



Contents lists available at ScienceDirect

## Int. J. Human-Computer Studies

journal homepage: [www.elsevier.com/locate/ijhcs](http://www.elsevier.com/locate/ijhcs)Categorization of seismic sources by auditory display: A blind test<sup>☆</sup>Arthur Paté<sup>a,b,\*</sup>, Lapo Boschi<sup>c,d</sup>, Jean-Loïc Le Carrou<sup>a,b</sup>, Benjamin Holtzman<sup>e</sup><sup>a</sup> Sorbonne Universités, UPMC Univ Paris 06, UMR 7190, Institut Jean Le Rond d'Alembert, équipe LAM, 11, rue de Lourmel, F-75015 Paris, France<sup>b</sup> CNRS UMR 7190, Institut Jean Le Rond d'Alembert, équipe LAM, 11, rue de Lourmel, F-75015 Paris, France<sup>c</sup> Sorbonne Universités, UPMC Univ Paris 06, UMR 7193, Institut des Sciences de la Terre Paris (iSTeP), F-75005 Paris, France<sup>d</sup> CNRS, UMR 7193, Institut des Sciences de la Terre Paris (iSTeP), F-75005 Paris, France<sup>e</sup> Lamont Doherty Earth Observatory, Columbia University, New York, USA

## ARTICLE INFO

## Article history:

Received 22 January 2015

Received in revised form

23 June 2015

Accepted 13 August 2015

Available online 2 September 2015

## Keywords:

Seismology

Sonification

Categorization

Perception

Psychoacoustics

Acoustics

Geology

Wave

Auditory

## ABSTRACT

Recordings of the Earth's surface oscillation as a function of time (seismograms) can be sonified by compressing time so that most of the signal's frequency spectrum falls in the audible range. The pattern-recognition capabilities of the human auditory system can then be applied to the auditory analysis of seismic data. In this experiment, we sonify a set of seismograms associated with a magnitude-5.6 Oklahoma earthquake recorded at 17 broadband stations within a radius of ~300 km from the epicenter, and a group of volunteers listen to our sonified seismic data set via headphones. Most of the subjects have never heard a sonified seismogram before. Given the lack of studies on this subject, we prefer to make no preliminary hypotheses on the categorization criteria employed by the listeners: we follow the "free categorization" approach, asking listeners to simply group sounds that they perceive as "similar." We find that listeners tend to group together sonified seismograms sharing one or more underlying physical parameters, including source–receiver distance, source–receiver azimuth, and, possibly, crustal structure between source and receiver and/or at the receiver. This suggests that, if trained to do so, human listeners can recognize subtle features in sonified seismic signals. It remains to be determined whether auditory analysis can complement or lead to improvements upon the standard visual and computational approaches in specific tasks of geophysical interest.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Seismologist Hugo Benioff first implemented a technique to accelerate seismograms to the range of audible frequency, compiling a set of sonified seismograms that was commercially released in 1953 in the form of an LP album (Karney, 2015). It was then suggested that earthquakes could be discriminated from man-made explosions by simply listening to the associated sonified time series, exploiting the high resolving power of the human auditory system (Speeth, 1961; Frantti and Levereault, 1965). This proposed approach was never put into practice: with the advent of digital seismology in the 1970s, automated software could accurately estimate hypocenter locations and source mechanisms by processing large seismic databases (Dziwonski et al., 1981; Ekström et al., 2012).

In principle, auditory analysis could contribute to current research topics in seismology. Humans have a powerful facility to understand physical characteristics of a process through sound,

such as the mechanical nature of an impact. We assess the materials involved (metal, wood, glass, plastic) and the magnitude of forces involved. The question for this study is: Can we use this ability to characterize physical aspects of an unknown process, such as an earthquake? And for future work, can this ability be trained? How do we build this capacity, and then integrate it into practical analysis?

Research on the nature of earthquake rupture could also benefit from auditory display. As an example, the *Source Inversion Validation initiative* (Mai et al., 2012) was a blind test of multiple methods of dynamic and kinematic inversion of seismic observations (Ruiz and Madariaga, 2013) to produce a map of an earthquake rupture (the displacement along a fault during a single seismic event). Results show that the current signal processing techniques do not lead to robust models of earthquake rupture (the displacement along fault during a single seismic event). It is worthwhile to explore whether auditory analysis can help discriminating signals originating from different types of earthquake rupture. Study of human auditory analysis could lead to improvements in the signal processing algorithms for analysis of seismograms. Such improvements could also aid methods of engineering and real time control of fluid pumping in

<sup>☆</sup>This paper has been recommended for acceptance by E. Motta.

\* Corresponding author.

E-mail address: [pate.arthur.lam@gmail.com](mailto:pate.arthur.lam@gmail.com) (A. Paté).

reservoirs, for hydrocarbons extraction, carbon sequestration and geothermal energy systems. Other potential applications are discussed at the end of the paper.

Over the years, a small community of researchers has continued to sonify seismic data for a number of (often educational or artistic) applications (Steinbrugge, 1974; Hayward, 1994; Dombois, 2001, 2002; Meier and Saranti, 2008). However, even though interest around sonification seems now to be growing in seismology (Hermann et al., 2011; Dombois and Eckel, 2011; Michael, 2011; Kilb et al., 2012; Peng et al., 2012; Holtzman et al., 2014) as well as other disciplines (Cowen, 2015; Worrall, 2009), the capability of the human auditory system (Benade, 1990; Hartmann, 1999; Roederer, 2008) to recognize patterns in seismic sound has not been studied quantitatively. No study so far has dealt with the discrimination of sonified seismic signals by human listeners, or, more generally, with our strategies (if any) of hearing, listening to, recognize, organize, or process such signals. The unique experiment of (Speeth, 1961) explored the human ability to distinguish sonified records of explosions vs. seismic events: a relatively simple, and very specific task.

In the experiment presented here, we proposed the listeners to categorize freely a set of sonified seismic data. As explained in Section 3, no information on the nature of such data (other than the fact that they were recordings of earthquakes) was provided, and the only criterion for grouping the data was their perceived “similarity.” Since all signals were generated by the same seismic event, we expected listeners to discriminate based on source–receiver distance, source–receiver azimuth and/or crustal structure between the source and the receiver. In the free-categorization approach, however, no specific hypothesis is tested directly, and it is *a priori* possible for a listener to group data according to a valid criterion not anticipated by the researchers.

Individual audio signals used in this study are produced by simple time-compression of seismic signals and are administered to listeners monophonically (the same signal is played through the two channels of the headphones, in phase) one at a time. This deliberately simple study is a first step towards the auditory analysis of spatialized seismic

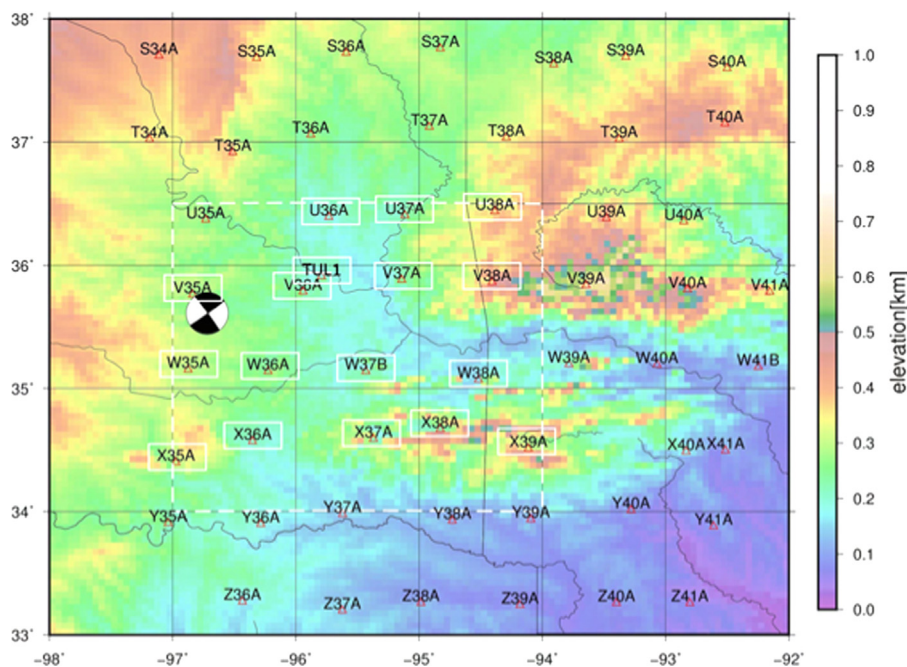
data; preliminary experiments in the spatialization of seismic sounds are described by Holtzman et al. (2014).

## 2. Seismic signals

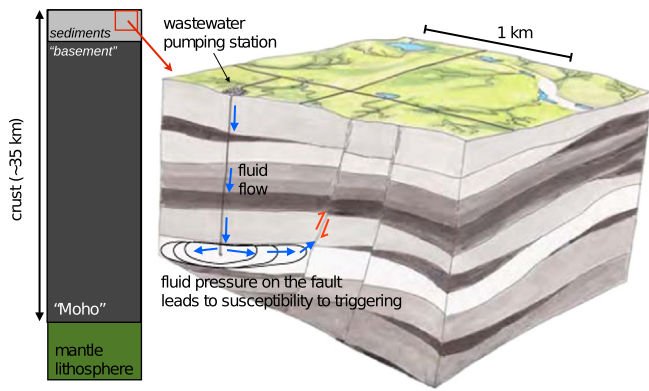
### 2.1. Brief geology and seismology overview

Our newly compiled database of sonified seismograms is based on records of a recent sequence of 40 Oklahoma earthquakes of magnitude ranging between 3 and 5, recorded by 17 stations at local epicentral distances (Fig. 1). All data were collected in the framework of the USArray experiment (Kerr, 2013) and were recorded by broadband seismic sensors. In order to achieve the best possible signal quality, we limited ourselves to the largest event (magnitude 5.6, November 6, 2011) in the sequence (Keranen et al., 2014). Throughout this study, only vertical-component records are used. The 17 stations contributing to our database are located at latitudes 34°N to 37°N and longitudes 94°W to 97°W. The earthquakes are demonstrably caused by injection of large volumes of wastewater from “hydrofracturing” (Keranen et al., 2014; van der Elst et al., 2013), for long-term storage, in formations that contained oil that was previously extracted. The high fluid pressures trigger earthquakes, particularly when the fluid accumulates on old, inactive faults, reactivating them (Fig. 2). These events have been selected for the large quantity and high quality of available data recorded locally at diverse azimuths and distances, for the reliability of hypocenter locations, and, after a preliminary auditory analysis, for the perceived quality of sonified signals.

The character of observed waveforms propagating through the region of interest is related to the properties of the underlying crust (Udías, 1999; Aki and Richards, 2002). These are best summarized by surface-wave phase velocities at different periods, each sampling a different depth range as illustrated, e.g., by Fry et al. (2010). The most recent and most complete surface-wave velocity model of North-America is that of Ekström (2013). We show in Fig. 3 a few examples



**Fig. 1.** Topography of the study area. The CMT focal mechanism (Dziewonski et al., 1981; Ekström et al., 2012) of the November 6, 2011, magnitude-5.6 event is plotted at the CMT epicenter location (compressional quadrants are shaded), suggesting a strike-slip fault with roughly SW-NE or SE-NW strike. Red triangles denote available seismic stations, whose names are specified. Different colors represent different elevations of the Earth's surface with respect to sea level. The dashed white line denotes the boundaries of our area of study. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)



**Fig. 2.** A qualitative representation of the sedimentary rocks of Oklahoma, where earthquakes have been triggered by wastewater injection. Seismic energy is reflected off the sediment/basement interface and the “Moho.” The sedimentary layers can have large velocity contrasts and add to the coda and complexity and character of the seismic wave and thus the sound.

of surface-wave phase-velocity perturbations according to Ekström (2013). The phase velocities of shorter-period waves in Fig. 3 are most sensitive to the thickness and elastic properties of young sedimentary rocks, forming a shallow layer that overlies the older, crystalline “basement”; 30–40 s Love and Rayleigh waves are strongly affected by variations in the depth to the Mohorovicic Discontinuity (also “Moho”), the interface between the crust and mantle, which has a high contrast in seismic wave speeds (on the order of 20%) at about 30–40 km depth in the area of study.

At most periods, both Love- and Rayleigh-wave velocities are highest in the north-eastern portion of the study area, and generally higher north of the 35°N parallel. A linear low-velocity anomaly can be distinguished to the south of and/or along the 35°N parallel, trending WSW-ESE. The south-east corner of the study area is, again, relatively fast at most periods.

The topography (Fig. 1) is relatively flat; some topographic highs to the south west are correlated with low long-period velocity of Rayleigh waves, and thus (probably) thicker-than-average crust. On the contrary, north of the 35°N parallel the topography is relatively high (>500 m) and so are crustal velocities.

The study of Reed et al. (2005) shows that the pattern of surface geology (age of sedimentary rocks) is correlated with that of topography, with older (Mississippian) outcrops found throughout the areas of higher topography in the eastern part of the study area, Pennsylvanian rocks along the 35°N parallel and 96°W meridian, and Permian further west, where topography again grows.

## 2.2. Sonification of seismic signals

All seismograms were sonified by changing their sampling frequency, from 40 Hz to 6000 Hz. The resulting signals (converted to wav files) are accordingly played  $6000/40 = 150$  times faster than their actual speed, and so translated to the audible frequency range. These sounds are somewhat reminiscent of gunshots, with a short release of impulsive, non-harmonic sound. Much of the signal that is usually analyzed by seismologists falls within the “attack” and in the first part of the “coda” (or “resonance”). The audio signals (impulse and coda) presented to the subjects have a 2 s-duration, corresponding to seismic signals of duration 300 s. The 300 s-long seismic signals to be sonified start 10 s (about 7 ms in audio scale) before the first important peak (the P wave) on the seismogram. Fig. 4 shows two sonified signals, with both audio and seismic wave scales. Fig. 4a shows the recording of the selected magnitude-5.6 event, by a station (V35A) which is close to the event location. The P- and S-wave arrivals do not have time to separate as they cover the short distance from

source to receiver (Udías, 1999; Aki and Richards, 2002), and, when sonified, the recording is characterized by a single “detonation.” Fig. 4b is the recording of the same event, by a station (W38A) which is far from the event location. The P- and S-wave arrivals are now well separated, and two distinct apparent “detonations” might be typically recognized in the sonified recording.

The dynamic range of seismic signals is greater than the dynamic range of audio signals, so the sonified signals have to be normalized. Each sonified signal was normalized with respect to its maximal value. This way, even though signal attenuates quickly as spherical seismic waves propagate away from the source, signals recorded at relatively large distances from the epicenter can still be heard and analyzed. This means, however, that the lower signal-to-noise ratio of lower-amplitude signals results in artificially large background noise; in other words, while the maximum amplitude (loudest peak) of the sonified signal is constant throughout our sonified database, signal-to-noise ratio systematically grows with decreasing source–receiver distance. Nearby events are not “louder”, but have a better signal-to-noise ratio than far-away events. This is a delicate issue that will be the subject of further study.

## 3. Experimental protocol

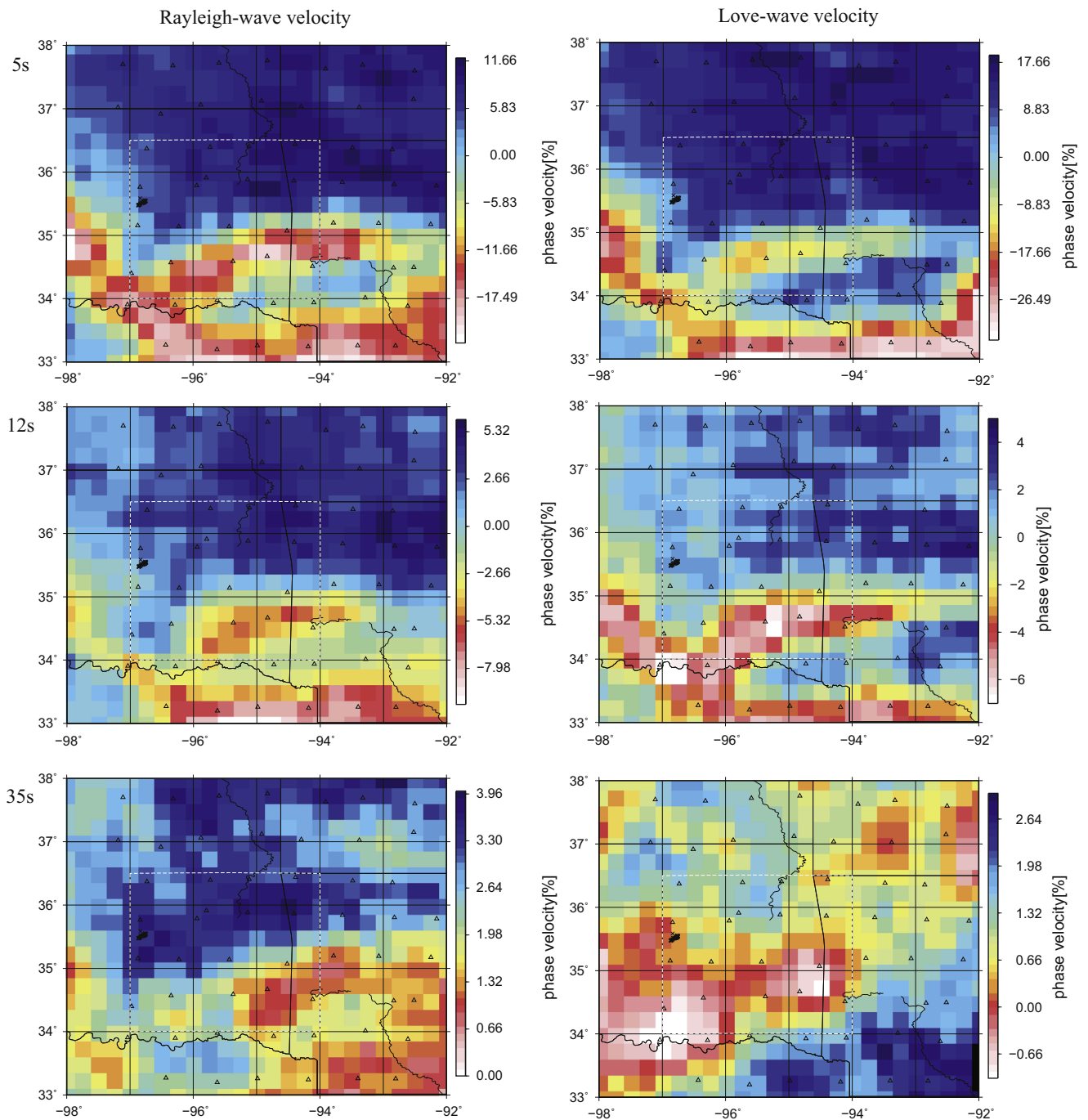
This section describes the experimental protocol used to test the sensitivity of the human auditory system to sonified seismic signals. First a short overview of the theoretical background (the categorization and the natural categories theory) from which the experimental method derives is given (Section 3.1). The experimental method is described in Section 3.2.

### 3.1. Theoretical background

Classic psychophysical methods are based on an underlying physical theory and are only valid as far as controlled physical parameters (e.g., the magnitude of an earthquake and amplitude of the corresponding signal) can be assumed to be linked to measurable psychological parameters (e.g., perceived “loudness”). This is not necessarily the case in this study: because our experiment is the first of its kind, we have no *a priori* knowledge of human responses to sonified seismic signal to rely on; we do not know what the relevant psychological parameters for the sonified seismic signals are, and have no elements on which to base assumptions.

The “free sorting” or “free categorization” method is more appropriate to the present scenario, as it requires little or no prior knowledge of the relationship between physical and psychological parameters. Free categorization task is now a classic method for instance in sensory analysis (Ballester et al., 2009; Dubois, 2009) or acoustics (Gaillard, 2000; Parizet and Koehl, 2011; Guastavino, 2007; Ballester et al., 2009; Morel et al., 2012; Paté et al., 2014). The method is based on the theory of categorization, which is comprehensively described in the field of cognition and cognitive psychology in Rosch and Lloyd (1978) and Dubois (1991): the human cognition process is described with the concept of *natural* categories, as opposed to *formal* categories, defined as follows.

Two objects (or stimuli) are placed in the same *formal* category if and only if they both fulfill a set of properties. Two objects do not belong to the same formal category if they differ in one property. This approach implies the assumption that the subject must have an analytical approach making him analyze each stimulus as a sum of independent properties. But, while the physical description of sound objects is classically based on a list of independent physical parameters, the corresponding perceptual parameters may be dependent/linked: for example, sound intensity and frequency are independent physical properties, but the



**Fig. 3.** Rayleigh-wave (left) and Love-wave (right) phase-velocity variations (with respect to the regional average) according to (Ekström, 2013), at periods of (top to bottom) 5, 12 and 35 s. The region is the same as that depicted in Fig. 1.

perceived sound intensity changes with frequency (Benade, 1990; Roederer, 2008).

On the other hand, two objects can be put into the same *natural* category if they share some features, and separated into different categories if they differ in some other respects: this aspect of the theory of natural categories can take into account the diversity of objects within a group, and the relationship between attributes, which are not assumed to be independent, and can be unknown *a priori*.

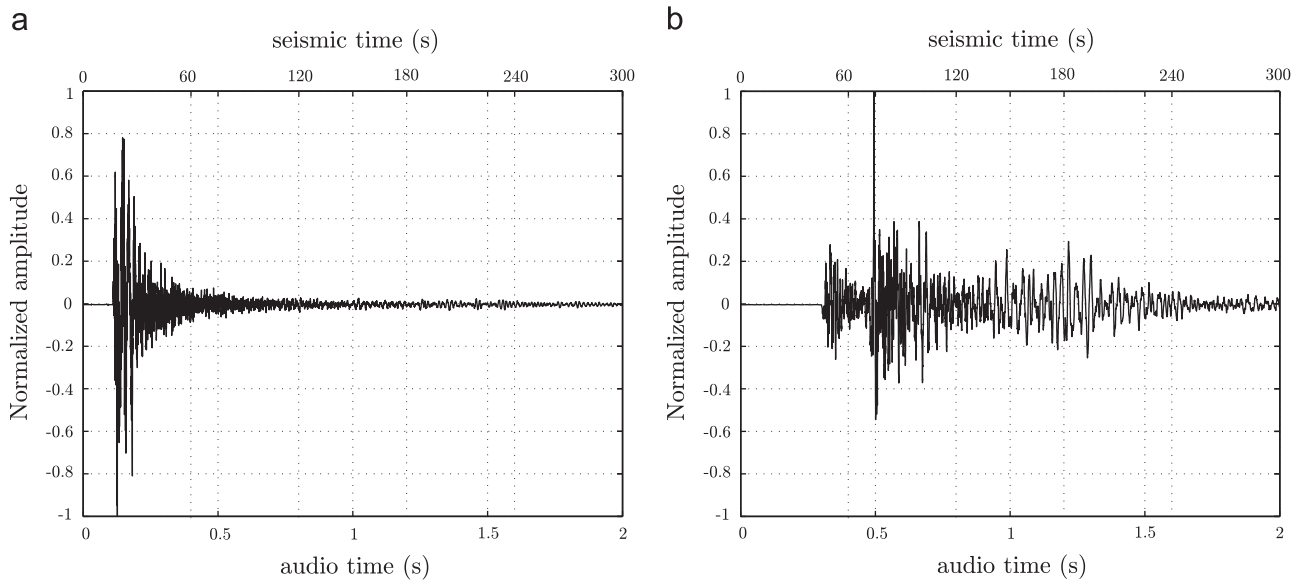
### 3.2. Free categorization method

We apply the free categorization method, consistent with the theory of categorization outlined in Section 3.1. The recordings by

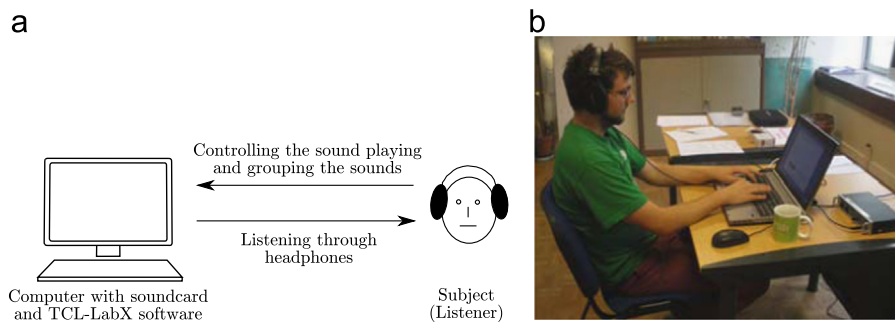
17 stations of the array described in Section 2.1 are sonified as described in Section 2.2: 17 sound samples are then to be listened to and categorized by the subjects. In the following, we take the psychoacoustical perspective and we use the word “stimulus” for “sound sample”. Two examples of such stimuli, corresponding to stations V35A (near the epicenter) and W38A (far from the epicenter), are shown in Fig. 4.

Of the  $K=24$  subjects who took part in the experiment, none reported to have prior expertise in the listening of sonified seismic signals. All subjects nevertheless had backgrounds in Earth sciences, acoustics or sound engineering, so have an *a priori* expertise in either listening or seismology.





**Fig. 4.** Recording of the magnitude 5.6 Oklahoma event, made at (a) station V35A (short propagation distance), and (b) station W38A (long propagation distance). Notice that the same onset time is used for all seismic records, which explains the slight delay between P-wave arrivals in (a) and (b).



**Fig. 5.** (a) Drawing of the experimental setup and (b) example of subject during the listening test.

### 3.2.1. Experimental setup

An external soundcard *RME Fireface UCX* connected to a computer was used for playing the sounds. *Sennheiser HD380 Pro* headphones were plugged to the output of the soundcard as illustrated in Fig. 5. For this preliminary study, involving only monophonic signals and no spatialization, listening through headphones is preferable to loudspeakers: the setup is simpler, and room effects are eliminated. The subjects could set and change the sound level in the headphones at any time during the test.

All tests were conducted on the same computer, where subjects used Pascal Gaillard's software *TCL-LabX* (Gaillard, 2014) to complete the proposed free categorization exercise. The *TCL-LabX* graphic interface displays the 17 stimuli as 17 small square icons, numbered randomly. A stimulus can be played back by double clicking on the corresponding icon, and a click-and-drag operation allows the subject to move each icon within the entire interface area. A screenshot of the interface is shown in Fig. 6.

### 3.2.2. Task and instructions

The subject is asked to sort the 17 stimuli into groups (the words “group” and “category” are used as synonyms throughout this study) that include stimuli perceived as similar by the subject. We report below a English translation<sup>1</sup> of the instructions:

<sup>1</sup> The original French instructions are: *Nous vous demandons de procéder à un tri des extraits sonores qui vous sont présentés. Pourriez-vous regrouper les extraits qui se ressemblent et placer dans des groupes différents ceux qui vous semblent différents ? Vous faites autant de groupes que vous le souhaitez.*

*Please sort the sound samples presented to you. You can group the samples which seem similar to you, and put in different groups those which seem different to you. You may form as many groups as you wish.*

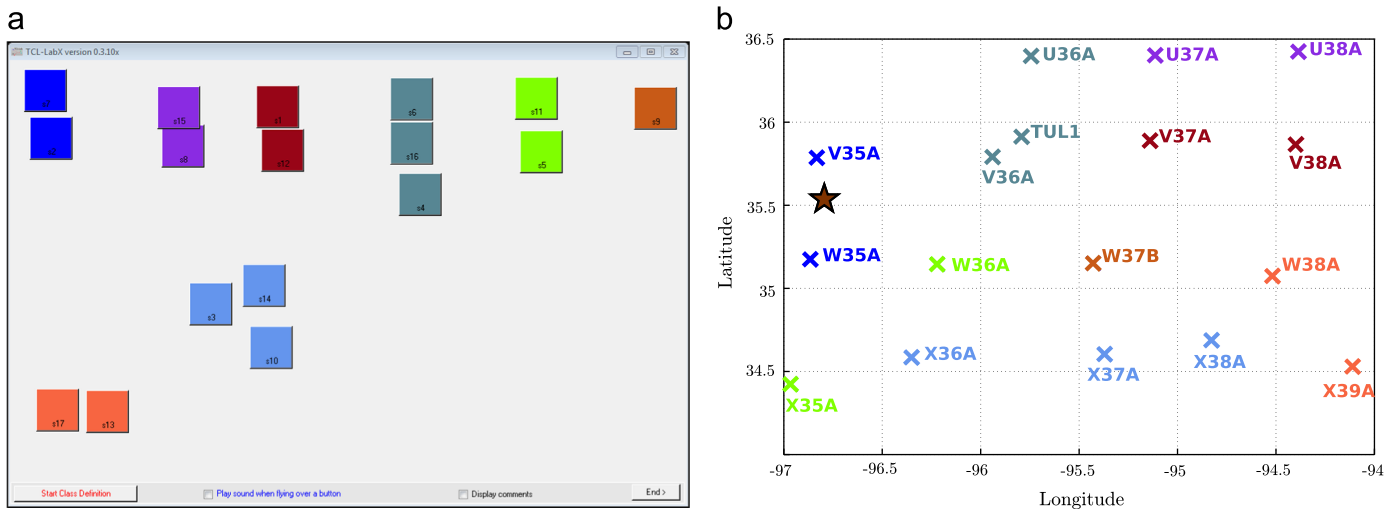
These instructions leave the subject quite free in selecting the similarity criteria, the number of groups to be formed, etc. The subject is allowed to group all stimuli into one group if no significant difference is perceived, and can form only groups that contain a single stimulus (“singleton” groups), if all stimuli are perceived as individuals sharing no particular similarity. Yet, all stimuli must belong to a group, and no stimulus can belong to two different groups. The subject is *not* asked to use a specific strategy in the spatial arrangement of the groups on the interface plane. In particular, the subject does not have to put far one from another groups that are perceived as more dissimilar: the position of groups and icons is not taken into account in the analysis phase, only the stimulus-group association matters.

No other information about the nature of the data (other than that they were originated from seismograms) were provided. Prior knowledge of the origin of the sound samples potentially would have induced expectations and listening strategies, reducing the “freedom” that was given to subjects. However, no subject reported the use of his/her knowledge of seismology, and no one knew what particular earthquakes were used.

24 subjects (16 males, 8 females) took part in the experiment. The mean duration of the test was 15.5 min (standard deviation: 9.5 min, maximal duration: 44.7 min, minimal duration: 6.5 min).



**Fig. 6.** Screenshot of the computer interface during the test. Each of the 17 square icons represents a sound stimulus. The subject first moves (click-and-drag operation) the icons within the interface area (this figure) and second gives each icon a color corresponding to the group to which it belongs (see Fig. 7a for the second stage).



**Fig. 7.** The partition made by one subject. (a) Screenshot of the test interface taken at the end of the test; icons belonging to the same group have the same color. (b) The same partition displayed on the latitude/longitude plane. Station locations are indicated by colored crosses accompanied by station names, while a star denotes the epicenter location; the color code is the same as in (a).

A single stimulus was played 9.5 (mean value) times in average (standard deviation: 2.1): the most often played stimulus is the recording from station V38A (13.6 times in average), the less often played stimulus is the recording from station W38A (6.8 times in average).

#### 4. Analysis of experimental data

Each of our 24 subjects provided a “sorting” (or “partition”) of the 17 stimuli. The output data of the test then consists of 24 sets of 1–17 categories/groups of stimuli. We call “individual partition” each of

those 24 partitions. Fig. 7 shows an example of an individual partition: this sample partition can be displayed as the final organization of icons in the test interface (Fig. 7a), or on the latitude/longitude plane (Fig. 7b). The data can be analyzed in several ways; after briefly quantifying the global inter-subject consensus (Section 4.1), we describe two analysis approaches in Sections 4.2 and 4.3, and discuss their outcomes in Section 4.4.

##### 4.1. Inter-subject consensus

In order to characterize differences between different partitions, the Rand index can be used (Rand, 1971). One value of the Rand index is computed for each possible pair of subjects; with  $K$

subjects, this amounts  $K(K - 1)/2$  pairs to be compared, i.e. 276 pairs if  $K=24$ . For subjects A and B, the Rand index is defined as  $R = (a + b)/(a + b + c + d)$ , where

- $a$  is the number of pairs of stimuli put in the same category by subject A and in the same category by subject B;
- $b$  is the number of pairs of stimuli put in different categories by subject A and in different categories by subject B;
- $c$  is the number of pairs of stimuli put in the same category by subject A, and in different categories by subject B;
- $d$  is the number of pairs of stimuli put in different categories by subject A, and in the same category by subject B.

In practice,  $R$  can take values from 0 to 1, with  $R=0$  corresponding to total disagreement between the partitions of subjects A and B ( $a=b=0$ ), and  $R=1$  to total agreement ( $c=d=0$ ).

#### 4.2. Tree analysis

The tree analysis aims at defining a perceptual distance between stimuli, and to represent this distance on an “additive tree”. We follow the classic criteria described by, e.g., Paté et al. (2014) and Guastavino (2003), which can be summarized as follows:

- (i) Compute an “individual” co-occurrence matrix  $M^k$  (square matrix whose size is defined by the number of stimuli, i.e.  $17 \times 17$  in the present case) for each subject  $k$  ( $k = 1, 2, \dots, 24$ ):
  - $M_{ij}^k = 1$  if stimuli  $i$  and  $j$  are in the same group according to subject  $k$ ;
  - $M_{ij}^k = 0$  if stimuli  $i$  and  $j$  are in different groups according to subject  $k$ .
- (ii) Calculate the total co-occurrence matrix, defined as the sum of all  $K$  individual co-occurrence matrices:  $M_{ij} = \sum_{k=1}^K M_{ij}^k$  ( $M_{ij}$  is large if stimuli  $i$  and  $j$  are often grouped together, and 0 if they are never grouped together).
- (iii) Convert the co-occurrence measure into the distance matrix  $D$  such that  $D_{ij} = 1 - M_{ij}/17$  ( $D_{ij}$  is small if  $M_{ij}$  is large;  $D_{ij} = 0$  if stimuli  $i$  and  $j$  are always grouped together, and 1 if they are never grouped together).

The matrix  $D$  is a “consensual” measure of perceptual distance, since it expresses a consensus among subjects, and smooths the differences between subjects. The information contained in  $D$  can be visualized by a “tree,” as described by Barthélémy and Guénoche (1991): each sound stimulus is represented by a “leaf,” and leaves are linked together through “branches,” whose length is proportional to the perceptual distance  $D$  between leaves/stimuli. For instance, if one has to climb (or descend) along many and/or long branches to go from leaf A to leaf B, that means that stimuli A and B have been perceived as very different by the subjects. We find the best-fitting tree to our  $D$  via Jacques Poitevineau's *Addtree* software (Poitevineau, 2014a).

#### 4.3. Central partition analysis

The central partition analysis consists essentially of identifying one or more “consensual” partitions, each representing an “average” of the individual partitions produced by an ensemble of subjects. As opposed to the tree analysis, of Section 4.2, this approach does not provide a notion of “distance” between stimuli. Each central partition then describes a consensus between subjects; this approach to data analysis does not allow one to visualize differences in subjects' responses, nor the individual listening behaviours or grouping strategies.

The central partition corresponding to a group of  $K$  subjects can be identified by the following procedure, based on Marcotorchino and Michaud (1982) and Brenac (2002), which we implement via Jacques Poitevineau's *wpartcent* software (Poitevineau, 2014b):

- (i) Compute the “individual” co-occurrence matrix  $M^k$ , defined in Section 4.2, for each subject  $k$  ( $k = 1, \dots, K$ ).
- (ii) Calculate the total co-occurrence matrix, as defined in Section 4.2, associated with the  $K$  subjects.
- (iii) Determine a matrix  $C$  of size  $N \times N$  with  $N$  the number of stimuli (i.e., 17), such that  $\sum_{i=1}^N \sum_{j=1}^N (2M_{ij} - K)C_{ij}$  is maximum; the matrix  $C$  is sought via optimization (Poitevineau, 2014b) within the space of all real  $N \times N$  matrices sharing the following properties:
  - $C_{ij}$  is either 0 or 1;
  - $C_{ij} = C_{ji}$  (symmetry);
  - $C_{ij} + C_{jk} - C_{ik} \leq 1$  (transitivity);
  - $C_{ii} = 1$  (reflexivity).
- (iv) The central partition is defined based on the so determined matrix  $C$ , placing stimuli  $i$  and  $j$  in the same group if  $C_{ij} = 1$ , and in different groups otherwise.

To understand the algorithm, notice, e.g., that for large  $M_{ij}$ , optimization will tend to pick  $C_{ij} = 1$  rather than 0; conversely, if  $M_{ij}$  is close to 0, its contribution to the function to be optimized is negative, and it will be preferable to pick  $C_{ij} = 0$ . If  $M_{ij}$  is neither large nor small, its contribution is close to negligible and either value of  $C_{ij}$  might be picked.

#### 4.4. Outcome of the analysis

Instructed to form groups of stimuli, the subjects actually tried and managed to do so. Fig. 8a shows the distribution of the number of categories in individual partitions. As we can see, no subject chose to form one single group containing all stimuli, or to form as many singleton groups as there were stimuli. While differences between stimuli within a group might be perceived, subjects have nevertheless recognized common properties, that allowed them to group the stimuli together.

Fig. 8b shows the distribution of the number of stimuli in categories. It can be seen that 34 categories (over a total amount of 132 categories when summing over the 24 individual partitions) contain only 1 stimulus, so that the grouping of stimuli according to the identification of common features or similarity may not always be the universal strategy to evaluate such stimuli. However, the large majority of the other categories contain 2–6 stimuli. We conclude that the subjects succeeded in producing a categorization of the sound stimuli.

It can be noticed that out of 34 singleton categories, 7 (21% of the singleton categories) consist of stimulus  $U38A$ , 5 (15%) of stimulus  $W37B$ , and 3 (9%) of stimulus  $V35A$ : those seem to be the most unique stimuli in the present 17-stimuli-corpus. On the contrary, stimulus  $TUL1$  was never placed alone into a singleton category: all subjects identified common properties between stimulus  $TUL1$  and others (in particular stimulus  $V36A$ , which is put only 1 time in a singleton category, and 19 times in the same category as  $TUL1$ ).

##### 4.4.1. Measured inter-subject consensus

Fig. 9 shows the distribution of the  $R$  over the 276 pairs of partitions, ranging from 0.38 (fairly good agreement) to 0.96 (very strong agreement) with an average value of 0.76. In summary,  $R$  is fairly high, indicating that there is some consensus amongst subjects in categorizing the stimuli. No “outliers” are found, and the entire database will be the object of our further analysis in the following.

##### 4.4.2. Outcome of the tree analysis

Fig. 10 shows the matrix  $D$  defined in Section 4.2, computed from all 24 partitions, visualized via the *Addtree* software. It is apparent from Fig. 10 that some stimuli are often grouped together (e.g.,  $X38A$ ,  $X37A$ ,  $X39A$ , and  $W38A$ ), other only very rarely (e.g.,

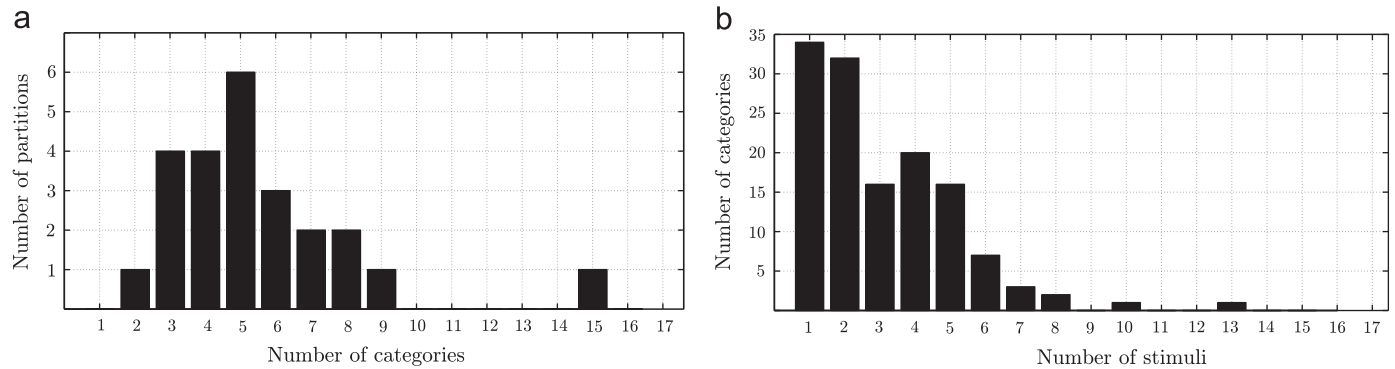


Fig. 8. Histogram showing (a) the distribution of the number of categories in partitions and (b) the distribution of the number of stimuli in categories.

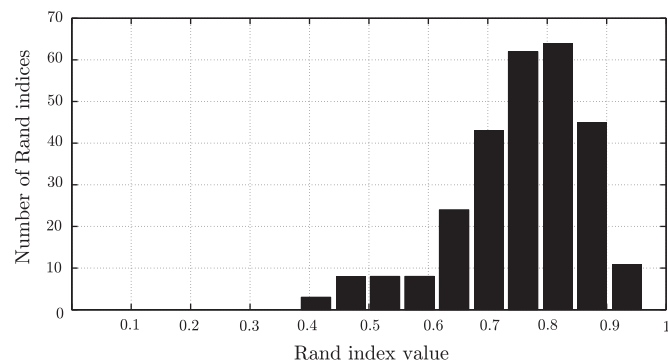


Fig. 9. Histogram of the Rand indices computed from the 276 pairs of partitions.

W35A, U37A, and X39A). Some “consensual” groups can be identified:

- group a: X38A, X37A, X39A, W38A;
- group b: W35A, V35A (less frequently W36A);
- group c: TUL1, V35A (less frequently U36A);
- group d: V37A, V38A;
- group e: U37A, U38A.

Members of groups d and e are all relatively close in Fig. 10 and these two groups could also be merged in a single, larger group. Keeping in mind the geographical distribution of receivers (Fig. 1), one can already notice that group b essentially consists of sounds from stations close to the epicenter, group c of sounds from intermediate-distance stations, and groups a, d and e of sounds from relatively far stations. Sections 4.4.3 and 5 below confirm that source–station distance is likely to play an important role in the categorization strategy.

#### 4.4.3. Outcome of the central partition analysis

The central partition analysis (Section 4.3) allows to define a single partition, most representative of the choices made by the entire population of 24 listeners. We visualize such central partition in Fig. 11 on a latitude/longitude plane, each color denoting one category.

With the exception of group {X35A, W37B}, these “average” categories include stimuli from stations that lie close one to another: on the latitude/longitude plane, the groups appear more as “lumps” of stations than as a “scattering” of stations. This indicates that stimuli may be grouped according to spatial proximity or common geological properties of the corresponding stations.

The geographical organization of the average groups seems to confirm the importance of the source–receiver distance as a grouping criterion: group {V35A, W35A} is made up of stimuli from stations close to the source, groups {V36A, TUL1, U36A} and {X35A, W37B} are

made up of stimuli from stations at a medium distance from the source, and groups {U37A, U38A} and {V37A, V38A} are made up of stimuli from stations far from the source. However, there must be other parameters relevant to explain the groupings, as for example category {X36A, X37A, X38A, X39A, W38A} is made up of stimuli from stations at very different distances from the source. Section 5 investigates the relation between the perceptual similarities and the geophysical parameters.

## 5. Linking perceived similarities to geophysical parameters

The seismograms we record and sonify are known to be controlled by several geophysical parameters (e.g., source–receiver distance, Earth structure, topography at the receiver, and the mechanical properties of the seismic source), resulting in sound features that the auditory system seems able to discriminate.

The two stations that are most often grouped together are V35A and W35A, which 21 out of 24 listeners have placed in the same category, and which are closest to the epicenter. At such a short source–receiver distance, P and S waves hit the receiver almost simultaneously, which is not the case at larger source–receiver distances included in our data set. Preliminary analysis of the subjects’ textual comments (which we shall analyze in detail in a separate publication) confirms that most listeners have used the lack of separation between the P- and S-wave arrivals as a criterion for categorization.

Stations TUL1 and V36A, which are very close to one another, were placed in the same category by 19 out of 24 subjects. 10 of these subjects have also included station U36A in the same group. The resulting consensual category {TUL1, U36A, X36A} includes signals that are characterized by similar crustal structure and distance between source and receiver, as well as similar receiver elevation.

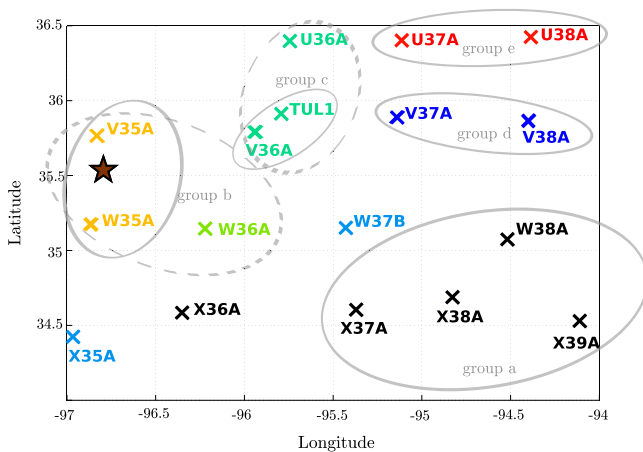
Fig. 1 and the top panels of Fig. 3 (Rayleigh- and Love-wave velocity at 5 s period, corresponding to shallow crust) show that stations X36A, X37A, X38A, X39A and W38A sit on similar terrain, characterized by low surface-wave velocity and (with the exception of station X36A) relatively high topography (roughly 300–500 m above sea level). These five stations were recognized as a single category by 10 subjects (plus one additional subject who also included other stimuli in the category). 7 other subjects exclude X36A from what is otherwise the same category: notice that station X36A lies on slightly different terrain and is much closer to the epicenter.

Stimuli associated with stations V37A, V38A, U37A and U38A, which all lie in the highest-topography area of the study region (Fig. 1), are perceived as relatively similar (Fig. 10), and 7 listeners have chosen to place them all in one category. There is, however, a pronounced tendency to form two separate categories, one consisting of U37A and U38A (grouped together 17 times), the other of





**Fig. 10.** Additive tree representing the perceptual distances between stimuli, computed from all 24 partitions. Each line (or “branch”) ends in a stimulus or “leaf,” or is connected to another branch. The perceptual distance between two stimuli is proportional to the distance between the corresponding leaves, measured along the branches that connect them. Leaves are labelled with station names, and epicenter-station distance in km. Gray circles indicate the groups identified in the analysis (a–e): solid lines denote groups of stimuli often grouped together, and dashed lines add stimuli perceived as similar as well, but less frequently included in those groups.



**Fig. 11.** The central partition computed from the 24 individual partitions, displayed on the latitude/longitude plane. Station names and locations (crosses) share the same color if the corresponding stimuli are found to belong to the same category in the central partition. The brown star indicates the epicenter localization. For comparison purpose, the gray circles correspond to the groups identified in the tree analysis (Fig. 10). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

V37A and V38A (16 times). The underlying Earth’s structure might slightly differ between northern stations U37A and U38A, and southern stations V37A and V38A (Fig. 3).

In summary, Fig. 11 suggests that audio signals have been categorized by most listeners in terms of (i) distance between source and

receiver, and (ii) geographic location of the receiver. Source–receiver distance controls the delay between the arrival of P and S waves, which is a prominent feature of audio signals (as a result, stations V35A and W35A are almost always grouped together). Source–receiver distance, however, is not the only parameter affecting the listeners’ categorization strategies; for instance, stations V37A and X37A, which lie at approximately the same distance from the epicenter, are never or rarely grouped together; the same holds for V38 and W38.

Receiver location determines the Earth structure sampled by seismic waves, resulting in different frequency content, dispersion, scattering of the waveform and, consequently, of the sonified seismogram. We have identified in Fig. 11 a tendency to group together stations that lie on similar terrain; it is harder to determine whether ray paths that sample similar terrain also result in sonified seismograms that are perceived as similar: this will be the subject of future work.

Receivers that lie at the same distance from a seismic source, but at different azimuths (e.g., again, V37A and X37A), stand in a different geometrical relationship with the seismic fault: Fig. 1 shows, e.g., that stations X35A and X36A are in the focal mechanism’s “compressional quadrant” while the majority of stations to the east lie in the “extensive quadrant” (Aki and Richards, 2002; Dziewonski et al., 1981; Ekström et al., 2012). It is likely that this also affects the seismic waveform in a way that is reflected by the sonified seismograms, but the available data do not yet allow us to make any strong inferences on this issue, which will have to be explored with *ad-hoc* experiments.

## 6. Conclusions and future work

Our analysis shows that listeners perceive “clues” in sonified seismic signals, and use such clues to try to discriminate them. While the number of listeners is too small for a rigorous statistical analysis, there are clear similarities between the responses of different listeners.

The way subjects categorize sonified seismic events seems to be consistent with some physical properties of these events. The tree analysis (Fig. 10) showed that, to a first approximation, the groups are defined according to the epicenter-station distance. This is presumably related to the auditory discrimination of P- and S-wave onset times. There are, however, exceptions to this grouping strategy, suggesting that the subjects make use of other clues in the signal in order to perform the categorization task. The central partition analysis (Fig. 11) shows that the average categories (computed from all subjects) are well explained in terms of geographical and geological similarities: the average categories tend to group together stimuli from stations which are close one to another (in Fig. 11 we see more “lumped” groups than “scattered” groups). This is in agreement with the fact that source-receiver azimuth, underlying Earth structure, station elevation all affect the seismic waveform, and thus its sonified counterpart.

In summary, we have found that the effects of several physical parameters that affect a seismic waveform can also be identified, by the human auditory system, in sonified seismograms. Disentangling the role of each parameter is a nontrivial problem requiring further work. We shall address it by new experiments, using “synthetic” seismic waveforms based on very simple theoretical seismic models; controlled-source (rather than earthquake) data from a particularly homogeneous region could also be employed. We plan to form a database of “seismic sounds” that only differ by one or few, well-known parameters; the sensitivity of listeners to each parameter/combination of parameters will thus be quantified. The performance of listeners at specific tasks will serve to evaluate more precisely the applicability of auditory analysis to real seismic signals.

Plans for future work include the application of current auditory scene synthesis (Berkhout et al., 1993; Corteel, 2007; Spors et al., 2011) techniques to seismic sonification, representing the seismic wave field by a three-dimensional soundscape in which listeners can move freely, making use of a set of simultaneous seismic records from a dense broadband array. It can be envisaged that the location of a receiver in real space will define that of a source in virtual space, and seismic wave propagation observed in areas that are well covered by seismic instruments will be represented fairly accurately by a synthesized soundscape. The spatialization can be achieved by current spatial audio technology, or even binaurally if the listener's HRTF (head-related transfer function) is convolved with the synthesized sounds, while the listener's motion is tracked in real time to render the soundfield according to his/her viewpoint and orientation. The ability to explore the seismic wave field in three dimensions will facilitate auditory analysis and presumably lead to better performance (Wightman and Kistler, 1999; Begault et al., 2001).

Furthermore, while most classic seismology applications only involve seismic records spanning not more than a few hours after an event, reconstructing the impulse response of the Earth's crust from ambient signal by two-receiver interferometry (Campillo and Paul, 2003; Stehly et al., 2006) requires the analysis of very long continuous records, from several days to an entire year depending on frequency range and source distribution; the time-acceleration, of a factor  $\sim 10^3$ , inherent to seismogram sonification makes it possible to “play” such records in a reasonable amount of time, so that open questions on the nature of the observed seismic background signal could then be addressed by listening experiments.

## Acknowledgments

We gratefully acknowledge financial support from INSU-CNRS which made our work possible. We would like to thank François Baudin, Nicolas Bellahsen, Etienne Caylou, Delphine Chadeaux, Elia d'Acremont, Danièle Dubois, Hugo Dujourdy, Augustin Ernoult, Claudia Fritz, Clément Gariel, Hugues Genevois, Ambroise Harivel, Philippe Huchon, Elsa Jauffret, Emanuel Kaestle, Félicie Korostelev, Alexis Plunder, Laurent Quartier, Alain Rabaute, Alexandre Roy, Alexandre Schubnel, Pauline Thierry, Camille Vauthrin, and Jean-Louis Waflart for taking part in the psychoacoustic tests. Thanks to Nolan Lem and Pascal Gaillard for fruitful discussion, and to Danièle Dubois for her precious, insightful analysis of our data. Thanks also to Nolan Lem, Meredith Nettles, Hannah Rabinowitz, and Heather Savage, for discussions and being subjects of a preliminary test at the Lamont-Doherty Earth Observatory, and to Katie Keranen, Patty Lin and Nolan Lem for data and other assistance. Lapo Boschi is grateful to Florian Dombois and Olivier Warusfel for some very interesting discussions, that inspired part of this study.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.ijhcs.2015.08.002>.

## References

- Aki, K., Richards, P.G., 2002. *Quantitative Seismology*, second ed. University Science Books, Sausalito, CA.
- Ballester, Jordi, Abdi, Herv, Langlois, Jennifer, Peyron, Dominique, Valentin, Dominique, 2009. The odor of colors: can wine experts and novices distinguish the odors of white, red, and ros wines? *Chemosens. Percept.* 2 (4), 203–213.
- Barthélémy, J.-P., Guénoche, A., 1991. *Trees and Proximity Representations*. J. Wiley, Chichester, New York.
- Begault, D., Wenzel, E.M., Anderson, M.R., 2001. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *J. Audio Eng. Soc.* 49, 904–916.
- Benade, A.H., 1990. *Fundamentals of Musical Acoustics*. Dover Publications, New York.
- Berkhout, A.J., de Vries, D., Vogel, J.-J., 1993. Acoustic control by wave field synthesis. *J. Acoust. Soc. Am.* 93, 2765–2778.
- Brenac, T., 2002. Méthodes de partition centrale appliquées à l'étude de catégories cognitives (Central partition methods applied to the study of cognitive categories). In: Poitevineau, J. (Ed.), *Arbres, Classes, Distances (Trees, Classes, Distances)—Cahiers du LCPE (6)*. LCPE, Paris, France (Chapter 5).
- Campillo, M., Paul, A., 2003. Long-range correlations in the diffuse seismic coda. *Science* 299, 547–549.
- Corteel, E., 2007. *Synthesis of directional sources using wave field synthesis, possibilities, and limitations*. EURASIP J. Adv. Signal Process., 90509.
- Cowen, R., 2015. Sound bytes. *Sci. Am.* 312, 44–47.
- Dombois, Florian, 2001. Using audification in planetary seismology. In: Proceedings of the International Conference on Auditory Display.
- Dombois, Florian, 2002. Auditory seismology—on free oscillations, focal mechanisms, explosions and synthetic seismograms. In: Proceedings of the International Conference on Auditory Display.
- Dombois, Florian, Eckel, Gerhard, 2011. Audification. In: Thomas Hermann, Andy Hunt, John G. Neuhoff (Eds.), *The Sonification Handbook*. Logos Verlag, Berlin, Germany (Chapter 12).
- Dubois, D., 1991. *Sémantique et cognition (Semantics and Cognition)*. Editions du CNRS, Paris, France.
- Dubois, D., 2009. *Le Sentir et le Dire : concepts et méthodes en psychologie et linguistique cognitives (Feel and Say: Concepts and Methods in Cognitive Psychology and Linguistics)*. l'Harmattan, Paris, France.
- Dziewonski, A.M., Chou, T.A., Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *J. Geophys. Res.* 86, 2825–2852. <http://dx.doi.org/10.1029/JB086iB04p02825>.
- Ekström, G., 2013. Love and Rayleigh phase-velocity maps, 5–40 s, of the western and central USA from USArray data. *Earth Planet. Sci. Lett.* 402, 42–49. <http://dx.doi.org/10.1016/j.epsl.2013.11.022>.
- Ekström, G., Nettles, M., Dziewonski, A.M., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes. *Phys. Earth Planet. Inter.* 200–201, 1–9. <http://dx.doi.org/10.1016/j.pepi.2012.04.002>.

- Frantti, G.E., Levereault, L.A., 1965. Auditory discrimination of seismic signals from earthquakes and explosions. *Bull. Seismol. Soc. Am.* 55 (1), 1–25.
- Fry, B., Deschamps, F., Kissling, E., Stehly, L., Giardini, D., 2010. Layered azimuthal anisotropy of rayleigh wave phase velocities in the European Alpine lithosphere inferred from ambient noise. *Earth Planet. Sci. Lett.* 297 (1–2), 95–102.
- Gaillard, P., 2000. Etude de la perception des transitoires d'attaques des sons de steeldrums : particularités acoustiques, transformation par synthèse et catégorisation (Study of the perception of attack transients of steeldrums' sounds: acoustic characteristics, synthesis and categorization). Ph.D. Thesis, Université de Toulouse II - Le Mirail.
- Gaillard, P., 2014. Website for *TCL-LabX* (<http://petra.univ-tlse2.fr/tcl-labx>).
- Guastavino, C., 2003. Etude sémantique et acoustique de la perception des basses fréquences dans l'environnement urbain (Semantic and acoustical study of the perception of low frequencies in the urban environment). Ph.D. Thesis, Université Pierre et Marie Curie.
- Guastavino, C., 2007. Categorisation of environmental sounds. *Can. J. Exp. Psychol.* 61 (1), 54–63.
- Hartmann, W.M., 1999. How we localize sound. *Phys. Today* 11, 24–29.
- Hayward, Chris, 1994. Listening to the earth sing. In: Gregory Kramer (Ed.), *Auditory Display—Sonification, Audification, and Auditory Interfaces*. Addison–Wesley, Reading, Massachusetts, pp. 369–404 (Chapter 15).
- Hermann, Thomas, Hunt, Andy, Neuhoff, John G., 2011. *The Sonification Handbook*. Logos Verlag, Berlin.
- Holtzman, Benjamin, Candler, Jason, Turk, Matthew, Peter, Daniel, 2014. Seismic sound lab: sights, sounds and perception of the earth as an acoustic space. In: *Sound, Music, and Motion*. Springer International Publishing, Cham, Switzerland, pp. 161–174.
- Karney, M.S., December 2006. Natural radio (news, comments and letters about natural radio) ([http://naturalradiolab.com/PDF/NR2006\\_12.pdf](http://naturalradiolab.com/PDF/NR2006_12.pdf)), (retrieved on 02.06.2015).
- Keranen, Kathleen, Weingarten, Matthew, Abers, Geoffrey A., Bekins, Barbara, Ge, Shemin, 2014. Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. *Science* 345, 448–451.
- Kerr, Richard, 2013. Geophysical exploration linking deep earth and backyard geology. *Science* 340, 1283–1285.
- Kilb, Debi, Peng, Zhigang, Simpson, David, Michael, Andrew Meghan, Fisher, Meghan, Rohrlück, Daniel, 2012. Listen, watch, learn: SeisSound video products. *Seismol. Res. Lett.: Electron. Seismol.* 83 (2), 281–286.
- Mai, P.M., Schorlemmer, D., Page, M., 2012. The source inversion validation (SIV) initiative: a collaborative study on uncertainty quantification in earthquake source inversions. In: EGU General Assembly.
- Marcotorchino, J.-F., Michaud, P., 1982. Agrégation de similarités en classification automatique (Aggregation of Similarities in Automatic Classification). *Rev. Stat. Appl.* 30 (2), 21–44.
- Meier, Manuela, Saranti, Anna, 2008. Sonic exploration with earthquake data. In: *Proceedings of the International Conference on Auditory Display*.
- Michael, A.J., 2011. Earthquake sounds. In: Gupta, H. (Ed.), *Encyclopedia of Solid Earth Geophysics*. Springer, Dordrecht, The Netherlands.
- Morel, J., Marquis-Favre, C., Pierrette, M., 2012. Road traffic in urban areas: a perceptual and cognitive typology of pass-by noise. *Acta Acust. United Acust.* 98, 166–178.
- Parizet, E., Koehl, V., 2011. Application of free sorting tasks to sound quality experiments. *Appl. Acoust.* 73, 61–65.
- Paté, A., Le Carrou, J.-L., Navarret, B., Dubois, D., Fabre, B., 2014. Influence of the electric guitar's fingerboard wood on guitarist's perception. *Acta Acust. United Acust.* 101 (2), 347–359.
- Peng, Zhigang, Aiken, Chastity, Kilb, Debi, Shelly, David R., Enescu, Bogdan, 2012. Listening to the 2011 magnitude 9.0 Tohoku-Oki, Japan, earthquake. *Seismol. Res. Lett.* 83 (2), 287–293.
- Poitevineau, J., 2014a. Website for addtree (<http://petra.univ-tlse2.fr/tcl-labx>).
- Poitevineau, J., 2014b. Website for wpartcent (<http://petra.univ-tlse2.fr/tcl-labx>).
- Rand, W.M., 1971. Objective criteria for the evaluation of clustering methods. *J. Am. Stat. Assoc.* 66 (336), 846–850.
- Reed, J.C., Wheeler, J.O., Tucholke, B.E., 2005. Decade of North American Geology—Geologic Map of North America. Published by the Geological Society of America, Boulder, Colorado, available online at (<http://ngmdb.usgs.gov/gmna>).
- Roederer, J.G., 2008. *The Physics and Psychophysics of Music: An Introduction*, fourth ed. Springer, New York, USA.
- Rosch, E., Lloyd, B.B., 1978. *Cognition and Categorization*. Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Ruiz, S., Madariaga, R., 2013. Kinematic and dynamic inversion of the 2008 Northern Iwate earthquake. *Bull. Seismol. Soc. Am.* 103, 694–708.
- Speeth, Sheridan D., 1961. Seismometer sounds. *J. Acoust. Soc. Am.* 33 (7), 909–916.
- Spors, S., Helwani, K., Ahrens, J., 2011. Local sound field synthesis by virtual acoustic scattering and time-reversal. In: 131st Audio Engineering Society Convention.
- Stehly, L., Campillo, M., Shapiro, N.M., 2006. A study of the seismic noise from its long-range correlation properties. *J. Geophys. Res.* 111 (B10).
- Steinbrugge, K.V., 1974. A catalog of earthquake related sounds. *Bull. Seismol. Soc. Am.* 64 (5), 1409–1418.
- Udias, A., 1999. *Principles of Seismology*. Cambridge University Press, Cambridge, UK.
- van der Elst, Nicholas J., Savage, Heather M., Keranen, Katie M., Abers, Geoffrey A., 2013. Enhanced remote earthquake triggering at fluid-injection sites in the midwestern United States. *Science* 341 (6142), 164–167.
- Wightman, F.L., Kistler, D.J., 1999. Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J. Acoust. Soc. Am.* 105, 2841–2853.
- Worrall, D., 2009. Information sonification: concepts, instruments and techniques. (Ph.D. thesis), University of Canberra.