, a metrical training phase which consisted of using the judgment scale and a final evaluation phase.

a. Semantic training (two sessions). The objective of this phase was to establish a list of sensory descriptors. The task of the judges was, first, to describe their sensations based on their own individual listenings. Following this, the experimenter discussed the initial list of descriptions with the judges, in order to reduce the list by grouping together those expressing the same sensations.

b. Sensory learning of the sensory profile (two sessions). The judges were trained, on the one hand, to improve their sensory acuity and, on the other hand, to ensure that a consensus could be reached regarding each of the descriptors. To this end, the judges carried out tests in which the sounds were classified based on three intensity levels (weak, medium, and strong). The judges' individual classifications were then compared and discussed in order to reach a consensus regarding them.

c. Metrical learning of the sensory profile (two sessions). The judges memorized the extremes of the classified sensations and practiced using the judgment scale. The experimenter set up discussions, as required, to improve the consensus within the group, highlighting difficulties (for example, a non-discriminating descriptor) and verifying the correct use of the scale (use of the limits, the distribution of the sounds over the scale,…).

d. Sensory profile evaluation (two sessions). Finally, the subjects carried out an evaluation of the sounds, descriptor by descriptor and sound by sound. These evaluations were conducted two to three times to verify the panel's consistency. The order of presentation of the sound samples was based on a Latin squares design. The GUI was controlled by Fizz software.

e. Sensory descriptors. The sensory descriptors embody the analytical properties that can originate from classical descriptors of sounds such as perceived intensity, frequency or timbre, or they may be generated from onomatopoeia specifically characterizing the door-closure sounds, like BON'M or KE. Details of their definition are presented in Table I along with the related type of scale. Certain descriptors, such as the intensity, are measured on a continuous scale, others such as the properties BON'M KE, and ESPACHOC have continuous values only if they are present in the sound as they are only perceivable in certain sounds. Finally, the descriptors "gris-gris" and "Triangle" have binary values, since they are either present in the sound or absent from it.

B. Performance of the panel

For each continuously valued sensory descriptor, the panel's performance was measured with the aid of three criteria: The discriminability of the sounds, the consistency of the judges from one session to the next and the consensus between the judges. To test these three criteria, a three-factor (Sound, Judge, Consistency) analysis of variance was carried out. The factor Judge, was set as random to take into account the Judge Sound interaction in the effect of Sound. The results of the analyses are shown in Table II.

The analyses reveal that the effect of Sound is significant, meaning that the sounds were well discriminated. The

TABLE I. Definition of analytical properties of door-closure sounds obtained from the sensory panel.

Descriptors	Definition	Type
INTENSE	Possessing a high sound level	Continuous
LONG	Possessing a long duration	Continuous
BON'M	Presenting a loud "BON'M" sound.	Abs/Continuous
KE	Begins with a sound pronounced as "KE"	Abs/Continuous
ESPACHOC	Presenting clear silences between two impacts	Abs/Continuous
Triangle	Presenting a loud high note at the end of a sound	Specific
gri-gri	Presenting strong oscillations	Specific

TABLE II. Performance of the panel obtained from an analysis of variance based on the three criteria: Discriminability of the sounds (Sound), repeatability for each of the judges from one session to the next (Judge) and consensus between the judges (Consistency).[a]

	Df	INTENSE	BON'M	ESPACHOC	LONG	KE
Sound	25	36.23***	29.01***	28.84***	14.09***	20.77***
Judge	8	13.48***	16.65***	12.8***	9.84***	24.41***
Consistency	2	1.31	0.47	0.82	3.36	1.09
Interaction 1–2	200	1.3*	2.76***	3.41***	3.99***	2.97***
Interaction 1–3	50	0.91	0.98	1.2	0.85	0.95
Interaction 2–3	16	7.23***	1	3.66***	1.99*	2.92***

[a]The p values are reported by ***($p < 0.001$), **($p < 0.01$), *($p < 0.05$) when coefficients are significant.

effect of Consistency as well as the Sound-Consistency interaction are not significant. The panel was, on the whole, repeatable despite a lack of consistency in the case of one or more judges regarding certain descriptors (the Judge-Consistency interaction is significant). Finally, the effect of Judge is significant, meaning that the judges did not all use the scales in the same manner. The Judge-Sound interaction is also significant, thereby revealing a problem of consensus. The individual performances of each judge were also examined with respect to the three previously mentioned criteria, i.e., discriminability, repeatability, and consensus. The results of the analyses are shown in Table III.

The ability to discriminate the sounds is confirmed by a one-factor analysis of variance for the nine subjects and the five continuous descriptors ($p < 0.001$) (except for judge 9 who did not evaluate the descriptor INTENSE). The consistency is measured by taking into account the standard deviation of all the scores attributed by a judge to all the sounds during all repeated evaluation sessions (Table III). The panel demonstrated good consistency for the descriptors, INTENSE, BON'M, ESPACHOC, and LONG, and a decrease in consistency for the descriptor KE (four subjects showed only modest consistency). The consensus was analyzed using the percentage of inertia presented by the first axis of the PCA (Sound, Judge) for each descriptor. It gives an estimation of the spread of the judges over each component. With an inertia of more than 70% on the first axis, we found that the judges were in good agreement overall regarding the descriptors ESPACHOC (80%), BON'M (79%), KE (74%), and INTENSE (88%), but they were less so for the descriptor LONG (65%).

TABLE III. Standard deviation of marks given to all the sounds and repetitions. The descriptor KE obtained the lowest consistency (Std > 2) for 4 subjects.

Subject	INTENSE	BON'M	ESPACHOC	LONG	KE
Subject 1	1.8	**2.08**	0.48	1.44	**2.06**
Subject 2	1.35	1.37	1.64	1.51	1.77
Subject 3	1.23	1.11	1.33	1.45	1.86
Subject 4	1.50	1.22	1.97	1.39	**2.13**
Subject 5	1.99	1.60	1.43	1.65	**2.07**
Subject 6	1.69	1.80	1.07	1.56	1.66
Subject 7	1.25	1.32	0.81	1.57	1.88
Subject 8	**2.06**	1.80	0.96	1.72	1.53
Subject 9	**X**	1.66	0.91	1.64	**2.06**
PANEL	**1.61**	**1.55**	**1.18**	**1.55**	**1.89**

Based on these analyses, we can conclude that the performances of the panel were satisfactory and can proceed to the consideration of the mean evaluations in the analytical characterization of the door-closure sounds.

C. Sensory profile

Certain sounds feature specificities, such as the specificity, "gris-gris" (Table I), which describes the vibrations in a sound, can be identified in the sounds s21 and s25, and the specificity, "Triangle," which makes reference to a high frequency component at the end of a sound, can be heard in the sounds s13 and s22. These specificities become detectable after a certain number of occurrences.

We obtained a characterization of the door-closure sounds using analytical listening. However, describing sounds by a set of analytical properties is only useful if these properties are related to the ecological properties, i.e., what the listeners perceive in their natural environment. A more detailed description of the analytical properties is therefore given in the next section where all the perceptual properties are linked.

IV. A NETWORK OF PERCEPTUAL PROPERTIES

In this section, we describe how the evocations of quality and solidity can be linked to ecological properties and how "expert" and naive ecological properties can be linked to analytic properties to form a network of perceptual properties of car door sounds.

A. Link between ecological properties from naive listeners

The ecological evaluation from naive listeners revealed that 5 out of the 26 door closure sounds studied were characterized as "badly closed" doors. Although we do not have sufficient data to draw a definite conclusion regarding this property, we believe that it might have affected the judgment of perceived "quality" and "solidity." It is worth noting that these sounds differed from the others in distinct manners as pointed out by the experts: In two of the sounds the individual movements of the lock notches was heard with ease (s12, s16), a sound related to the joints was heard in sound s24 and sounds s21 and s25 contained vibrations. These sounds were therefore not included in the further analyses.

The relationships between the perceptual properties, "door weight" and "energy of closure" and the evocations "quality" and "solidity" were established by submitting the corpus of the 21 remaining closure sounds considered as reflecting "fully closed" doors to a linear regression analysis. We used the LASSO regression (Matzner-Løber, 2006) to estimate the models of these related data sets. The LASSO regression offers a solution to the problem of the collinearity of the variables by including a bias in the estimation of coefficients, hence facilitating the reduction of the variance. It yields more stable model coefficients compared with multiple linear regression analysis, and it is selective. In addition, the models are estimated using the "leave-one-out" cross-validation criterion to assure their robustness (Matzner-Løber, 2006). Various models are determined by training on all $n-1$ data sets, and based on these predictions, the coefficient of determination (R2val) is calculated. The regressions allow us to quantify the relative importance of the properties, "door weight" (P) and "energy of closure" (E) in the evaluations of the associations' quality and solidity [Eqs. (1) and (2)],

$$Estimated\_quality = 5.49 + 1.08P - 1.20E, \quad (1)$$

$$Estimated\_solidity = 5.43 + 1.34P - 1.38E, \quad (2)$$

where the coefficient of determination of the estimated quality is $R2val = 0.88$ and of the estimated solidity is $R2val = 0.92$.

These results reveal that the evocation of quality and solidity are obtained from a heavy door and a gentle gesture. The perceived energy of door closure had a greater negative influence on the quality (53%) than on the solidity (22%) (this association being effectively, more related to the object than to the action, whereas the quality of the vehicle can incorporate the notion of a well tuned door that can be closed gently). Thus, we obtain a perceptual characterization of the ecological properties of the door-closure sounds. The "quality" evoked by the sound of the door closure depends on the perceived "door weight" and the "energy of closure"; a heavy door linked to a gentle gesture evokes an impression of quality and vice versa. The impression of "solidity" depends mainly on the perceived weight of the door. A door may be judged as "badly closed" if one focuses on specific characteristics such as the slow movement of the lock's notches, vibrations or joint noise.

B. Link between expert ecological listening and analytic properties

The expert descriptions of the car door sounds can be directly compared to the analytical properties. Sounds s21 and s25 which, according to the automobile experts, contain vibrations are characterized by the presence of a "gri-gri," in sensory terms. Sounds s12 and s16, in which one can hear the sound of the movement of the lock notches, are the two sounds containing the most ESPACHOC. In contrast, the presence of joint noise is not characterized by the sensory panel, as only one sound (s24) contains this specificity. The presence of the lock sound is linked to the descriptors KE, BON'M, and INTENSE. In fact, we notice that those sounds featuring mainly the lock sound ("lock" sounds: s1, s6, s7, s12, s16, s18, s26) all score highly in relation to the property, KE (>4) (and vice versa). We also noted that sounds scoring weakly in the evaluation of KE and highly in the evaluation of BON'M and INTENSE correspond to those for which the lock sound was not distinctly heard and for which the impact of the door panel was clearly perceived ("closure" sounds: s3, s19, s20, s22, s23, s25). Therefore, we can say that the property, KE, is related to the presence of a lock and the property, BON'M, is related to the closure.

For sounds s12 and s16, the specificity, "badly closed", can be related to a strong ESPA-CHOC (movement of the lock notches); for sounds s21 and s25 it can be associated with the presence of "gri-gri" (vibrations). Sound s24, associated with a "badly closed" door, is characterized, not by an analytical property, but from an organic point of view as it is the only sound to contain a strong joint noise. However, too few sounds feature these specificities to allow us to generalize these observations. The specificity, TRIANGLE, does not affect the impression of a door that is badly closed.

C. Link between naive ecological listening and analytic properties

The links between the continuous natural properties, "door weight" and "energy of closure," and the analytical properties were determined by LASSO regression [Eqs. (3) and (4)], using the evaluations of the 26 door-closure sounds,

$$Est\_door\_weight = 5.01 + 0.76\,BONM - 0.72KE, \quad (3)$$

$$Est\_closure\_energy = 4.54 + 2.45\,INTENSE, \quad (4)$$

where the coefficient of determination of the estimated door-weight is $R2val = 0.81$ and of the estimated closure energy is $R2val = 0.93$.

The perceived weight of the door is linked to the analytical properties KE and BOMN, which is dependent on the pitch of the sound. The perceived energy of closure is linked to the descriptor INTENSE.

The evoked associations (quality and solidity), the ecological properties (door-weight, energy of closure, badly closed door and the organic sources), and the analytical properties (INTENSE, BON'M, KE, ESPACHOC, gri-gri) all contribute to the perceptual characterization of the door-closure sounds. The evoked associations are related to the ecological properties characteristic of the event as perceived in a real context, while the natural properties, in turn, can be estimated by analytical properties. Therefore, the various perceptual properties can be considered as forming part of a perceptual network presenting a complete perceptual description of the door-closure sound. If the analytical properties can be described by acoustic correlates, it should be possible to extract what we perceive from the acoustic signal as well.

V. CHARACTERIZATION OF DOOR-CLOSURE SOUNDS BY SIGNAL PARAMETERS

In the previous sections we have shown how different levels of perceptual properties can be linked together from the analytical (i.e., low-level) properties to evoked associations (i.e., high-level properties). The analytical properties constitute useful cues for the acoustic modeling of ecological properties. Modeling the features KE and BON'M of the sound seems more accessible as they more closely characterize the sound signal than the ecological properties associated with the door-closure sound, which are dependent upon high-level cognitive processing.

Beyond these specificities which each require a distinct analysis (detecting a flaw), not dealt with in this part of the study, the perceptual characterization of the door-closure sounds is also based on a balance between the properties KE and BON'M, which are themselves linked to the presence of the lock sound. After decomposing the door-closure sound "perceptually" into KE and BON'M components we present a method for decomposing the signal with respect to these perceptual descriptors to finally determine a list of associated acoustic descriptors for each of these components.

A. Signal analysis with analytical properties in mind

The door-closure sounds are composed partly of a KE component related to the "lock" and partly of a BON'M component related to the "closure" (sound of car door/trunk panel impact). Therefore, we were able to tackle our problem better by including a source separation stage using these two principal components. A method of source separation by physical models was considered too problematic to apply in this case given the number of inter-dependent sources contributing to the sound (joints, door, lock, etc.). Furthermore, our aim was to carry out a separation into "perceived" sources. In fact, that which the automobile expert perceives as a "lock" sound or the sensory expert as "KE" is not strictly related to the impact sound of lock notches on the strike plate. Joint noise or noise from the door rods could combine with it at the very moment it contributes to the metallic aspect of the sound. In contrast, filtering in the time-frequency domain did not succeed in separating the "lock" and "closure" sounds, since it distorted either the lock impact sound or the low frequency resonance. Actually, both "lock" and "closure" sounds are characterized by their impulsive behavior, implying broadband frequency contents that may overlap and make a separation by filtering impossible. Moreover, these two components occur almost simultaneously, meaning that a time separation is also impossible. Thus, we proposed to use an alternative signal decomposition approach called the Empirical Mode Decomposition (EMD) method that decomposes a signal into a number of independent basic modes entangled in the time-frequency domain. These modes can further be combined to reconstruct the lock and the closure contributions, even though they were mixed in the time-frequency domain.

1. Empirical Mode Decomposition

The EMD method was introduced by Huang *et al.* (1998) and has proved attractive because of the simplicity of its underlying principles, despite not permitting an analytic formulation. The essence of this method is to identify iteratively intrinsic mode functions (IMFs) of the signal, both amplitude and frequency modulated, separating locally (on an oscillatory scale) a "fast" component from a slower pattern. This is done by pointing out the located maxima and minima of the signal, then constructing the superior and inferior envelopes, and finally the mean envelope. The first mode is thus obtained. The algorithm is then processed on the rest of the signal, until the second mode is obtained. By finding EMD modes, the rest of the signal (the residue) has less and less extrema. The decomposition process stops when the last residue has only three extrema. The signal can be perfectly reproduced by simple addition of the modes. In the absence of *a priori* information regarding the decomposition, the iterative process of EMD explores a signal sequentially according to its "natural" scales. The EMD method is only defined algorithmically; it is studied from the point of view of its theory, its performance and its application, following numerical simulations using well-controlled signals. It also presents, in particular, the properties of locality, adaptability and time-frequency filtering (Flandrin and Gonçalves, 2004; Flandrin *et al.*, 2004). This method, intuitively used, was justified *a posteriori*.

2. Applying the decomposition to the door-closure sound

We applied the EMD method to decompose the door-closure sound into a "lock" sound and a "closure" sound. This decomposition is illustrated with three examples: A sound that is more "lock" in character (s26), a sound that is more "closure" in character (s20) and a sound that contains both, "lock/closure" (s4). We began by carrying out an EMD decomposition, then we recombined the EMD modes into two signals to construct the "lock" and "closure" sounds. These reconstructions of the "lock" and "closure" sounds were carried out by partial recombination of modes 1 to 6, and of modes 6 to the last mode, respectively. Figure 2 illustrates this process applied to the sound s4. Figure 3 clearly shows the interest of the EMD method, which enables a neat source separation that conserves important perceptual cues, linked to both their time and frequency supports. Sound s20 presents a much weaker "lock" than "closure" component. This could easily be perceived by ear. In contrast, sound s26 presents a much stronger "lock" than "closure" component. Sound s4 presents a balance of both.

The 26 door-closure sounds of our study were, thus, systematically decomposed into a "lock" sound and a "closure"

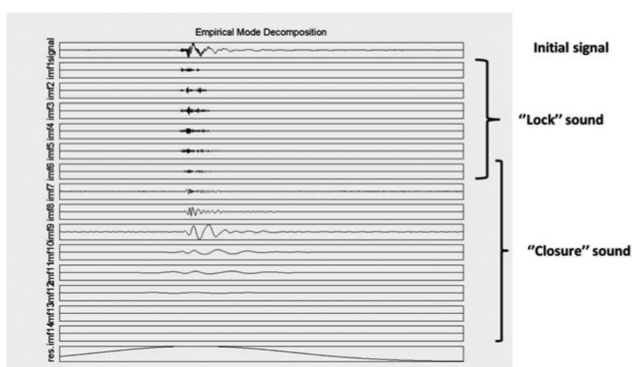

FIG. 2. Example of modes obtained by the Empirical Mode Decomposition (EMD) method applied to the sound s4. The "lock" contribution of this sound is obtained by adding modes 1 to 6 and the "closure" contribution by addition modes 6 to the last mode.

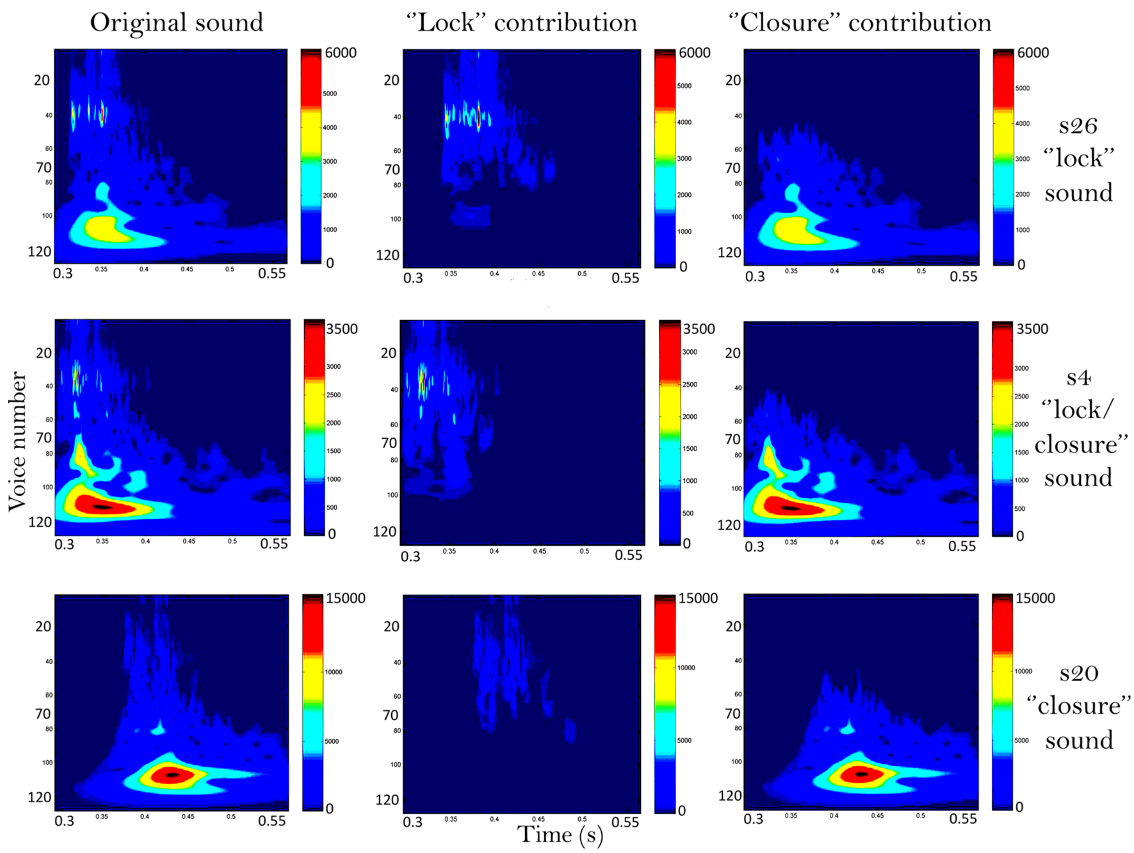

FIG. 3. (Color online) Wavelet transform of an original door-closure sound (1st column), the "lock" contribution (2nd column) and the "closure" contribution (3rd column) of the sounds s26 (1st row), s4 (2nd row) and s20 (3rd row). The sound s26 is characterized by the listeners as a "Lock" sound, s4 as a "Lock/closure" sound while s20 is characterized as a "Closure" sound.

sound. The success of the decompositions were validated by expert listening to the "door," "lock," and "closure" signals.

B. Acoustic analysis of lock and closure contributions

The 26 door-closure signals of the perceptual study were, to begin with, decomposed into "lock" sounds and "closure" sounds. The "lock" sound is composed mainly of three rapidly damped impact sounds while the "closure" sound is composed of a lower frequency impact with a slower attack and damped portion.

1. Time-ERB representation

The "lock" and "closure" components of the s4 door-closure sound were analyzed by a time-ERB representation (Patterson *et al.*, 1995). The impact signals were analyzed using time-scale representations from which the initial spectral distributions and the damping laws were extracted. The time-scale representation used was taken from (Aramaki and Kronland-Martinet, 2006) and takes into account basilar membrane filtering. The 41 analysis windows were centered at each ERB interval, and their bandwidth was defined such that two successive windows overlap at the point where the maximum amplitude decreases by a factor of $\sqrt{2}$. Gaussian analysis windows were selected as they have the advantage of minimizing the Gabor-Heisenberg uncertainty principle. Then, the main impacts were identified. Three main impacts identified from the largest values of high frequency energy densities ($ERB > 20$; $f > 1800$ Hz) as a function of time created the "lock" sound. The selected peaks had to be separated by at least 13 ms so that the impacts would be perceptually salient and would not be affected by temporal masking. The single impact of the "closure" component was isolated as the largest value of high frequency energy density ($ERB > 3$; $f > 100$ Hz) as a function of time. In our example (Fig. 4), the impacts of the "lock" component are separated by 17.4 and 21.7 ms, while the impact of the "closure" component appears just after the second "lock" component impact. The impact of the "closure" sound and the second impact of the "lock" sound may have similar physical origins, but they are dealt with independently in the analysis.

2. Acoustic descriptors

The levels, for each ERB band, of the three impacts of the "lock" sound were calculated over a duration of 4.5 ms, while those of the lower frequency "closure" impacts were calculated over a duration of 13.6 ms. The damping coefficients, for each ERB band, of the last "lock" impact and the "closure" impact were also calculated. As the first "lock" impacts were not temporally distinguishable, their damping could not be measured easily and were masked by the following impacts. The damping coefficients of these impacts were modeled by constants indexed by the inter-impact durations. To distinguish these impact sounds, a damping coefficient (a) was chosen so that the signal amplitude was reduced by 20 dB with the arrival of the following impact:

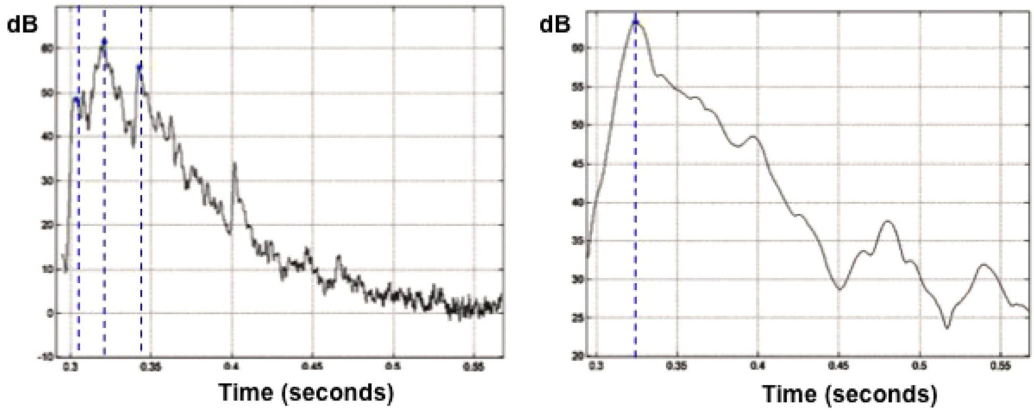

FIG. 4. (Color online) Impacts of the "lock" contribution (left) and the "closure" contribution (right) for the sound s4. The maximum values of the 3 impacts of the "lock" component and the impact of the "closure" component are indicated by vertical, dotted lines.

$$a = \ln(10)/d, \quad (5)$$

where d is the inter-impact duration.

Each door-closure sound was decomposed into four impacts, i.e., three impacts for the "lock" component and one impact for the "closure" component. The three impacts of the "lock" component were characterized by their temporal position (Tclic), the level (L) of the initial spectral distribution, the spectral spread (Spread), the sub-band of the maximum energy (Bmax), and the damping law of the last impact sound for which we calculated the average band and the bandwidth (Bandmean, Bandnum). The impact of the "closure" component was characterized by the temporal position (Tclose), the level (Lclose) of the initial spectral distribution, the spectral centroid (SCclose), and the damping law approximated by an exponential function (two parameters: A and B). An overview of the acoustic descriptors is given in Table IV.

C. Relation between perceptual properties and signal parameters

The aim of our study was to predict the perceptual properties on the basis of linear combinations of the acoustic descriptors. For each perceptual property, descriptors were selected with the help of step-by-step regressions and linear relations were established by Ridge regression with the "leave-one-out" cross-validation criterion (Matzner-Løber, 2006). The Ridge regression offered a solution to the problem of the colinearity of the acoustic descriptors by allowing a bias in the estimation of coefficients to reduce variance. The cross-validation ensured that the models conserved their robustness when dealing with new stimuli; it consisted in determining different models by learning for all n-1 datasets and in calculating the coefficient of determination from these predictions.

The relationship between the natural properties, the evoked associations and the acoustic descriptors were

TABLE IV. Definition of acoustic descriptors of door-closure sounds.

Acoustic Descriptors	Definition
Lclic	Mean level of the 3 "lock" sound impacts
L1	Initial level of the 1st "lock" sound impact
Bmax1	ERB sub-band of the peak energy of the 1st "lock" sound impact
Spread1	Spectral spread of the initial spectral distribution of 1st "lock" sound impact
L2	Initial level of the 2nd "lock" sound impact.
Bmax2	ERB sub-band of peak energy of 2nd "lock" sound impact.
Spread2	Spectral spread of the initial spectral distribution of 2nd "lock" sound impact
L3	Initial level of the 3rd "lock" sound impact.
Bmax3	ERB sub-band of the peak energy of the 3rd "lock" sound impact
Spread3	Spectral spread of the initial spectral distribution of 3rd "lock" sound impact
Bandmean	Mean ERB sub-band of minimum damping of 3rd "lock" sound impact
Bandnum	Number of ERB sub-bands of minimum damping of 3rd "lock" sound impact
Lclose	Initial level of the "closure" sound impact
SCclose	Initial spectral centroid of the "closure" sound impact
Slope	Decay coefficient of the initial spectral distribution of the "closure" sound impact
AB	Approximation function coefficient of the damping law of the "closure" sound impact. $= Aexp(B \times ERB)$
Tclic3-Tclic2	"Inter-impact" duration between the 3rd and the 2nd "lock sound impact
Tclic2-Tclic1	"Inter-impact" durations between the 2nd and the 1st "lock sound impact
Tclic1-Tclose	"Inter-impact" durations between the 1st "lock sound impact and the "closure" sound impact

TABLE V. Prediction models of perceptual properties and associated R2val (coefficient of determination obtained by cross validation).

Estimated property	Prediction models	R2val
INTENSE	$0.28 Lclose + 0.81 SCclose - 25$	0.84
BON'M	$-0.42(L3 - Lclose) + 0.14 L1 - 10$	0.72
KE	$0.48(Lclic - Lclose) - 0.04(Tclic3 - Tclose) - 0.09(Tclic3 - Tclic2) + 10$	0.78
Door weight	$-0.26(L3 - Lclose) - 0.12 Bandmean + 0.80 Slope + 5$	0.83
Energy of closure	$0.32 Lclose + 0.68 SCclose - 27$	0.87
Solidity	$-(L3 - Lclose) - L2 + Slope$	0.62
Quality	$-0.40 Lclic + 0.41 L1 + 21$	0.75

established for all the 26 sounds studied, with the exception of sounds s12 and s21 that featured strong vibrations (s21) and had widely spaced lock notches (s12). These specificities were excluded from our study. The relationships obtained and the associated coefficients of determination (R2val) are summarized below (Table V).

The perceptual properties were estimated from the acoustic descriptors L1, L2, L3, Lclic. Bandmean, Lclose, SCclose, Slope, Tclic3-Tclose, Tclic3-Tclic2. The quality of our estimations (predicted vs measured data) is illustrated in Fig. 5 (for analytical properties), and Fig. 6 (for natural properties and evoked associations).

VI. NETWORK

The acoustic descriptors selected characterize the network of perceptual properties of the door-closure sounds as follows:

(1) The related properties, INTENSE and "Energy of closure" were both estimated using the initial closure level (Lclose) and the initial spectral centroid of the impact of the "closure" component (SC close). The more forcefully the door is closed, the more significant the contribution of the "closure" component compared to the "lock" component.

(2) The KE property emerges when the "lock" component (Lclick) increases in significance with respect to the "closure" component (Lclose) and the inter-impact durations (Tclick3-Tclose, Tclick3-Tclick2). The greater the temporal distance between the impacts, the more the sound is evaluated in terms of KE. This is true particularly in the case of the last impact of the "lock" component and the impact of the "closure" component.

(3) The importance of the BON'M property contribution is inversely related to the level of the last impact of the "lock" component (L3) and directly related to the level of the first impact (L1) and the initial level of the closure component (Lclose). Thus, the BON'M property emerges when we hear a single impact associated with the low frequency resonance.

(4) The "Door weight" property is estimated from the level of the last impact of the "lock" component (L3) relative to that of the "closure" component (Lclose). As for the BON'M property, the Lclose property has to be weak to give the impression of a heavy door. It is also dependent on the frequencies related to the minimum damping coefficient of the last "lock" impact (Bandmean); the lower the frequencies, the greater the impression of a heavy door. Finally, this property is linked to the decay coefficient of the initial spectral spread of the "closure"

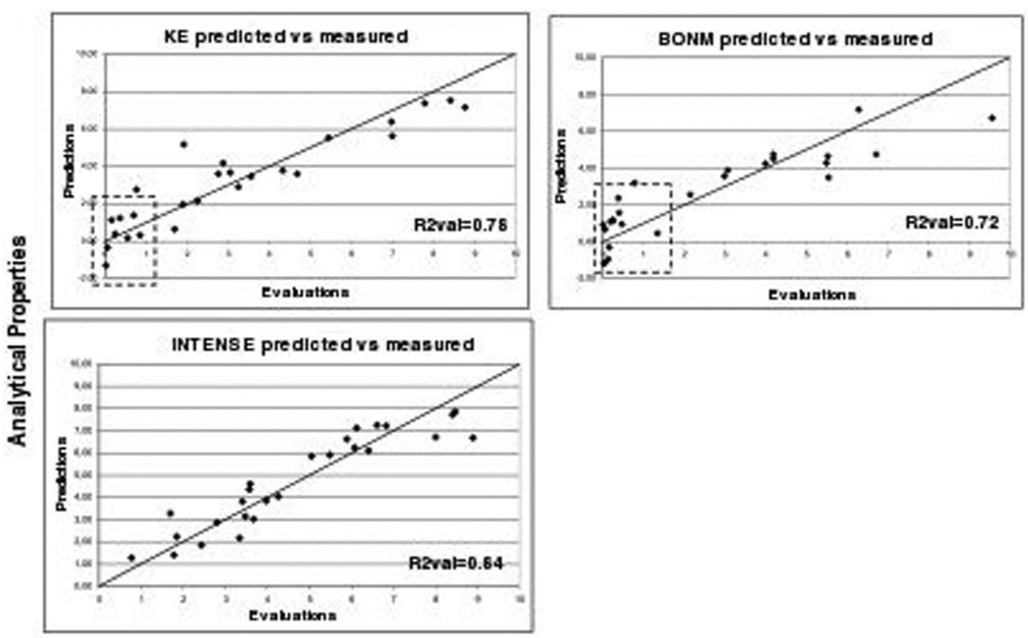

FIG. 5. Predicted versus measured analytical properties for the descriptors KE, BONM and INTENSE.

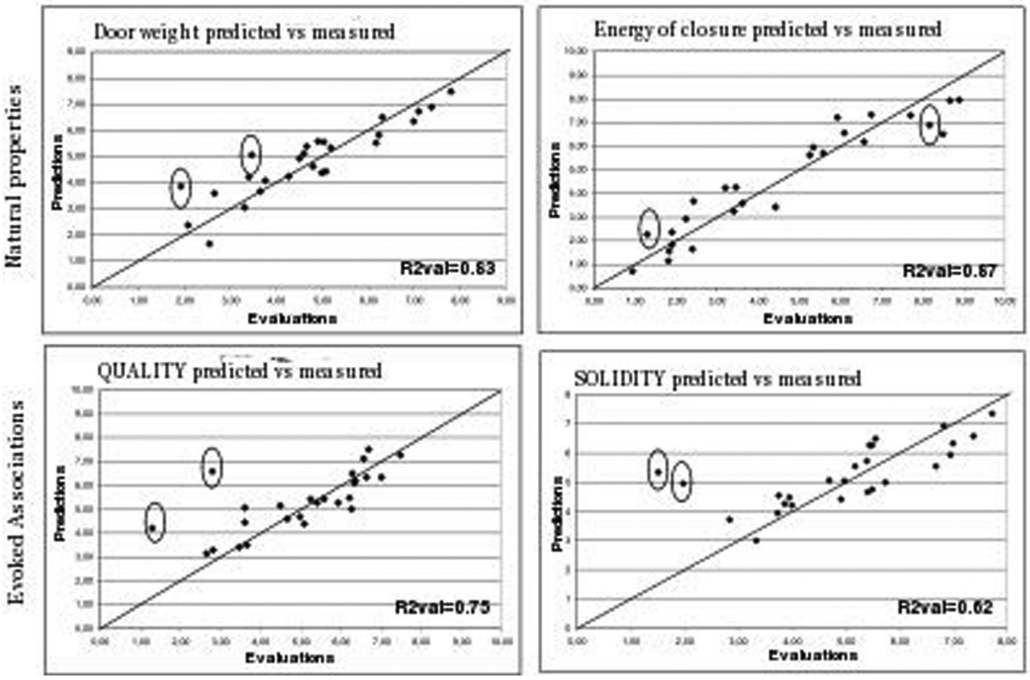

FIG. 6. Predicted versus measured natural properties and evoked associations. The data marked with circles correspond to sounds s12 and s21 that strongly differ from the other sounds by their marked vibrations (s21) and widely spaced lock notches (s12) and are therefore not taken into account in the models.

component (Slope). The lower the slope value, the greater the number of high frequencies in the "closure" component and the greater the impression of a heavy door. Although not very intuitive, this result can be explained by the fact that door sounds characterized by high "energy of closure" and with a lower slope value generally feature BON'M to a greater degree than sounds of low "energy of closure." Thus, they are associated with the sound of a heavy door.

(5) The evoked impression of "solidity" is associated with the level of the last impact of the "lock" component (L3) relative to the level of the "closure" component (Lclose), as for the properties, BON'M and "Door weight," and to the slope of the initial spectral distribution of the "closure" component (Slope), as for the property "Door weight." The "solidity" is also inversely related to the level of the second "lock" impact (L2); the weaker this impact, the more the sound evokes an impression of solidity.

(6) Finally, the evoked impression of "quality" is negatively related to mean impact level of the three "lock" components (Lclick) and positively so to the level of the first impact of the "lock" component (L1). Therefore, the higher the level of the first impact, the more the sound evokes an impression of quality.

These results enabled to link acoustic descriptors and perceptual properties as shown in Fig. 7

VII. CONCLUSION AND DISCUSSION

The aim of this work was to propose a predicative model of the relationship between the perceived car quality and solidity induced by door-closure sounds and acoustic

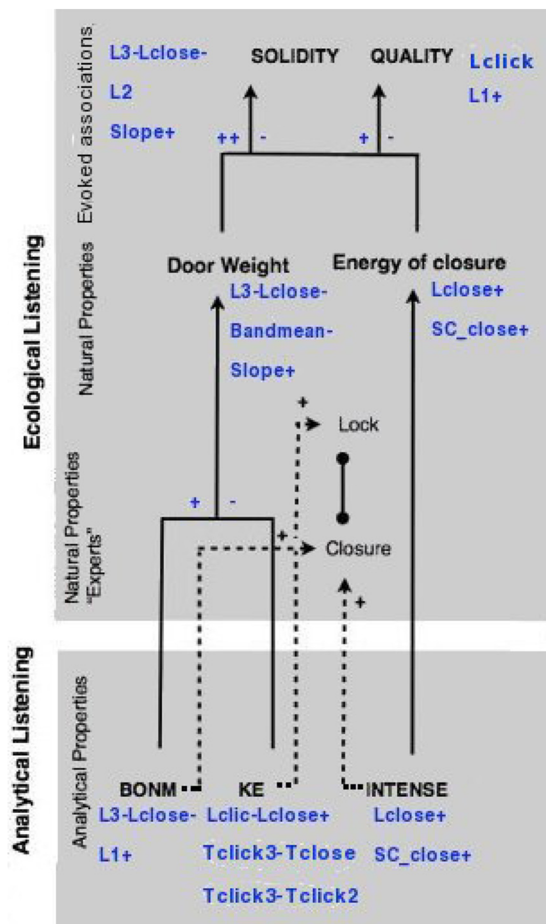

FIG. 7. (Color online) Diagram of perceptual properties and acoustic correlates.

descriptors. Perceptual judgments of real sounds obtained from recordings of car doors (different brands and segments) related to high-and low-level perceptual attributes were therefore investigated as well as the corresponding acoustical and mechanical properties.

The perceptual judgment of car doors were obtained from several listening tests. Following Gaver's definition (Gaver, 1993a), both analytical and ecological listening was considered.

In the case of analytical listening, the sound signal is described without reference to the event. Even if analytical listening is generally considered artificial and does not play a role in the perceptual process as assumed by direct perception theory, it is used here as a help to reveal perceptually relevant signal morphologies. To incite subjects to characterize sounds in this way, a sensory analysis method was used. This approach revealed that the door-closure sound is described mainly by the intensity and by the onomatopoeia, BONM and KE, hereby reflecting the perceptual importance of a low-frequency contribution (BONM) and a high-frequency contribution (KE) in the car door signal.

In the case of ecological listening, the event associated with the sound characterized by a set of natural properties as well as the evoked associations is described. The ecological description of a door-closure sound depends on the listener's need to understand the world with which he or she interacts; the "heavy" door is not a satisfactory description for a physicist, but it is well understood by the individual who is accustomed to closing car doors and who, therefore, infers its associated action from the sound heard (mirror neurons). Ecological listening, therefore, depends on the listener's expertise. Consequently, listening tests with both naive and expert listeners were performed. The evaluations obtained from this kind of listening strongly depends on context-related factors: A comparative study of evaluations made in real and laboratory settings showed a strong role of a vehicle's image in evaluations of its quality based on the door-closure sound (Bezat *et al.*, 2006; Bezat, 2007). The laboratory setting used in the present study (where only auditory stimuli were presented) enabled the listeners to focus on ecological properties in a more "impoverished," controlled setting and, thus, to achieve more precise quantifications. The listening tests revealed that naive listeners were able to discriminate the doors by their quality, solidity, energy of closure, door weight, and door-closure effectiveness in a coherent manner. This is in line with previous studies by Kuwano *et al*. (2006) that resulted in coherent quality evaluations of car door sounds across subjects from semantic differential tests based on a predefined adjective scale. The expert listeners identified the organic sources of the door-closure sound (i.e., lock and closure contributions) in contrast to the more macroscopic evaluations of the naive listeners.

These different listening tests enabled to establish a network of perceptual properties of door-closure sounds (Fig. 7). It was hereby shown that the impression of solidity and quality was linked to the sensation of a heavy door and a gentle gesture, which in turn was characterized by the sensory panel as sounds without flaws (ex. vibrations), low-pitched with little lock presence (strong BONM and weak KE) and of low intensity. The influence of the weight of the car door on the perceived quality was also confirmed by Scholl and Amman (1999) in a study where car door noises were evaluated after physical modifications of the car door sources.

To further associate acoustic properties to the perceptual evaluations, the perceived analytic properties as well as the expert evaluations of the car door sounds brought important cues to the identification of perceptually relevant signal parameters. As already mentioned, the analytical evaluations revealed the presence of high and low-frequency impacts. This was coherent with the expert evaluations that focused on the presence of the lock (high-frequency) and closure (low-frequency) contribution, which suggests that the sensory panel participants unconsciously identified the lock and closure contribution, respectively, by the KE and the BONM onomatopoeias. These two aspects are not easy to distinguish from traditional signal analysis, since they have overlapping frequency contents and occur almost simultaneously. An empirical mode decomposition method (EMD) turned out to be efficient to separate the lock and closure contributions in a satisfactory manner from a perceptual point of view. The principle behind this method is to iteratively identify signal modes modulated both in amplitude and frequency by locally separating a rapid contribution from a slower one. Hence, the closure sound is obtained from the six first modes and the lock sound from higher-order modes.

Intuitively, analogies between the method and the functioning of the basilar membrane can be drawn. Moore and Glasberg (1983) suggested describing the patterns of excitation at the level of the basilar membrane as the output of a bank of auditory filters whose bandwidths (critical bands) increase as a function of the center-frequency constant $\Delta f/f$. The filtering carried out by EMD uses a constant $\Delta f/f$ (Flandrin *et al.*, 2004), and its procedure proves "simpler" than existing models of basilar membrane filtering. In addition, like the basilar membrane, this filtering is adaptive; it can adapt itself to the principal components of a signal. Finally, it allows nonlinear decompositions.

Further investigations on the lock and closure parts of the signal (separated by the EMD decomposition) revealed that the lock contribution is mainly composed of three rapidly damped impacts, while the closure sound is composed of one low-frequency impact. The energy level, temporal positions, spectral content and damping laws were considered as main acoustic descriptors for these signals and were related to the perceptual properties by a Ridge regression method. The perception of quality and solidity could hereby be linked to the relative level of certain impacts and the decay of the closure component. The perceptual importance of the decay of impact sounds is in line with previous "perceptual" models of impact sounds (Aramaki *et al.*, 2010b; Aramaki *et al.*, 2010a; Aramaki *et al.*, 2011; Giordano and McAdams, 2006; McAdams *et al.*, 2010).

The statistical relationship between perceptual properties and acoustic descriptors allow those working in industry to carry out a diagnostic of a given door-closure sound and, in particular, to position it on a scale of perceived quality.

These relations convey the optimal balance of acoustic descriptors obtained from a corpus of sounds, but they do not describe the causal relations between the descriptors and the perceptual properties. In particular, they do not account for the specific effects of acoustic descriptors for a given sound and do not contribute to specifying the changes required to improve a certain door-closure sound. Rather than wanting to just predict the perceptual properties, our aim is to control them using the acoustic signal. Therefore, an analysis-synthesis approach is envisaged. Note that a perfect resynthesis will not be needed in this case, since the aim is to identify the principal perceptual properties characteristic of these sounds. According to Liénard (2002), "the higher level description of a signal is […] complete if it is capable of reconstructing, by synthesis, a signal that is perceptually equivalent to the initial signal." Carrying out perceptual tests on these controlled synthesized sounds would allow us to study, in a systematic fashion, the relations between the acoustic signal and the perceptual properties. This aspect will be presented in a companion article.

ACKNOWLEDGMENTS

We would like to thank Mitsuko Aramaki for her precious suggestions related to the manuscript.

Aramaki, M., Baillères, H., Brancheriau, L., Kronland-Martinet, R., and Ystad, S. (**2007**). "Sound quality assessment of wood for xylophone bars," J. Acoust. Soc. Am. **121**(4), 2407–2420.

Aramaki, M., Besson, M., Kronland-Martinet, R., and Ystad, S. (**2011**). "Controlling the perceived material in an impact sound synthesizer," IEEE Trans. Audio Speech Language Processing **19**(2), 301–314.

Aramaki, M., Gondre, C., Kronland-Martinet, R., Voinier, T., and Ystad, S. (**2010a**). "Imagine the sounds: An intuitive control of an impact sound synthesizer," in *Auditory Display, Lecture Notes in Computer Science*, edited by S. Ystad, M. Aramaki, R. Kronland-Martinet, and K. Jensen (Springer-Verlag, Berlin), Vol. 5954, pp. 408–421.

Aramaki, M., and Kronland-Martinet, R. (**2006**). "Analysis-synthesis of impact sounds by real-time dynamic filtering," IEEE Trans. Audio Speech Language Processing **14**(2), 695–705.

Aramaki, M., Marie, C., Kronland-Martinet, R., Ystad, S., and Besson, M. (**2010b**). "Sound categorization and conceptual priming for nonlinguistic and linguistic sounds," J. Cognit. Neurosci. **22**(11), 2555–2569.

Bezat, M. (**2007**). "Perception des bruits d'impact: Application au bruit de fermeture de porte automobile" ("Perception of impact noises: Application to car-door closure noise"), Ph.D. dissertation, Université Aix-Marseille I, Provence, France.

Bezat, M.-C., Roussarie, V., Kronland-Martinet, R., Ystad, S., and McAdams, S. (**2006**). "Perceptual analyses of action-related impact sounds," in *Euronoise 2006* (Tampere, Finland).

Blommer, M., Yang, B., and Vandenbrink, K. (**2005**). "Detection and classifying secondary impacts in door closing sounds," in *Proceedings of the 2005 Noise and Vibration Conference, Society of Automotive Engineers* (Traverse City, MI), No. 2005-01-2471.

Champagne, A., and Amman, S. (**1995**). "Vehicle closure sound quality," SAE Technical Paper 951370.

Flandrin, P., and Gonçalves, P. (**2004**). "Empirical mode decompositions as data-driven wavelet-like expansions," Int. J. Wavelets Multiresolution Inf. Processing **2**(4), 477–496.

Flandrin, P., Rilling, G., and Gonçalves, P. (**2004**). "Empirical mode decomposition as a filter bank," IEEE Signal Processing Lett. **11**(2), 112–114.

Freed, D. J. (**1990**). "Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events," J. Acoust. Soc. Am. **87**(1), 311–322.

Fridrich, J. R. (**1999**). "Percentile frequency method for evaluating impulsive sounds," in *Proceedings of the 1999 Noise and Vibration Conference, Society of Automotive Engineers* (Traverse City, MI), No. 1999-01-1851.

Fridrich, J. R. (**2007**). "Measuring the pitch of door closing sounds the sound quality issue of door 'thump,'" in *Proceedings of the 2007 Noise and Vibration Conference, Society of Automotive Engineers* (Traverse City, MI), No. 2007-01-2216.

Gaver, W. W. (**1993a**). "How do we hear in the world? Explorations in ecological acoustics," Ecol. Psychol. **5**(4), 285–313.

Gaver, W. W. (**1993b**). "What in the world do we hear? An ecological approach to auditory event perception," Ecol. Psychol. **5**(1), 1–29.

Giordano, B. L., and McAdams, S. (**2006**). "Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates," J. Acoust. Soc. Am. **119**(2), 1171–1181.

Hamilton, D. (**1999**). "Sound quality of impulsive noises: An applied study if automotive door closing sounds," in *Proceedings of the 1999 Noise and Vibration Conference, Society of Automotive Engineers* (Traverse City, MI), No. 1999-01-1684.

Hojberg, K. (**1991**). "Measuring the loudness of door slams," in *Proceedings of the 1991 Noise and Vibration Conference, Society of Automotive Engineers* (Traverse City, MI), No. 911090.

Huang, N. E., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Yen, N., Tung, C., and Liu, H. H. (**1998**). "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis," Proc. R. Soc. A London, **454**(1971), 903–995.

Klatzky, R. L., and Pai, D. K. (**2000**). "Perception of material from contact sounds," Presence: Teleoperators Virtual Environ. **9**(4), 399–410.

Kuwano, S., Fastl, H., Namba, S., Nakamura, S., and Uchida, H. (**2006**). "Quality of door sounds of passenger cars," Acoust. Sci. Technol. **27**(5), 309–312.

Lakatos, S., McAdams, S., and Chaigne, A. (**1997**). "The representation of auditory source characteristics: Simple geometric form," Percept. Psychophys. **59**, 1180–1190.

Lemaitre, G., Dessein, A., Susini, P., and Aura, K. (**2011**). "Vocal imitations and the identification of sound events," Ecol. Psychol. **4**(23), 267–307.

Liénard, J. (**2002**). *Variabilité et multicatégorisation (Variability and Multicategorization)* Invariants et variabilité dans les sciences cognitives (Invariants and variabilities in cognitive science) (Presses de la maison des sciences de l'homme, Paris), 298 p.

Lutfi, R. A., and Oh, E. L. (**1997**). "Auditory discrimination of material changes in a struckclamped bar," J. Acoust. Soc. Am. **102**(6), 3647–3656.

Malen, D. E., and Hancock, M. (**1995**). "Engineering for the customer: Combining preference and physical system models. Part II application," J. Eng. Design **6**(4), 329–341.

Matzner-Løber, É. (**2006**). *Régression: Théorie et applications et applications (Regression: Theory and Applications)* Statistique et Probabilités Appliquées (Statistics and Applied Probabilities) (Springer Verlag, Paris), 317 p.

McAdams, S., and Bigand, E. (**1993**). *Thinking in Sound: The Cognitive Psychology of Human Audition* (Oxford University Press, New York), 368 p.

McAdams, S., Chaigne, A., and Roussarie, V. (**2004**). "Psychomechanics of simulated sound sources," J. Acoust. Soc. Am. **115**(3), 1306–1320.

McAdams, S., Roussarie, V., Chaigne, A., and Giordano, B. L. (**2010**). "The psychomechanics of simulated sound sources: Material properties of impacted thin plates," J. Acoust. Soc. Am. **128**(3), 1401–1413.

Merer, A., Aramaki, M., Ystad, S., and Kronland-Martinet, R. (**2013**). "Perceptual characterization of motion evoked by sounds for synthesis control purposes," ACM Trans. Appl. Percept. **10**(1), 1–24.

Merer, A., Ystad, S., Aramaki, M., and Kronland-Martinet, R. (**2011**), *Abstract Sounds and Their Applications in Audio and Perception Research* (Springer-Verlag, Berlin), pp. 269–297.

Micoulaud-Franchi, J. A., Aramaki, M., Merer, A., Cermolacce, M., Ystad, S., Kronland-Martinet, R., Naudin, J., and Vion-Dury, J. (**2012**). "Toward an exploration of feeling of strangeness in schizophrenia: perspectives on acousmatic and everyday listening," J. Abnormal Psychol. **131**(3), 628–640.

Moore, B. C. J., and Glasberg, B. R. (**1983**). "Suggested formulae for calculating auditory filter bandwidths and excitation patterns," J. Acoust. Soc. Am. **74**(3), 750–753.

Parizet, E., Guyader, E., and Nosulenko, V. (**2008**). "Analysis of car door closing sound quality," Appl. Acoust. **69**(1), 12–22.

Patterson, R. D., Allerhand, M. H., and Giguere, C. (**1995**). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," J. Acoust. Soc. Am. **98**(4), 1890–1894.

Petniunas, A., Otto, N. C., Amman, S., and Simpson, R. (**1999**). "Door system design for improved closure sound quality," in *Proceedings of the SAE Noise and Vibration Conference and Exposition* (Traverse City, MI).

Poirson, E. (**2006**). "Prise en compte des perceptions de l'utilisateur en conception de produit. Application aux instruments de musique de type cuivre" ("Taking into account the perception of the user when conceiving a product. Applications on brass-like instruments") Ph.D. thesis, Institut de Recherche en Communications et en Cybernétique, Nantes, France.

Repp, B. H. (**1987**). "The sound of two hands clapping: An exploratory study," J. Acoust. Soc. Am. **81**(4), 1100–1109.

Roussarie, V., Richard, F., and Bezat, M.-C. (**2004**). "Validation of auditory attributes using analysis synthesis method," in *Congráes Francais d'Acoustique/DAGA* (Strasbourg, Germany).

Schaeffer, P. (**1966**). *Traité des Objets Musicaux (Treaty of Musical Objects)* (Seuil, Paris VI), 672 p.

Scholl, D., and Amman, S. (**1999**). "A new wavelet technique for transient sound visualization and application to automotive door closing events," in *Proceedings of the SAE Noise and Vibration Conference and Exposition* (Traverse City, MI).

Schön, D., Kronland-Martinet, R., Ystad, S., and Besson, M. (**2010**). "The evocative power of sounds: Conceptual priming between words and nonverbal sounds," J. Cognit. Neurosci. **22**(5), 1026–1035.

Sciabica, J., Bezat, M., Roussarie, V., Kronland-Martinet, R., and Ystad, S. (**2009**). "Towards the timbre modeling of interior car sound," in *15th International Conference on Auditory Display* (Copenhagen, Denmark).

Shin, T., Park, D., and Lee, S. (**2013**). "Objective evaluation of door-closing sound quality based on physiological acoustics," Int. J. Automotive Technol. **14**, 133–141.

Stautner, J. P. (**1983**). "Analysis and synthesis of music using the auditory transform," M.S. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Thoret, E., Aramaki, M., Kronland-Martinet, R., Velay, J.-L., and Ystad, S. (**2014**). "From sound to shape: Auditory perception of drawing movements," J. Exp. Psychol. **40**, 983–994.

Traube, C. and D'Alessandro, N. (**2005**). "Synthèse de l'imitation vocale et representation graphique des gestes phonétiques sous-jacents à la description du timbre de la guitare classique" ("Synthesis of voice imitation and graphical representations of phonetic gestures qualifying timbre of classical guitars"), in *Proceedings of the Conference on Interdisciplinarity Musicology (CIM05)* (Montreal, Quebec).

Traube, C., and Depalle, P. (**2004**). "Timbral analogies between vowels and plucked strings tones," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP04)* (Montreal, Quebec), vol. 4, pp. 293–296.

Van den Doel, K., and Pai, D. K. (**1996**). "The sounds of physical shapes," Presence **7**(4), 382–395.

Warren, W. H., and Verbrugge, R. R. (**1984**). "Auditory perception of breaking and bouncing events: A case study in ecological acoustics," J. Exp. Psychol. **10**(5), 704–712.

Wildes, R. P., and Richards, W. A. (**1988**). *Recovering Material Properties From Sound* (MIT Press, Cambridge, MA), Chap. 25, pp. 356–363.