



HAL
open science

Auditory display of seismic data: On the use of experts' categorizations and verbal descriptions as heuristics for geoscience

Arthur Paté, Lapo Boschi, Danièle Dubois, Jean-Loic Le Carrou, Benjamin Holtzman

► To cite this version:

Arthur Paté, Lapo Boschi, Danièle Dubois, Jean-Loic Le Carrou, Benjamin Holtzman. Auditory display of seismic data: On the use of experts' categorizations and verbal descriptions as heuristics for geoscience. *Journal of the Acoustical Society of America*, 2017, 141, pp.2143 - 2162. 10.1121/1.4978441 . hal-01522831

HAL Id: hal-01522831

<https://hal.sorbonne-universite.fr/hal-01522831v1>

Submitted on 15 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

11 **Abstract:** Auditory display can complement visual representations in order to better interpret scientific
 12 data. A previous article showed that the free categorization of “audified seismic signals” operated by listeners can
 13 be explained by various geophysical parameters. The present article confirms this result and shows that cogni-
 14 tive representations of listeners can be used as heuristics for the characterization of seismic signals. Free sorting
 15 tests are conducted with audified seismic signals, with the earthquake/seismometer relative location, playback
 16 audification speed, and earthquake magnitude as controlled variables. The analysis is built on partitions (cat-
 17 egories) and verbal comments (categorization criteria). Participants from different backgrounds (acousticians
 18 or geoscientists) are contrasted in order to investigate the role of the participants’ expertise. Sounds resulting
 19 from different earthquake/station distances or azimuths, crustal structure and topography along the path of
 20 the seismic wave, earthquake magnitude, are found to a) be sorted into different categories, b) elicit different
 21 verbal descriptions mainly focused on the perceived number of events, frequency content, and background noise
 22 level. Building on these perceptual results, acoustic descriptors are computed and geophysical interpretations
 23 are proposed in order to match the verbal descriptions. Another result is the robustness of the categories with
 24 respect to the audification speed factor.

25 **Keywords:** free sorting; categorization; sonification; perception; auditory display; verbalization

26

27 **PACS:** 43.66.Lj; 43.40.Ph; 43.75.Cd

28 I Introduction

29 The development of seismology as a scientific discipline has traditionally been based upon graphical tools
 30 (through visualization of empirical data on graphs, lists, and figures^{1;2}), and primarily upon the visual analysis
 31 of seismograms³, which are representations of recordings of the oscillations of a point at the Earth’s surface. The
 32 advent of digitized data acquisition and the development of modern signal processing techniques has facilitated
 33 the representation of seismic data (or of potentially any data eliciting no *a priori* modality of display) through
 34 other sensory modalities. The work presented here deals with the auditory representation (“auditory display”)
 35 of seismic data.

36 Many instances of educational^{9;10;11;12;13;14;15;16} or artistic^{6;7;8} uses of seismic data “sonification” have been
 37 reported. However, to our knowledge, it has only been used twice for scientific research purposes: Speeth⁴ and
 38 then Frantti and Levereault⁵ accelerated seismic signals in order to shift the frequencies to the audible range,
 39 and trained people to tell “natural” earthquakes (“double-couple” sources) from explosions by listening to accel-
 40 erated seismograms. This promising approach has not found practical applications, because of the development
 41 of digital seismology in the 1970s and the focus on mathematical processing with computers, largely bypassing
 42 the direct analysis of seismograms by human observers.

43 Auditory display as a scientific research field has grown considerably during the last decade¹⁷. Auditory
 44 representation of data has proven efficient for *e.g.* solar wind ion composition¹⁸, stem-cell classification¹⁹,

45 recognition of patterns in stock market data²⁰, or in the physiological processes of trees²¹. A previous article²²
46 has shown that the free categorization of audified seismic signals conducted by listeners is consistent with some
47 geophysical parameters (distance between epicenter and recording station, Earth's structure). If it seems now
48 to be accepted that auditory display can complement visual display contributing to the interpretation of sci-
49 entific data, further investigations are necessary to more precisely identify the structures of *a priori* cognitive
50 representations that are involved when humans are exposed to auditory display, and the properties that are
51 processed when the data (earthquake recordings in our case) are transformed into acoustic signals.

52 While the classification freely performed by listeners has been shown to agree with some conceptual repre-
53 sentations of geoscientists²² (categories made by listeners match categories based on geophysical parameters),
54 it is now necessary to access the "intensional" definition of the categories, *i.e.* to identify the criteria used for
55 the categorization. The aim of the present article is to show that cognitive representations of listeners, when
56 extracted using proper testing and analysis methods, can be used as heuristics in order to identify relevant
57 features for the discrimination and characterization of seismic signals. Thus at this stage we are not claiming
58 general results about how humans perceive sonified seismic data (for this reason, statistical analysis would not
59 be appropriate for this study), but rather at exploring the ability of some expert listeners (in sound *per se*, or
60 in geophysics) to bring novel description of the data that can be used by geoscientists.

61 Following a first experiment²², here referred to as T1, investigating the effect of event/station relative loca-
62 tion, we apply the same approach in 3 further experiments to consolidate the results previously obtained and
63 acquire more precise knowledge on auditory categories for earthquake recordings, through the investigation of
64 other seismic parameters (magnitude in T2, audification speed factor in T3a and T3b). More importantly, while
65 the previous study was limited to similarity measurement (only based on co-occurrences of stimuli in categories),
66 here a thorough cognitive analysis is conducted on the verbal data collected at the end of the tests. This ap-
67 proach allows us to get at the relevant characteristics of the stimuli mentioned by the listeners, and therefore
68 to guide our exploration of the seismic data. The verbal analysis, and its use to access and understand the
69 categorization criteria, constitutes the original contribution of the present article with respect to the previous
70 one. Following an inductive approach, the analysis of the comments associated to the categories is used to access
71 and understand the categorization criteria, common or different between ensembles¹ of subjects (acousticians
72 or geoscientists, the first ones being trained in listening and analyzing any acoustic signal as such, the second
73 ones being experts in earthquakes). These criteria are further used to elicit and suggest relevant parameters for
74 the description of the categories in terms of physical parameters.

75 The second contribution of this article is to take into account the expertise of the listeners. This question has
76 been widely discussed in the literature, mostly for musical expertise and exposure to familiar sounds: Trained
77 musicians and non-musicians were shown to have similar results on musical processing tasks²³ (although trained
78 musicians' answers are more accurate²⁴), mostly because both groups have been exposed to music on a every-
79 day basis. Yet it is known that in sorting tasks the expertise of listeners can change the way the categories are
80 formed²⁵ or the level of categorization²⁶. Furthermore the focus is here on sounds any of the two ensembles of

81 participants has never been exposed to, and the question of the use of prior knowledge (either on sound or seis-
 82 mic data) on such signals remains open. The previous knowledge involved in subjects' perceptual processing is
 83 investigated here through a subject-centered approach of cognition and categorization as "acts of meaning"^{27;28}
 84 (*i.e.* "the nature and cultural shaping of meaning-making, as the central place it plays in human actions"²⁹).
 85 For that purpose we contrasted two ensembles of subjects exposed to the "same signals": "geoscientists" — ex-
 86 perts in visual analysis of seismograms, but not trained in processing (earthquakes as) acoustic signals — , and
 87 acousticians — experts in acoustic signal processing but without background in seismology. In this situated
 88 approach of cognition, categories resulting from individual sensory experience are not conceived as "information
 89 processing" filtered by the human senses but as a meaning-making process involving different types of knowl-
 90 edge³⁰, among which individual experience, knowhow, academic and scientific knowledge. The exploration of
 91 sensory categories cannot therefore rely only on the scientific knowledge of the world (as given by geoscience
 92 or acoustics in our case), but has to identify the categories as sets of properties making sense to the user ("*ad*
 93 *hoc* categories"^{31;32}). Such categories as individual cognitive constructions not only include perceived physical
 94 characteristics (bottom-up processes, signal processing), but also memorized properties (top-down processes,
 95 signal interpretation) in context (*i.e.* depending on the subject's goal, cognitive orientation and attention, and
 96 expertise)². Within this theoretical framework, the physical characteristics of the stimuli as defined in terms of
 97 dimensions may be psychologically meaningful only if relevant for discriminating categories. For example cate-
 98 gories of everyday sounds are not structured along the dimension of intensity as an independent variable but in
 99 close interaction with the source identification³³, categories for soundscapes are not structured along intensity
 100 as a physical (abstracted) dimension of the acoustic signal but remain embodied within the experience, concern,
 101 and identification of the source³⁴.

102 A major empirical consequence of this theoretical positioning is that the physical description cannot be used
 103 as the reference for evaluating the human categories as deviations "errors" from the "true" representations given
 104 by physics. Therefore at this exploratory step of research, we stand apart from the psychophysical paradigm
 105 which attributes to the physical description the referential value in defining *a priori* what information is to be
 106 processed by humans³⁰. We rather focus on a subject-centered paradigm, in that we explore the ability of dif-
 107 ferent expert listeners to bring their specific descriptions of the data. The present approach is inductive, aiming
 108 at providing new hypotheses for future hypothetico-deductive studies, which would consolidate the hypothesis
 109 through further and more canonical experimental setting including statistical analysis. Again, statistical anal-
 110 ysis is not relevant to our approach.

111 As in the previous paper²², the audio stimuli result from a time-compression of seismic signals (this method
 112 called "audification" is a particular case of sonification^{17;35}), and are presented to over headphones. Note that
 113 the seismic recordings themselves and even more so the audified signals are observable reproductions of concep-
 114 tual representations elaborated from present day scientific knowledge and technologies: We implicitly assume
 115 some adequacy between these representations and the vibration of a point on the Earth's surface but it must
 116 be kept in mind that we actually deal with a specific representation of a complex phenomenon.

117 In the rest of the paper, Sec. II describes the database and the production of stimuli, Sec. III describes the
118 experimental method, Sec. IV describes the analysis method for the categories and their verbal description, and
119 Sec. V present the results as perceptual descriptions of the categories of stimuli. Building on the perceptual re-
120 sults, acoustic descriptors are eventually computed in order to match the verbal descriptions: they are presented
121 in Sec. VI.

122 II From seismic signals to audio stimuli

123 A The database

124 The database used in this study consists of broadband recordings (sampling frequency $F_{s,0} = 40$ Hz, recording
125 stations of the USArray experiment³⁶, all stations with nearly identical mechanical characteristics and spectral
126 sensitivity) of the Earth’s oscillations, made at the locations depicted in Fig. 1, of a sequence of 40 injection-
127 triggered (“hydrofracturing”^{37;38}) earthquakes (or “seismic events”) in Oklahoma that occurred in November
128 2011. Fig. 1 is a topographic map of the study area, showing the location of the stations and epicenters, as well
129 as the Earth’s elevation. Fig. 1 shows that central and western stations are located on relatively flat terrain,
130 whereas northern-eastern and northern-western stations have a higher elevation. Topography differences are
131 indicative of crustal structure heterogeneities^{22;39} (mechanic properties of rocks, crustal thickness, *etc.*) so that
132 this region is not seismologically homogeneous. In order to give a more synthetic view of the seismic events and
133 stations, Fig. 2a presents a map of the seismic stations used in this study, and Fig. 2b presents a map of the
134 earthquakes of the database.

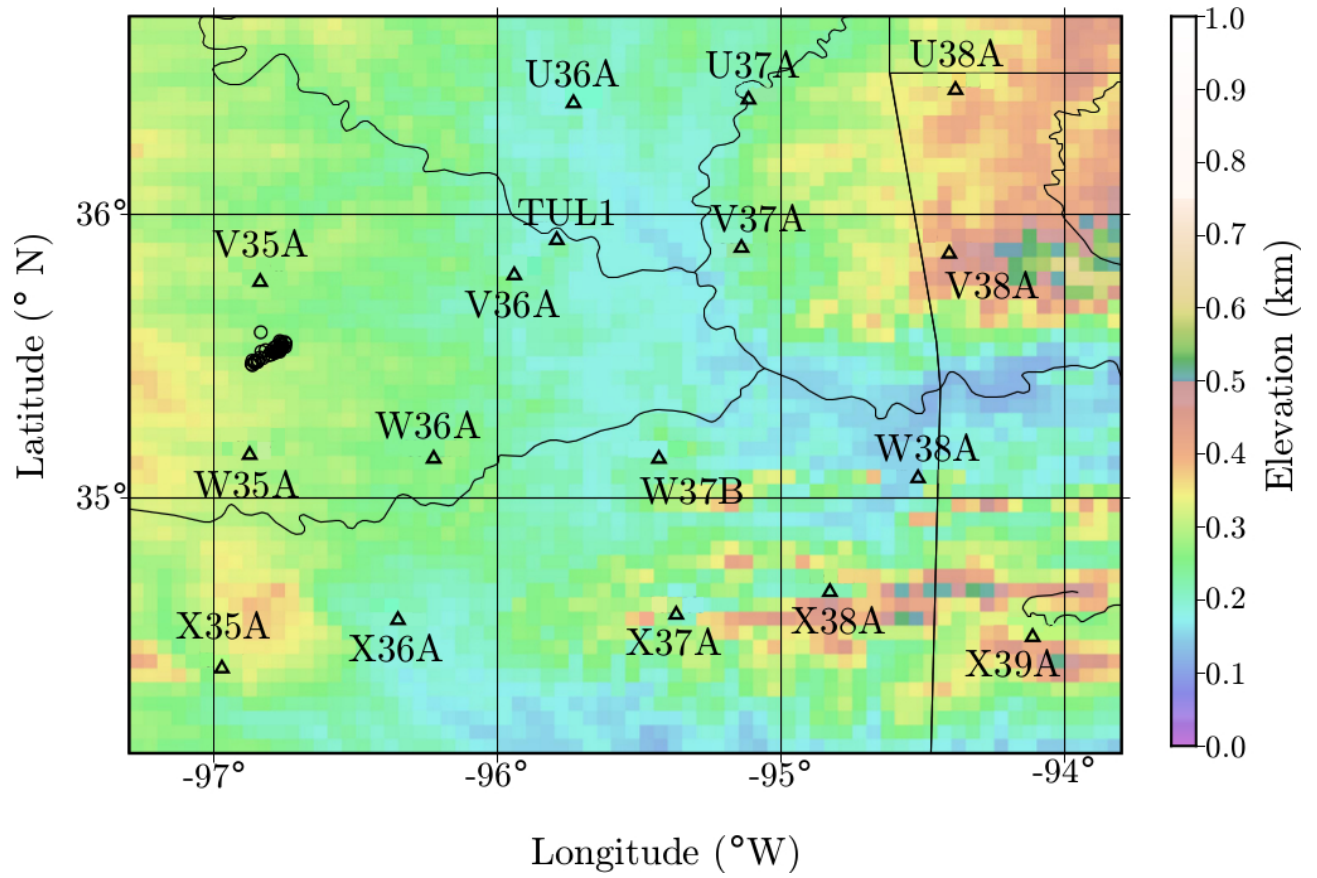


Figure 1: (color online) Topography of the study area. Black triangles denote available seismic stations, which are labelled. The color scale corresponds to the elevation of the Earth's surface with respect to sea level. Black circles on the left side denote the epicenters of the seismic events of the database.

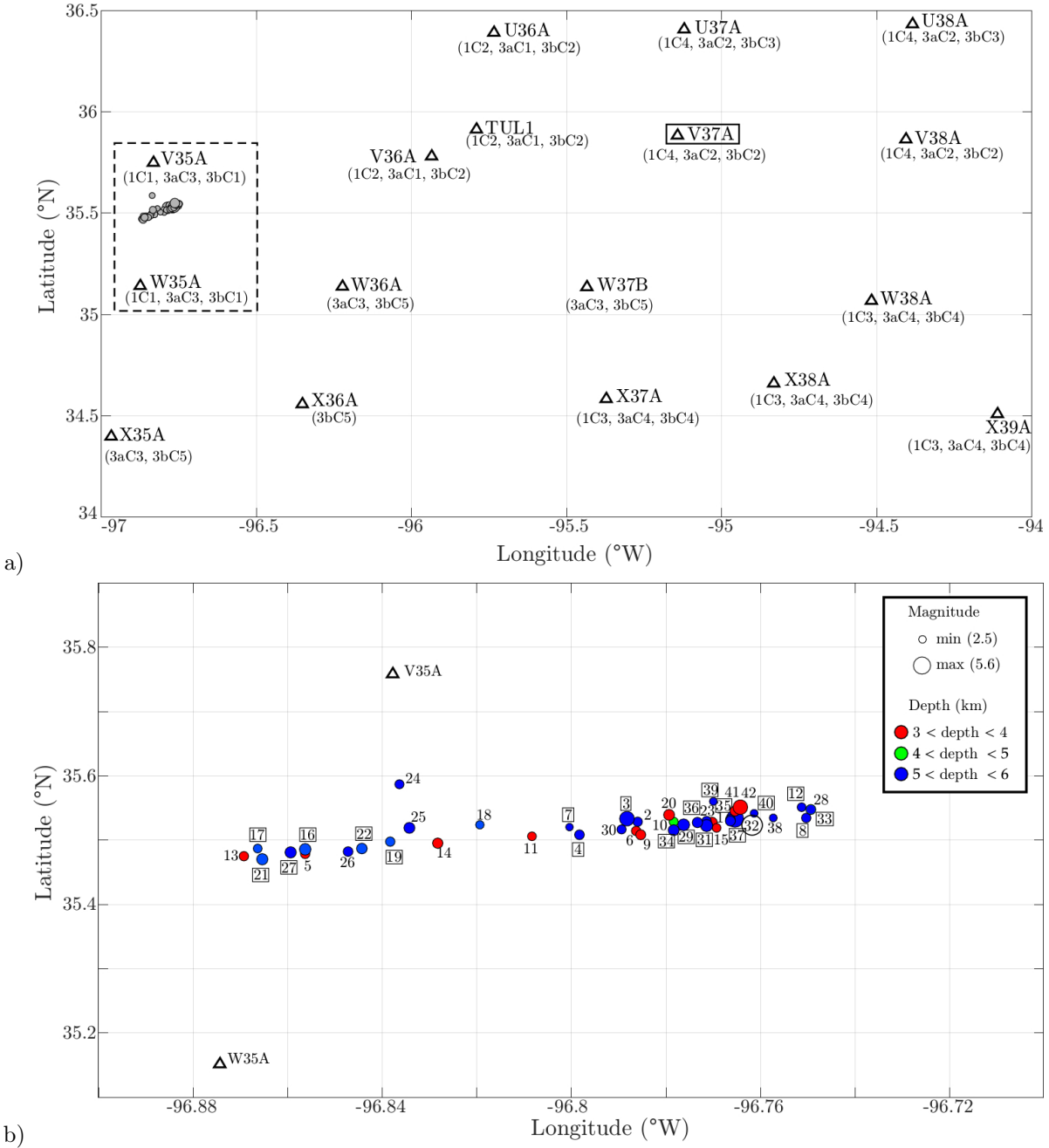


Figure 2: (color online) Map of the seismic stations (black triangles and names, followed by the consensual categories in which the resulting audio signals are put in T1, T3a and T3b, see Sec. V for more details) and the seismic events of the database (filled circles: the size is proportional to the magnitude). a) Global view, the selected station for T2 is enclosed in a solid line rectangle, the dashed line rectangle indicates the area of the epicenters, magnified below; b) Magnified view centered on the seismic event epicenters, the selected event for tests T1, T3a and T3b is circled, the selected events for T2 are enclosed in a rectangle. The events are numbered in decreasing order of appearance (1 is the latest, 42 is the earliest) during the 4 days of recording. Different colors indicate the depth of the events.

135 B The stimuli

136 The recording stations are 3-component sensors, measuring ground vibrations in the vertical (normal to the
 137 Earth’s surface) and the two orthogonal horizontal directions. The audible acoustic waves are unidimensional;
 138 only the vertical component of the seismic recordings is investigated. Based on the similarity in the nature of
 139 seismic and acoustic signals (zero-mean, decreasing amplitude), the most direct sonification method is used,
 140 that is “audification”. In the present case, the inaudible content of seismic recordings has to be translated to
 141 audible range. Audification then consists of playing the recorded samples at a quicker rate, which is implemented
 142 through an increase of the sampling frequency ($F_s > F_{s,0}$, the ratio $F_s/F_{s,0}$ is the speed factor). The dynamic
 143 range of seismic signals is much wider than that of audio signals, so that, in practice, signals associated with
 144 lower-magnitude events in our database would be too quiet to be heard; audified signals are therefore all
 145 normalized, each with respect to its maximum amplitude value. This means that some of the information that
 146 is contained in the signals, and that in principle could contribute to their auditive interpretation, is lost. The
 147 related issues are addressed below, in our discussion of experimental results. Four free sorting tasks are carried
 148 out. The tests are numbered T1, T2, T3a and T3b according to their order of presentation.

149 **Stimuli for T1** The variable is the event/station relative location (distance and azimuth of the station with
 150 respect to the event), so recordings of the same event by 17 stations are used. In order to maximize the signal-
 151 to-noise ratio the seismic event with the highest magnitude is selected: Event number 32 (magnitude 5.6, depth
 152 5.2 km, circled in Fig. 2b.) Recordings of this event from the stations plotted in Fig. 2a are audified with a
 153 speed factor of 150 ($F_s = 6,000$ Hz). All 17 signals are trimmed so as to obtain a duration of 2s for the resulting
 154 audio stimuli.

155 **Stimuli for T2** The variable is the magnitude of the earthquakes. Twenty-two seismic events out of 42 are
 156 selected since they have the same estimated epicenter depth (5 km) and present magnitudes ranging from 2.5 to
 157 4.8, which is typical of the database (magnitudes of the 42 events range from 2.5 to 4.8 with only one event with
 158 higher magnitude, that is 5.6). The seismic recordings used are from station *V37A*, located at an intermediate
 159 distance from the events. The seismic signals are audified with a speed factor of 150 ($F_s = 6,000$ Hz).

160 **Stimuli for T3a** The selected stations and event for T3a are the same as in T1, except that the seismic
 161 recordings are audified with a speed factor of 250 ($F_s = 10,000$ Hz).

162 **Stimuli for T3b** The selected stations and event for T3b are the same as in T1, except that the seismic
 163 recordings are audified with a speed factor of 350 ($F_s = 14,000$ Hz).

164 Another test variable must be made explicit here: The role of previous knowledge and expertise is investigated
 165 in all 4 tests by having 2 ensembles of subjects (geoscientists and acousticians) involved. Table 1 summarizes the
 166 effects (station location, earthquake magnitude, sonification speed factor, expertise of listeners) tested in each

167 test, either as a variable or as a fixed effect. Tests T1, T3a and T3b, presenting the same stimuli with different
 168 levels of the speed factor as a fixed effect, can be directly compared. Additionally and as a guide for the readers,
 169 all stimuli are available online³. In general, they can be roughly and informally described as a gunshot-like
 170 sound with decay over a broadband background noise. This decaying part is called the “coda”, and is known by
 171 geoscientists as containing all the information about wave propagation, path, scattering, attenuation.

172 III Experimental procedure

173 Classic psychophysical methods mainly rely on exclusively bottom-up models of stimulus processing, and only
 174 involve stimuli that are controlled and designed along independent physical parameters and that determine the
 175 perceptual answers, measured along dimensional indices. These methods are challenged by the two following
 176 observations. First, human perception is influenced by bottom-up (signal-driven) processes but also by top-
 177 down processes which depend on the memorized knowledge and expertise of the participants. Second, since no
 178 previous experiment dealt with the perception of audified seismic data except for our previous study²², it is
 179 impossible to *a priori* decide what acoustical parameters are relevant for exploring the psychological processing
 180 (the study precisely aims at discovering it). The free sorting task^{40;41;42;43;44;45} is chosen in this study, because it
 181 can address these two remarks. The contribution of participants’ previous knowledge is explored by contrasting
 182 two ensembles of subjects.

183 A Theoretical background

184 Following Rosch’s seminal work⁴⁶ on the structure of so-called “natural” categories, we aim to identify both
 185 the extensional structure of the categories as the list of their members, as well as their intensional structure as
 186 sets of properties defining the categories. Unlike well-defined categories as elaborated in scientific knowledge
 187 (with clear-cut binary membership, *i.e.* an item is either a member of the category, or is not a member), the
 188 extensional structure of natural categories is defined by similarity and distance from a prototypical exemplar.
 189 The intensional description relies on this prototype, which is defined as the exemplar gathering most of the
 190 properties of the category. The other exemplars (stimuli) are distributed along similarity (“family” resemblance)
 191 within the set of properties that they “more or less” share with one another. If a lot of psychological literature
 192 has been devoted to developing various models of categories constructed along prototype and similarity^{47;48},
 193 previous research has been mainly concerned with acquired and shared established knowledge on different objects
 194 but has more rarely dealt with experiential knowledge and individually-constructed categories. Participants in
 195 this study are either acousticians or geoscientists (see Sec. B), having different education and knowledge: if
 196 acousticians are trained to describe sounds as objects *per se*, geoscientists process the “same” sounds as acoustic
 197 representations of seismic signals (as “sounds of”). The question is to identify how this difference in expertise
 198 influences the categories and the categorization criteria.

199 B Free sorting experiment

200 1 Participants

201 Each participant is presented with tests T1, T2 and either T3a or T3b, always in this order. 24 participants
 202 (15 acousticians, 9 geoscientists) took part in T1 and T2. From these 24 participants, 11 (8 acousticians, 3
 203 geoscientists) in T3a and 12 (6 acousticians, 6 geoscientists) in T3b. One participant (acoustician) did not take
 204 part in T3. People in the ensemble “acousticians” are either faculty/staff of the LAM team at the d’Alembert
 205 Institute or professional sound technicians. People in the ensemble “geoscientists” are faculty/staff of the Earth
 206 Sciences Institute ISTE \P . Note that the time and availability constraints did not allow us to have as many geo-
 207 scientists as acousticians participating in the test. Note also that the assignment of T3a or T3b to a participant
 208 was randomized, resulting in a non-balanced number of acousticians and geoscientists for T3b. Note that at this
 209 stage of investigation the goal of the research is to find out whether such a differential approach is contrasting
 210 different ensembles of subjects is productive even with a small number of subjects, for the potential development
 211 of further research with more quantitative data allowing statistical computation.

212

213 2 Procedure

214 The instructions given to the participants for each test are as follows⁴:

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

218 Each of the N stimuli has to belong to only one group. The participants are free to form as many groups as
 219 they want and can put any number of stimuli in a single group. The participants are told that the stimuli
 220 originated from seismic recordings, but no other information on their nature is given. Only the assignment
 221 of each stimulus to a group is taken into account in the analysis: The spatial arrangement of the groups and
 222 the icons within the interface area is neglected. At the end of the sorting, each participant is asked to type a
 223 comment for each group he/she made. The categorization and the verbal description provide complementary
 224 characterizations of the stimuli: descriptions allow us to identify the characterization of stimuli as properties
 225 along which the categorization has been processed.

226 3 Experimental setup

227 The tests are run on a laptop equipped with a *RME Fireface UCX* soundcard. The stimuli are played back
 228 through *Sennheiser HD380 Pro* headphones. Audio stimuli are monophonic, each ear being exposed to the same
 229 signal, in phase. The participants can set and change the sound level in the headphones at any time during the
 230 test (but no participant did it). The *TCL-LabX*⁴⁹ software is used for the free sorting interface. The graphic

231 interface displays each stimulus as a small square icon. Illustrations of the graphic interface and of the test
 232 setup are given in the previous study²². The N icons for a test with N stimuli are randomly numbered from 1
 233 to N . A double click on an icon launches the stimulus playback, and the icon can be moved within the entire
 234 interface area with a click-and-drag operation. Each stimulus can be played back as many times as wanted.

235 4 Output and duration

236 The output of each test, for each participant, is referred to in the following as a “partition;” it consists of a suite
 237 of groups of stimuli, accompanied by a verbal description of each group in this partition.

238 The mean duration for test T1 was 13.6 ± 9.1 minutes, and for T2, T3a and T3b respectively 9.3 ± 3.7 min.,
 239 5.3 ± 1.5 min. and 5.3 ± 2.1 min.

240 IV Method of data analysis

241 The data analyzed are of two kinds, and as such impose different types of processing in order to evaluate
 242 their robustness and interpret them within the theoretical framework discussed in Sec. A: Individual partitions
 243 resulting from the sorting task (Sec. A), and b) verbal comments for each individual partition (Sec. B).

244 A Categories of stimuli

245 Partitions reflect similarities and differences between stimuli as evaluated by the subjects: Stimuli within a
 246 category are more similar to one another than stimuli sorted into different categories. Individual partitions are
 247 added up, with the number of subjects grouping together a certain pair of stimuli functioning as a metric of
 248 the similarity between those two stimuli. It is very important to note that this measure of similarity a) relies
 249 on the consensus between subjects, and b) processes stimuli as whole and indivisible items. In other words
 250 we get a representation “in extension” of the categories (*i.e.* an explicit list of its members). Consequently a
 251 projection of the stimuli on a multidimensional space is not necessary relevant, before further investigations
 252 of the “intensional” description of the stimuli (*i.e.* as sets of dimensional properties or other characteristics).
 253 It is therefore not adequate to use statistics relying on distributions of values on dimensions and Gaussian
 254 assumptions, and we prefer other mathematical metrics developed through the close collaboration between
 255 mathematicians and psychologists in order to account for classification analysis^{50;51;52}.

256 Only a concise description of the mathematical method of analysis is given here: A more detailed description
 257 can be found in the literature^{22;41;45;50;53;54;55}. For each test, a perceptual distance between stimuli is defined
 258 as follows:

259 1. A co-occurrence matrix M^k is defined and computed for each participant ($k=1, \dots, N_s$, where N_s is the
 260 number of participants). M^k is a square matrix of size N , where N is the number of stimuli:

- 261 • $M_{ij}^k = 1$ if stimuli i and j are in the same group according to participant k ,

262 • $M_{ij}^k = 0$ if stimuli i and j are in different groups according to participant k ;

263 2. The total co-occurrence matrix is computed: $M_{ij} = \sum_{k=1}^{k=N_s} M_{ij}^k$ (the more often stimuli i and j are
264 grouped together, *i.e.* the more subjects having grouped them together, the larger M_{ij});

265 3. The distance matrix D is defined as: $D_{ij} = 1 - M_{ij}/N$ (the more often stimuli i and j are grouped
266 together, the smaller D_{ij} ; $0 \leq D_{ij} \leq 1$).

267 The values in D are “consensual” measures of perceptual distance between stimuli, *i.e.* they represent a consensus
268 between the participants. They can be represented by an additive tree⁵⁰: the length of the branches (connecting
269 the leaves, or vertices, representing the stimuli) is proportional to the perceptual distance between stimuli.
270 Branches aggregate at “nodes”, enabling to consider categories at different levels of generality/inclusion. The
271 orientation of the branches is arbitrary, only the distance along branches matters. The distances in D are fitted
272 to an additive tree distance by means of the *Addtree* software⁵⁶.

273 The resulting trees are represented in Figs. 3 through 6 (in Sec. V) for each test. They take account of
274 data from all participants (acousticians and geoscientists). On the trees, the consensual categories are identified
275 visually as the most compact clusters of leaves/stimuli. These identified consensual categories are circled in
276 Figs. 3 through 6, and numbered for clarity (numbers are arbitrary). Note that the identification of these
277 clusters depends on the experimenter^{22;41;45} and might slightly change for another experimenter, however we
278 believe the visual identification of consensual categories to be robust enough for our purpose. The same analysis
279 is conducted separately for acousticians and geoscientists (the trees from separated ensembles of participants
280 are not shown here for brevity, but are available as supplementary material⁵. There is no major differences
281 between the categories of both ensembles of participants in the structural properties of the categories, but
282 there are differences in their verbal descriptions, interestingly showing that participants process along the same
283 bottom-up constraints but conceptualized in a different way that the verbal analysis will make explicit (see
284 Sec. A).

285 B Verbal comments

286 In this section the verbal descriptions of the consensual categories identified on the trees in Sec. A are analyzed.
287 The method has already been applied in the literature^{22;45;57;58}. Different ensembles of participants are formed
288 and independently studied: all participants (“all”), “acousticians” and “geoscientists”. In the two latter cases,
289 the consensual categories described are not those shown in Figs. 3 to 6 but the ones computed from the total
290 co-occurrence matrix of all acousticians and the total co-occurrence matrix of all geoscientists respectively.

291 As far as the verbal comments are concerned, our analysis relies on a “differential” conception of lexical
292 semantics, that considers that the meaning of a word (as a “lexical form” or “significant”) is not given by its
293 referential value (a label on pre-existing things it refers to in the physical world), but mainly relies on a consensus
294 between speakers to attribute its form to something or to a concept. For example in scientific discourses, scientific

295 concepts are named by “terms”, *i.e.* words whose meaning is an explicit rectification of a common sense word
 296 negotiated and accepted in the scientific community. Same words used in different communities may have
 297 different meanings⁵⁹, and individual variations in the meaning assigned to words are a common and well-known
 298 phenomenon. In other words, variations in meaning attributed by different individuals cannot be considered as
 299 errors with respect to a “true” meaning, but rather as data the analysis has to account for. It follows that it is,
 300 at least, problematic, to apply classical statistics (*e.g.* averaging, test of significant differences) to our data, and
 301 we refrain from doing so in this study. Furthermore, there is too little verbal data to undertake any statistical
 302 analysis.

303 In verbalization tasks, expert participants (*e.g.* expert guitarists talking about how they perceive their
 304 instrument⁴⁵) have a specific use of lexicon, assigning to words different meanings than when they are used by
 305 non-experts or in a generic context. In those cases, it is necessary to undertake a linguistic analysis in order
 306 to identify the semantics of the words, using linguistic clues such as reformulations, definitions, *etc.*, present
 307 in the participants’ discourse when they are invited to describe their sensory experience of the stimuli. In the
 308 present case, the instructions are oriented on the “objectivity” of the stimuli (participants are told that the
 309 stimuli originated from seismic recordings) that encourage a more straightforward naming by the use of simple
 310 and common words. Furthermore the constraints associated with typing prevent participants from producing
 311 long sentences they might utter if speaking⁶⁰. Even in lack of substantial discourses as it is the case here it
 312 is worth keeping in mind the different background of the two ensembles of subjects. Due to the constraints of
 313 the task discussed above, geoscientists and acousticians are expected to produce short statements using simple,
 314 everyday-life words, with meanings slightly differing from the “common sense” meaning, as given by a dictionary.
 315 Geoscientists are expected to use common sense meaning for words referring to the sound itself but technical
 316 terms for referring to the sound as the “sound of an earthquake”, whereas acousticians are expected to use
 317 technical terms for the sound itself and common sense words when referring to the sound as the sound of an
 318 earthquake. In other words, the subjects all share the same language and culture, but slight differences in
 319 educational background and training may change the way they conceptualize and therefore verbally describe
 320 sounds.

321 The analysis of verbal comments⁶ is carried out by first organizing them according to the aspect of the
 322 stimuli they refer to. Words are grouped into semantic classes, which are labelled by a word picked from
 323 the corresponding class⁷. The comments can be split into those referring to: number of perceived *impacts*⁸
 324 in the stimuli (semantic class IMPACTS), frequency content of these impacts (semantic class FREQUENCIES⁹);
 325 *duration* or *speed* of the stimulus (semantic class DURATION / SPEED); distance from the presumed source of the
 326 stimulus (semantic class DISTANCE). Other aspects are identified and split into the classes REVERBERATION¹⁰
 327 (references to the part of the audio stimuli after the impacts), PERCUSSIVE (*sharp* and *clear impacts*), BOUNCING
 328 SOUND (chirp-like sound sometimes occurring after the impact), BASS BACKGROUND (*large amount of very low*
 329 *frequencies* before the *impact*; note that this does not describe the impacts, hence it is separated from the class
 330 FREQUENCIES), DEEP (referring to the supposed, perceived depth of the seismic event), AGGRESSIVE (words

331 sometimes referring to a very strong high-frequency content, but always bearing a judgment of unpleasantness),
 332 CRACKLING NOISE (crackings added to the background noise), and VOLUME (the perceived loudness of the
 333 stimulus). The semantic classes as well as the words assigned to these classes are presented in Tabs. 2, 3, 4 and
 334 5 (original French in normal font, lexical units separated by commas, English translation in italic font).

335 Verbal descriptions of the consensual categories defined in Sec. A are next formed by associating with them
 336 the comments made by each participant back to her/his own actual groups of stimuli. As individual groups often
 337 slightly differ from the consensual categories, a threshold has been defined: We gather in the verbal description
 338 of a consensual category the comments associated with each individual group sharing more than half of its
 339 stimuli with the consensual category. For example, if 2 out of 3 (or 3 out of 4, 3 out of 5, *etc.*) stimuli are
 340 common to an individual group and to the consensual category, the comments of this individual group are added
 341 to the verbal description of the consensual category.

342 The third and last step of our analysis consists of summarizing the lists of words formed above into a synthetic
 343 verbal description for each consensual category^{45;57}. For this purpose words in each semantic class are gathered
 344 into subclasses labelled by simple lexical units: *e.g.* in the semantic class IMPACTS, words *impacts*, *strokes* and
 345 *shots* are assumed to refer to the same aspect of sound, summarized by the label IMPACTS; in the semantic
 346 class FREQUENCIES *bass*, *low-frequency* and *dull* are summarized by the label “*bass*”. The number of comments
 347 under each subclass is then counted: A positive number (+1) is assigned to a comment showing the presence
 348 of the corresponding label, or sound aspect (*e.g. with reverberation* for the label/subclass REVERBERATION); a
 349 negative number (-1) is assigned to a comment showing the absence of the corresponding label (*e.g. without*
 350 *reverberation* for the label/subclass REVERBERATION). These numbers are then added together for each label.
 351 Note that absolutely no value judgement is made by the authors when using positive and negative numbers: this
 352 only depends on the arbitrary choice of the label word. Tabs. 6, 7, 8 and 9 show the presence or absence of the
 353 various sound aspects evaluated for tests T1, T2, T3a and T3b respectively. A positive (resp. negative) number
 354 indicates that the sound aspect in question is present (resp. absent) according to the majority of evaluations.
 355 The categories resulting from the comments of all participants are numbered for ease of reading.

356 V Results

357 This section first provides a general comparison of acousticians’ and geoscientists’ categories. Next, the results
 358 of the consensual categories and the verbal descriptions for each test are presented. The section ends with a
 359 discussion about the playback speed of the stimuli.

360 A How acousticians and geoscientists differ

361 One first result is that that acousticians and geoscientists differ both in the words they use and in the objects
 362 they refer to. This is evidence for the fact that listeners make use of their experience, memory and knowledge
 363 in their interpretation of the stimuli. Acousticians interpret, describe and name the stimuli as acoustic objects,

364 which they are used to listening to analytically (searching for “acoustical similarities”²⁵); whereas geoscientists
 365 interpret, describe and name the stimuli as cues referring to geophysical processes that make sense for them
 366 based on their knowledge (searching for “causal similarities”²⁵). This difference in conceptualization can be
 367 inferred from the following results:

368 **Number of words used** Acousticians use more words than geoscientists (average of 105 and 49 words per
 369 ensemble). This indicates that acousticians are trained to describe acoustic signals and have richer vo-
 370 cabularies available (not only common words they share with geoscientists such as *echo* but also technical
 371 acoustic terms such as *reverberation*). Similar observations can be found in the literature²⁵.

372 **Object described** It is interesting to note that when giving a description of the stimuli, acousticians tend to
 373 use the word *sound* in plural rather than in singular form (62.2% of the occurrences of *sound* are in plural
 374 form), whereas geoscientists prefer to use the singular form (only 17% of the occurrences of *sound* are in
 375 plural form). This means that acousticians identify different sounds in the one stimulus, and process in
 376 an analytical mode; whereas geoscientists process the stimulus as the sound representation of one seismic
 377 event, a meaningful unitary object for them.

378 **Words used** Geoscientists use a different vocabulary than acousticians, who are experts in describing sounds.
 379 For instance, geoscientists describe the impacts as *impact* or *sound*, as opposed to the expressions *impact*,
 380 *stroke*, *attack*, *wave*, *shot* employed by acousticians. According to their expertise, working customs and
 381 training, acousticians are more precise in their description of the spectral content, distinguishing medium
 382 frequencies from low-medium and high-medium, which geoscientists do not do. In the description of the
 383 BASS BACKGROUND, acousticians show a more accurate ability to describe the noise (*extremely low fre-*
 384 *quencies*, *sub basses*, *constant bass*, *continuous low-frequency sound*, *humming*, *<100Hz*, *throbbing*, *rumble*,
 385 *rolling*, *dull / deep / bass / low background*, *low-frequency noise*), while geoscientists are more succinct
 386 (*low-frequency background*, *rumble*, *vibrations*).

387 **Aspects of sound** Geoscientists and acousticians focus on different aspects of the stimuli. On the one hand,
 388 geoscientists do not make use of the word *balanced* (spectral balance between the frequencies), and do
 389 not write about the DISTANCE or the DEPTH of the earthquakes (presumably because those words refer to
 390 precise parameters of seismic events, which geoscientists felt that they were not able to estimate from the
 391 presented stimuli). Furthermore, the chirp-like sound sometimes occurring after the impacts (BOUNCING
 392 SOUND) has not been mentioned by the geoscientists. On the other hand, acousticians do not use the
 393 loudness of the stimuli (class VOLUME) as a categorization criterion.

394 **Selection of evaluation criteria** While acousticians evaluate systematically the same sound aspects for each
 395 test (apart from the DURATION / SPEED in T3b), geoscientists adjust their evaluation criteria depending
 396 on the test and therefore geophysical relevant parameters (BASS BACKGROUND evaluated only in T2 and

397 T3b, PERCUSSIVE only in T3a).

398

399 The above analysis is complicated by the sample size, which is small from a statistical standpoint, and by the
 400 slight disparity in the size of the “acousticians” and “geoscientists” ensembles. Nevertheless both ensembles
 401 produced comments that are relevant for distinguishing the consensual categories of stimuli, and hence attempt
 402 to reconstruct the categorization criteria. The next paragraphs summarize the differences between consensual
 403 categories for each test, focusing not on the similarities but on the differences, in particular when categories are
 404 contrasted according to a criterion.

405 B T1 (variable: event/station relative location, speed factor: 150)

406 All identified consensual categories in the tree in Fig. 3 group together stations which are close to one another²².
 407 Information about the categorization criteria is given by the verbal description associated with each consensual
 408 category (Tab. 6).

409 Consensual categories are first formed according to the perceived number of *impacts* in the stimuli. The
 410 physical interpretation of this number of *impacts* is quite straightforward: the longer the propagation distance
 411 between the event and the station, the more the P- and S- wave (travelling at different speeds) are temporally
 412 separated²². To a lesser extent the categorization relies on the event/station azimuth (*i.e.*, the propagation
 413 path between the event and the station). More in details, category 1C1 (stations closest to the event) includes
 414 stimuli where only one *impact* is perceived. Category 1C2 (grouping stations located at an intermediate dis-
 415 tance, North-East from the epicenter) includes stimuli where 2 *impacts* are distinguished but found to be close
 416 or very close one to another. Categories 1C3 and 1C4, consisting of stations further away, respectively located
 417 South-East and North-East from the epicenter, include stimuli where 2 *impacts* are perceived to be clearly
 418 separated.

419 A second criterion on which categorization relies is the frequency content, which can be related to the
 420 event/station azimuth. The spectrum of stimuli of categories 1C2 and 1C4 (North-East from the event) are
 421 perceived to have more *treble* and *medium frequencies*, whereas the frequency contents of categories 1C1 (close
 422 to the event) and 1C3 (South-East from the event) are respectively more in the *medium* to *low frequencies*.

423 A third criterion is related to the mention of the perceived *speed* and *distance* from the sound event(s),
 424 which can be related to the event/station azimuth, just as the evaluations of the FREQUENCIES. Sounds from
 425 categories 1C2 and 1C4 (both categories North-East from the event) are perceived as *fast* and *near*, whereas
 426 stimuli from category 1C3 (South-East from the event) are perceived as *slow* and *far*. It is possible that the
 427 perceived *speed* and *distance* refer to different frequency contents: further studies dealing with more complete
 428 verbalizations may make this point clearer.

429 Other criteria are not relevant to distinguish categories, but do contribute to the description of specific cate-
 430 gories. For stations North-East from the event, stimuli are described as more *percussive* when the event/station

431 distance increases (see the *percussive* evaluation of 1C1, 1C2 and 1C4). No straightforward geophysical inter-
 432 pretation can be provided for now for the fact that categories 1C2 and 1C3 are often described as having a *bass*
 433 *background* (this aspect is evaluated by acousticians only, so it might not be a relevant parameter, geophysically
 434 speaking), while categories 1C3 and 1C4 are characterized, among other things, as having *reverberation*.

435 It could have been hypothesized that the event/station distance is related to the perceived intensity or
 436 loudness of the stimuli. However, the stimuli had to be normalized according to amplitude, presumably making
 437 loudness differences so subtle that they could not be used as a relevant and discriminative criterion.

438 It should be noted that the categories of acousticians only, geoscientists only, and all participants are quite
 439 similar. Geoscientists and acousticians are able to produce relevant (*i.e.* allowing us to discriminate between
 440 categories) evaluations of the number of impacts and the spectral content, but the other aspects are explicitly
 441 considered by acousticians only.

442 C T2 (variable: magnitude, speed factor: 150)

443 The tree in Fig. 4 does not show clusters as clearly as the tree of T1 (Fig. 3), indicating a weaker consensus
 444 (all listeners generally tend to follow a broader range of different criteria than in T1, and acousticians and geo-
 445 scientists are less consistent with each other, as shown in Tab. 7). Note that the categories identified from the
 446 data of all participants are more similar to the categories made by acousticians, but simply because acousticians
 447 form the most numerous ensemble of participants.

448 A few consensual categories can be identified however on the tree in Fig. 4. We observe a tendency to group
 449 together stimuli coming from seismic events close to one another: Category 2C1 includes stimuli associated
 450 with seismic events East of latitude -96.8° , *E3* and *E34* (2C2) are at latitude -96.78° , *E22* and *E27* (2C3) are
 451 between latitudes -96.86° and -96.84° , *E4* and *E29* (2C5) are between -96.8° and -96.78° (see Fig. 2b). Excep-
 452 tions are categories 2C4 and 2C6, which include seismic events occurring on both sides of the cluster of seismic
 453 events. In general, the categories are not related to the magnitude, with the exception of 2C1, grouping the two
 454 lowest-magnitude seismic events, *E7* and *E40*, as well as two other seismic events of relatively small magnitude
 455 (<3.1).

456 The majority of the comments focuses on the presence/absence and level of the *bass background*, and stimuli
 457 in 2C1 have clearly a high level of *bass background*, whereas in all other categories the *bass background* is either
 458 very *low* or absent. Two interpretations can be proposed for the mention of the presence and level of this
 459 *bass background*. First, each audified seismogram is normalized with respect to its maximum amplitude: As
 460 a result the background level is raised for lower-amplitude events (with lower magnitude). Second, the (*bass*)
 461 *background*, as opposed to the number of *impacts*, depends on the medium through which the elastic waves
 462 travel. If one considers *e.g.* 2C1, the former explanation can apply because the two lowest-magnitude events
 463 (*E7*, magnitude 2.5; *E40*, magnitude 2.6, see Fig. 4) are included in this category, but the latter explanation
 464 can apply as well, since the events in 2C1 have an epicenter very close one to another.

465 Clearly in T2 the number of *impacts* is no longer relevant for discriminating stimuli. There are indeed only
 466 a few comments about the number of impacts. Furthermore, there is no clear consensus about the number of
 467 *impacts* in the stimuli of a category (*e.g.* it is not sure whether 2C2 has stimuli with 1 or 2 *impacts*, nor is it
 468 certain whether stimuli in 2C2 or 2C6 have clearly separated or very close *impacts*). This is not surprising to
 469 us, as here the variable is magnitude, and all stimuli are associated with one cluster of events very close to one
 470 another, and one (always the same) station.

471 Some differences can be seen in the frequency content of the stimuli: stimuli in 2C1 and 2C6 are perceived to
 472 have more *treble* and *bass*, stimuli in 2C4 and 2C5 are perceived to have less *bass*, stimuli in 2C3 are perceived
 473 to have more *treble* and no *bass*.

474 An occasionally perceived and mentioned *bouncing sound* allows us to make a difference between the cat-
 475 egories: 2C2, 2C3, 2C4 and 2C6 have stimuli with this *bouncing sound*, whereas 2C1 does not have. The
 476 geophysical origin of this *bouncing sound* is however unclear.

477 D T3a (variable: event/station relative location, speed factor: 250)

478 Tests T3a and T3b are aimed at determining whether perception changes when the audification frequency (speed
 479 factor) changes. This section and the following one (Sec. E) show the results obtained for tests with the same
 480 stimuli as T1 but with different speed factors, and Sec. F compares the results of T1, T3a and T3b in order to
 481 identify the influence of the speed factor.

482 As in the case of T1, each consensual category includes stimuli associated with stations located close to one
 483 another (Fig. 5). The relevance of the criterion IMPACTS (see Tab. 8) and the mapping between event/station
 484 distance and category is not as clear as in T1. Yet the event/station distance and azimuth remain the main
 485 criteria for categorization (in decreasing order of importance). More in details, there is no consensus between
 486 participants about the number of *impacts* heard in the stimuli of category 3aC1 (stations at an intermediate
 487 distance, North-East from the event), and for 3aC2 (stations at intermediate to long distance, North-East from
 488 the event) and 3aC3 (stations at close to intermediate distance from the event, spanning around it) the consensus
 489 is weak. Only stimuli of 3aC4 (stations far, South-East from the event), are perceived as certainly having 2
 490 (*clearly separated*) *impacts*. This difficulty at identifying the number of *impacts*, and therefore at grouping
 491 according to the number of *impacts*, may be explained by the change in the stimuli playback speed: Increasing
 492 the playback speed makes the *impacts* temporally much closer, making them harder to resolve.

493 The descriptions of the FREQUENCIES, although hard to interpret, are quite close to those provided in T1.
 494 Sounds in 3aC2 (intermediate to long distance, North-East from the event) and 3aC3 (short to intermediate
 495 distance) have a spectrum with *treble* frequencies and no *bass*, whereas stimuli in 3aC4 (far, South-East from
 496 the event) have more *bass*.

497 The categories also differ according to the classes CRACKLING NOISE, BASS BACKGROUND, AGGRESSIVE.
 498 Further investigations would be necessary to precisely interpret these observations, but this goes beyond the

499 scope of the present study. Furthermore, the relation between the event/station distance and the *percussive*
500 aspect of sound, as identified in T1, is no longer observed.

501 Like in T1, the consensual categories of acousticians only and geoscientists only are quite similar to the
502 consensual categories identified when taking into account all participants. Like in T1 again, both geoscientists
503 and acousticians provide evaluations of the number of *impacts* and the spectral content, but the other aspects
504 (except PERCUSSIVE) are evaluated by acousticians only.

505 E T3b (variable: event/station relative location, speed factor: 350)

506 Just like in T1 and T3a, the consensual categories group together stimuli from stations located close to one
507 another (Fig. 6). Tab. 9 shows that the consensus about the number of impacts is also weaker than in T1.

508 Just like in T1 and T3a, the participants have focused first on the event/station distance, related to the
509 “number of impacts” criterion; and then on the event/station azimuth in order to group the stimuli, related
510 to the other criteria, as discussed below. Sounds in 3bC1 (smallest event/station distance) are perceived as
511 having one impact or two very close one to another, stimuli in 3bC2 and 3bC5 (intermediate event/station
512 distance) as having 2 impacts quite close one to another, stimuli in 3bC3 and 3bC4 (stations far from the event)
513 are perceived as having two impacts. Category 3bC2 shows a weaker consensus about the perceived temporal
514 distance between the impacts. Two sub-categories can be indeed identified in 3bC2: stations *V37A* and *V38A*
515 lie at a greater distance from the epicenter than stations *U36A*, *V36A* and *TUL1*.

516 Here, again, categories differ by the perceived spectral content of their stimuli. Quite like T1 and T3a, stimuli
517 in 3bC4 (far from the event, South-East from it) have more *bass*, stimuli in 3bC2 and 3bC3 (intermediate to
518 long distance from the event, North-East from it) are perceived to have more *treble* and *medium* frequencies,
519 stimuli from 3bC1 and 3bC5 (short to intermediate distance from the event) have more *treble* frequencies.

520 The categories also differ according to the classes VOLUME, REVERBERATION, CRACKLING NOISE, BASS
521 BACKGROUND: these classes require further investigations.

522 Like in T1 and T3a, the consensual categories of acousticians only and geoscientists only are similar to the
523 consensual categories when grouping all the participants. But contrarily to the previous tests, the geoscientists
524 mention all criteria: Increasing even more the playback speed may have made some aspects of sound more
525 salient and noticeable, even for participants not trained to analyze sound¹¹.

526 F Effects of the playback speed (T1, T3a, T3b)

527 Categories are robust with respect to the playback speed. The effect of the playback speed is indeed small in
528 comparison with the event/station relative location, as shown in Tab. 10 by the similarity of categories in T1,
529 T3a and T3b. In particular, subcategories $\{TUL1, U36A, V36A\}$, $\{U38A, U37A\}$, $\{V38A, V37A\}$, $\{V35A,$
530 $W35A\}$ and $\{X37A, X38A, X39A\}$ remain unchanged with speed factor variations. Signals from the stations
531 in each of these subcategories may share some strong similarities that remain to be interpreted in terms of

532 geophysical parameters. Increasing the speed factor with respect to T1 (T3a and T3b) pushes forward the
533 grouping of stimuli { *W36A*, *W37A*, *X35A*, *X36A* }. The speed factor applied in T3a favors the grouping of
534 stimuli { *V35A*, *W36A*, *W37B*, *X35A* }, whereas the speed factor of T3b favors the grouping of stimuli { *W36A*,
535 *W37B*, *X35A*, *X36A* } and { *TUL1*, *U36A*, *V36A*, *V37A*, *V38A* }. A higher speed factor enhances similarities
536 between stimuli, while a lower speed factor enhances the differences between stimuli. Again, these similarities
537 and differences have to be interpreted in terms of geophysical parameters: Different speed factors highlight
538 different aspects of the signals, that are translated in terms of perceived similarities and differences, so possible
539 developments of audition-based seismic data analysis methods may adapt the speed factor depending on the
540 feature of interest in the signals.

541 Additionally, it can be remarked that the consensus about the number of *impacts* is weaker in T3a and
542 T3b than in T1. This is most probably due to the speed factor (playback of the seismic time series) which is
543 increased from T1 to T3b: this necessarily reduces the time interval between the *impacts*, making them harder
544 to resolve/discriminate.

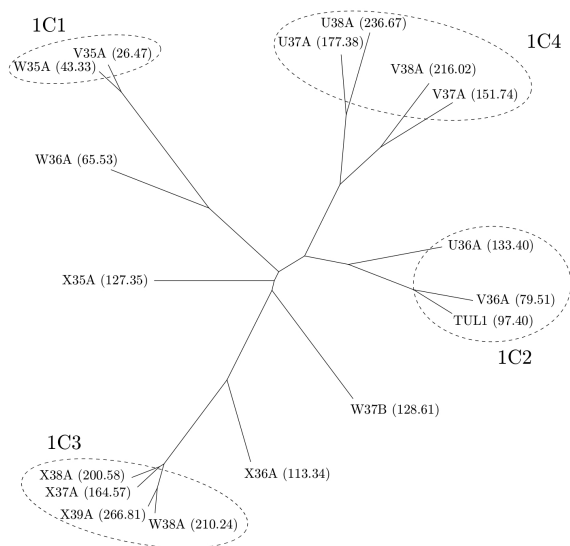


Figure 3: Additive tree for T1 (variable: event/station relative location, speed factor: 150), all participants. The identified consensual categories are numbered from 1C1 to 1C4. Leaves are labelled with the stimulus name, and the event/station distance in km in parentheses.

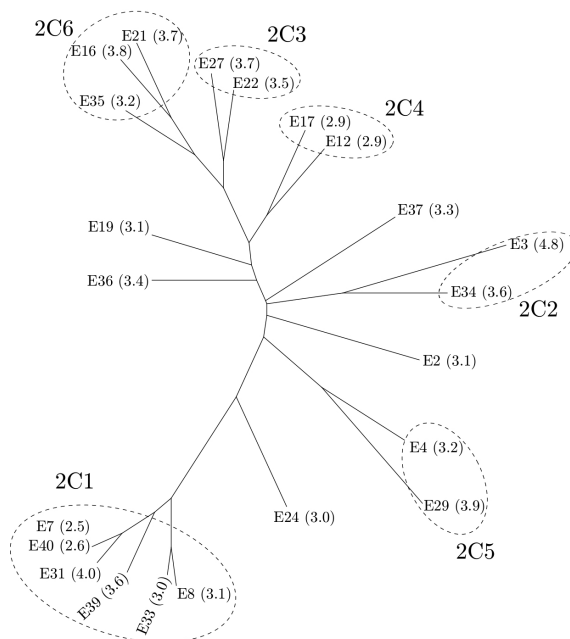


Figure 4: Additive tree for T2 (variable: magnitude, speed factor: 150), all participants. The identified consensual categories are numbered from 2C1 to 2C6. Leaves are labelled with the event number, and the magnitude in parentheses.

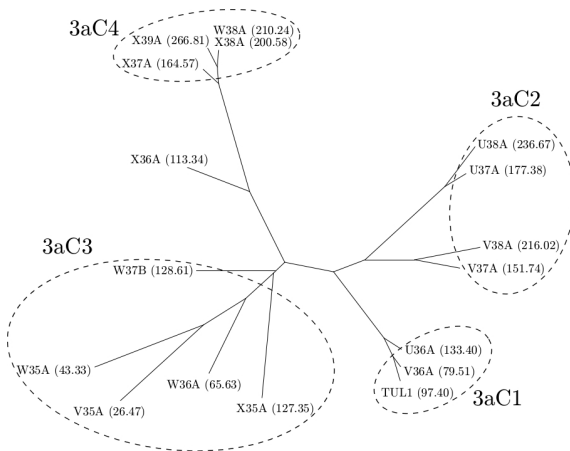


Figure 5: Additive tree for T3a (variable: event/station relative location, speed factor: 250), all participants. The identified consensual categories are numbered from 3aC1 to 3aC4. Leaves are labelled with the stimulus name, and the event/station distance in km in parentheses.

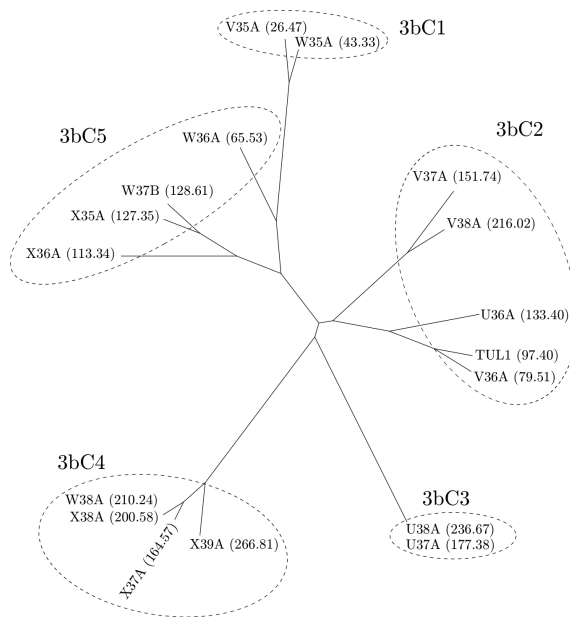


Figure 6: Additive tree for T3b (variable: event/station relative location, speed factor: 350), all participants. The identified consensual categories are numbered from 3bC1 to 3bC5. Leaves are labelled with the stimulus name, and the event/station distance in km in parentheses.

545 VI Acoustic descriptors

546 Some sound aspects have been shown in Sec. V to be particularly relevant for the perceptually- and cognitively-
 547 based categorization of stimuli. These aspects are: the number of IMPACTS (relevant for T1, T3a and T3b), the
 548 FREQUENCIES (relevant for T1, T3a and T3b), and the amount of BASS BACKGROUND (relevant for T2). Building
 549 on these observations, we searched for acoustic descriptors matching these perceptually relevant sound aspects.
 550 These descriptors are presented here. Results about the number of IMPACTS (Sec. A) and the FREQUENCIES
 551 are shown only for stimuli of T1 (similar results are found for stimuli of T3a and T3b, not shown here for
 552 the sake of brevity), and results about the BASS BACKGROUND are shown only for stimuli of T2. Note that
 553 these descriptors have been chosen after the linguistic analysis. These descriptors have also been chosen to be
 554 as simple as possible, hence they are classical and well-known acoustic descriptors. Importantly, the linguistic
 555 analysis has been conducted without any *a priori* on the nature of these acoustic descriptors (or on the fact
 556 that there was going to be any acoustic descriptors to derive).

557 A Number of IMPACTS and the temporal envelope

558 A simple way to visualize the number of *impacts* (note that we keep using the participants' wording) in the
 559 stimuli is the computation of the temporal envelope of the stimuli. The method used here is described in an
 560 article by D'Orazio *et al.*⁶² (computation with 500 iterations). Fig. 7 shows the envelopes of all stimuli of
 561 T1. On the one hand, stimuli in category 1C1, perceived as having one *impact* only, exhibit a rather smooth
 562 envelope with a main *impact* (higher value). On the other hand, stimuli in categories 1C3 and 1C4, perceived
 563 as having 2 *impacts*, have a more irregular envelope with a main *impact* preceded by a lower-amplitude event.
 564 Stimuli of category 1C2, for which the number of perceived *impacts* is not clear, have intermediate envelopes:
 565 The preceding lower-amplitude event does not clearly stand out from the main *impact*. Note that the envelopes
 566 are sketchy and somewhat “quantized” due to the envelope computation algorithm and to our parameters, but
 567 this “simplified” aspect is adequate for qualitative interpretation.

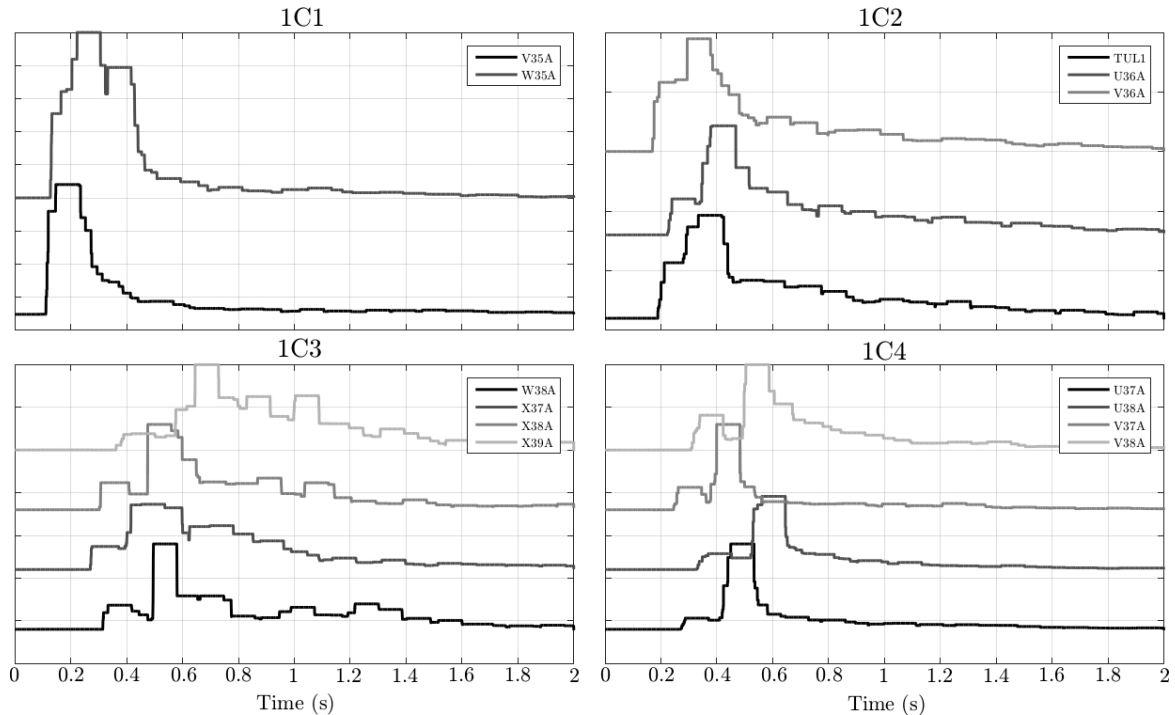


Figure 7: Envelopes of all stimuli of T1. Stimuli are grouped by consensual categories. For each category, the envelopes are shifted vertically for ease of reading.

568 B FREQUENCIES and the spectral centroid

569 Psychoacousticians usually describe the frequency content of sounds with the concept of spectral distribution of
 570 energy⁶³, which is classically illustrated by the spectral centroid. The spectral centroid is the “center of gravity”
 571 of the spectrum and is defined⁶⁴ as:

$$SC = \frac{\sum_{k=1}^N f_k a_k}{\sum_{k=1}^N a_k}, \quad (1)$$

572 where $f(k)$ and $a(k)$ are respectively the frequency and amplitude in bin k . Thus more energy in the low (resp.
 573 high) frequencies gives a lower (resp. higher) spectral centroid. Fig. 8 shows the spectral centroid computed
 574 on each stimulus of T1. Stimuli of category 1C3 have a lower spectral centroid, which is consistent with the
 575 verbalisations describing them as having more *bass*. Stimuli in categories 1C2 and 1C4 have higher spectral
 576 centroids and are judged as having more *treble* and *medium* frequencies. Stimuli in 1C1, perceived as having
 577 more *bass*, have intermediate spectral centroids.

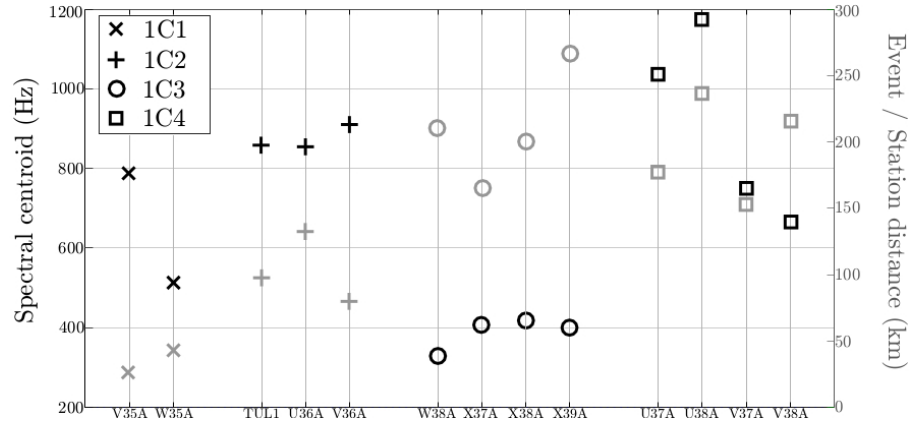


Figure 8: Spectral centroid (black) and event/station distance (gray) for all stimuli of T1. Stimuli are grouped by consensual categories.

578 C BASS BACKGROUND and the signal-to-noise ratio

579 The most natural descriptor of *background* noise is the signal-to-noise ratio SNR . For each stimulus of T2, the
 580 maximum value of the first 500 points of the audio signal is computed. The “noise part” ends and the “signal
 581 part” starts when the audio signal first exceeds 3 times this maximum value. Then the SNR is computed as:

$$SNR = 10 \log_{10} \left(\frac{\int Signal^2(t)}{\int Noise^2(t)} \right) \quad (2)$$

582 Fig. 9 shows the computed SNR value for each stimulus of T2. While stimuli of categories 2C2, 2C3, 2C4 and
 583 2C5 have similar SNR values, stimuli in category 2C1 clearly have lower SNR values. This is consistent with
 584 the perception: Only stimuli of 2C1 are perceived as having a large amount of *bass background*. Furthermore,
 585 Fig. 9 confirms that higher-amplitude events correspond to signals with a higher signal-to-noise ratio.

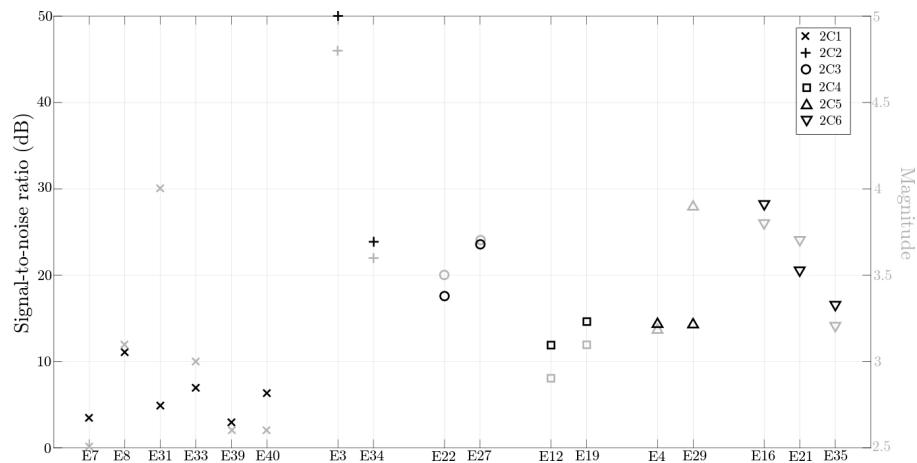


Figure 9: Signal-to-noise ratio (black) and magnitude (gray) for all stimuli/corresponding events of T2. Stimuli are grouped by consensual categories.

586 VII Conclusion

587 Expert human categorization of audified seismic signals is found to match geophysical parameters (event/station
 588 distance and azimuth), confirming previous results²². With respect to earlier work, limited to one of the four
 589 tests presented here, the present study contributes a thorough analysis of the categorization criteria used by
 590 participants. With the help of basic acoustical features, that have been derived and selected after the linguistic
 591 analysis, and that are shown to correspond to the verbal descriptions of stimuli, this makes the link between
 592 geophysical parameters and psychological responses clearer, and will facilitate future applications of auditory
 593 display as a didactic, data analysis and possibly research tool in seismology. Among the categorization criteria,
 594 one can mention the following important ones:

595 **Number of *impacts*** Participants primarily sort stimuli according to event/station distance: The number of
 596 *impacts* and the time difference between the *impacts* is directly related to the difference in arrival time
 597 between P- and S- waves²². The computed temporal envelope of the signal can help visualize, detect and
 598 confirm this temporal distance between seismic phases.

599 **Frequency content** The frequency content of the stimuli can be related to the medium through which seismic
 600 waves travel from the event to the station. The North-East and South-East regions of the investigated area
 601 differ in their elevation, local phase velocity variations, crustal structure and composition of ground, *etc.*,
 602 inducing different scattering, dissipation, attenuation, dispersion behaviors, that may act on the seismic or
 603 sonic waves as filters. As a result audified signals differ in their frequency content depending on whether
 604 they result from recordings made in one of these two regions: Signals from North-East are perceived to
 605 have more *treble* and *medium* frequencies, whereas stimuli from South-East are perceived as having more
 606 *low* frequencies; for similar event/station distances. It has been shown that the spectral centroid is a good
 607 indicator of the perceived frequency content: Stimuli perceived as having more *bass* (resp. *treble*) have
 608 lower (resp. higher) spectral centroids.

609 **Background noise** T2 showed that evaluating the loudness/volume balance between the *background noise*
 610 and the main *impact* can give clues about the magnitude of the seismic event. As mentioned in Sec. B,
 611 signals were normalized according to maximal amplitude, so that the same level of background noise
 612 ends up sounding louder for smaller-magnitude events than for higher-magnitude ones. The computed
 613 signal-to-noise ratio (SNR) has been shown to be a good indicator of the perceived *background noise*
 614 level. The issue of amplitude of the sonified vs. seismic signal is further complicated by the frequency-
 615 dependence of loudness as perceived by the human ear. Frequencies that carry important seismological
 616 information might systematically be underestimated or neglected by the auditory system. In future work,
 617 a frequency-dependent amplitude correction (equalization) of sonified data that accounts for this effect is
 618 envisaged.

619 Other sound aspects have been pointed out by participants, and may prove relevant for the interpretation in

620 terms of geophysical parameters, but they remain to be investigated further: In T1 stimuli from North-East
 621 (resp. South-East) stations are perceived as *fast* and *near* (resp. *slow* and *far*); acousticians notice a *bouncing*
 622 *sound* after the *impacts*, which probably derives from some properties of the coda of the seismic recordings;
 623 *crackling noises* are heard in T3a and T3b, maybe related to some geophysical features. Note that in order to
 624 reduce the dynamic range of seismic signals, all stimuli had to be normalized (the solution here was to normalize
 625 each signal separately according to its maximum amplitude), and as a result loudness differences may have been
 626 much reduced, preventing the subjects to use loudness as an informative and discriminative criterion.

627 As stated in the introduction, this study is aimed at grasping new ideas for making new hypotheses in joining
 628 researchers' expertise from different domains: acoustics and geophysics but also psycholinguistics. Future more
 629 systematic tests will assess and generalize the listeners' sensitivity to the identified acoustic parameters and
 630 their supposed relationship to the geophysical parameters.

631 A logical continuation of this test is to investigate more systematically how geophysical parameters are trans-
 632 lated into acoustical parameters and then into perceptual evaluations, using controlled-source experiments, *i.e.*
 633 using signals from laboratory experiments where the geophysical parameters (propagation medium, direction of
 634 failure, fault geometry) are known and controlled.

635 The categories are quite robust with respect to the playback speeds investigated here, but further tests
 636 should focus on performing a more complete parametric test of the playback speed, in order to identify whether
 637 certain playback speeds facilitate the resolution of impacts, the evaluation of spectral features, or of any other
 638 relevant features.

639 The investigation of the effect of training is also a good candidate for future research: participants may be
 640 trained beforehand to recognize some features, in order to get enhanced results in listening tasks. This may
 641 be applied to discrimination (*e.g.* threshold regarding the number of *impacts*) or recognition tasks (*e.g.* tell
 642 *impacts* from *echoes* resulting from reverberations on different Earth layers).

643 In future work, we also plan to address the issue of normalization. Normalization is needed if earthquakes of
 644 different magnitudes are investigated, due to the greater dynamics in seismic signals than in audio signals. It has
 645 been shown here that normalizing each signal according its maximum amplitude can help for the identification
 646 of some parameters (magnitude, by changing the signal-to-noise ratio), but on the other hand it can remove
 647 auditory clues for the identification of other parameters (distance, by reducing the loudness differences). Other
 648 normalization methods have to be tested (*e.g.* based on RMS value of the signal, or mapping a given seismic
 649 amplitude range to a given loudness range), depending on which seismic parameters are investigated.

650 The study presented here, developing a procedure aiming at employing user perception as heuristics in
 651 acoustics and geophysics, is a step towards the use of calibrated auditory display devices as complementary to
 652 or independent from visual devices. A further idea for future research is to compare human perception in two
 653 contexts, visual display (*e.g.* plot of the seismic wave) and auditory display, *e.g.* audification of the same signal:
 654 How can these two approaches to data display complement each other for the identification of some geophysical
 655 parameters?

656 We expect that the findings presented here will open the way to numerous applications and further develop-
 657 ments. The verbal description of stimuli may bring new ideas for the automated analysis of seismic data. This
 658 study (and previous ones) focused on the earthquakes themselves. Further studies may focus on the “background
 659 noise”, *i.e.* the soft seismic activity occurring before and after majors earthquakes. Higher speed factors (up to
 660 1,000 or even higher) may be considered for the audification of this background noise, making auditory analysis
 661 much faster than visual analysis: one day of data can be monitored in a couple of minutes only. It is possible
 662 that auditory display techniques help understanding aftershocks, seismic swarms, or even possible earthquake
 663 precursors.

664 Acknowledgements

665 We gratefully acknowledge financial support from INSU-CNRS which made our work possible. We would like
 666 to thank François Baudin, Nicolas Bellahsen, Etienne Caylou, Delphine Chadefaux, Elia d’Acremont, Hugo
 667 Dujourdy, Augustin Ernoult, Claudia Fritz, Clément Gariel, Hugues Genevois, Ambroise Harivel, Philippe
 668 Huchon, Elsa Jauffret, Emanuel Kaestle, Félicie Korostelev, Alexis Plunder, Laurent Quartier, Alain Rabaute,
 669 Alexandre Roy, Alexandre Schubnel, Pauline Thierry, Camille Vauthrin, and Jean-Louis Waffart for taking part
 670 in the psychoacoustic tests. Thanks to Nolan Lem and Pascal Gaillard for fruitful discussion. Thanks also to
 671 Nolan Lem, Meredith Nettles, Hannah Rabinowitz, and Heather Savage, for discussions and participation in a
 672 preliminary test at the Lamont-Doherty Earth Observatory, and to Katie Keranen, Patty Lin and Nolan Lem for
 673 data and other assistance. Matthew Vaughan is acknowledged for his work in setting up an early predecessor to
 674 this study at LDEO. Lapo Boschi is grateful to Florian Dombos and Olivier Warusfel for some very interesting
 675 discussions, that inspired part of this study.

676 Notes

678 ¹Throughout the manuscript, the following words — usually synonyms — will be used depending on what is being described:
 679 “ensemble” will denote a group of subjects of similar expertise (the ensemble of acousticians and the ensemble of geoscientists);
 680 “group” will denote a group of stimuli as produced by a subject; “category” will denote a group of stimuli, when added up over the
 681 subjects; and “class” will denote a group of words of similar meaning, that is a semantic class.

682 ²Signal processing refers to an object-centered conception of perception as “data driven”, which is mainly a bottom-up approach
 683 of perception as information processing extracted from the stimulus; whereas signal interpretation refers to a subject-centered
 684 conception in which it is the subject who is giving meaning to the stimulation along knowledge-based, top-down processing, “down
 685 to” the selection of the relevant characteristic to be perceived and to which it is concerned.

686 ³http://www.lam.jussieu.fr/Membres/Pate/Fichiers/Sounds_ArticleEQ/

687 ⁴This is an English translation of the original French instructions *Nous vous demandons de procéder à un tri des extraits sonores*
 688 *qui vous sont présentés. Pourriez-vous regrouper les extraits qui se ressemblent et placer dans des groupes différents ceux qui vous*
 689 *semblent différents ? Vous faites autant de groupes que vous le souhaitez.*

690 ⁵See supplementary material at ...

691 ⁶This analysis has been carried out by the first author.

692 ⁷In the following, words are italicized when they originate from the verbal comments of the participants, or capitalized when
693 they are labels of semantic classes.

694 ⁸In this article the words used by the participants are written in italic font; the labels of the semantic classes are written in
695 uppercase

696 ⁹Because this article aims at interpreting the participant’s wording in terms of physics, the word *frequencies*, borrowed from the
697 lexicon of acoustics, is preferred to the other words of the category. “*Frequencies*”, as an acoustical concept, stops being a word
698 and becomes a “term” (word with a semantic constructed by the speakers sharing a specific expertise)⁶¹.

699 ¹⁰For the same reason, the technical term “*Reverberation*” used in acoustics is chosen here to denote the semantic category.

700 ¹¹Geoscientists may have learnt and got trained in processing acoustics signals after the four experimental sessions (the type of
701 stimulus becoming somehow more familiar and therefore being processed in a more discriminative way). This hypothesis, to be
702 further tested, could be of interest for training geoscientists in acoustics signal processing in order to complement their previous
703 training in reading graphical representations of seismographs

704 References

705 [1] B. Latour, “Visualisation and Cognition: Thinking with Eyes and Hands,” in *Knowledge and Society*
706 *Studies in the Sociology of Culture Past and Present*, H. Kuklick (editor), Jai Press vol. 6, pp. 1–40 (1986).
707 Reprinting and revision in *Representation in Scientific Activity*, Michael Lynch and Steve Woolgar (editors),
708 MIT Press, Cambridge Mass, pp. 19–68 (1990).

709 [2] B. Latour, “Visualisation and Cognition: Drawing Things Together,” in *Knowledge and Society Studies in*
710 *the Sociology of Culture Past and Present* 6:1–40, H. Kuklick (Ed.), JAI Press, (2003).

711 [3] E. von Rebeur-Paschwitz, “The Earthquake of Tokio, April 18, 1889,” *Nature* **40**:294–295 (1889).

712 [4] S. D. Speeth, “Seismometer sounds,” *Journal of the acoustical society of America* **33**(7):909–916 (1961).

713 [5] G. E. Frantti and L. A. Levereault, “Auditory discrimination of seismic signals from earthquakes and
714 explosions,” *Bulletin of the seismological society of America*, **55**(1):1–25 (1965).

715 [6] M.S. Karney “Natural radio (news, comments and letters about natural radio),” December 2006, available
716 online at http://naturalradiolab.com/wp-content/uploads/2015/11/NR2006_12.pdf, retrieved on Decem-
717 ber 22nd 2016.

718 [7] F. Dombois, “Using audification in planetary seismology,” in *Proceedings of the international conference on*
719 *auditory display* (2001).

720 [8] F. Dombois, “Auditory seismology – on free oscillations, focal mechanisms, explosions and synthetic seis-
721 mograms,” in *Proceedings of the international conference on auditory display* (2002).

722 [9] B. Holtzman, J. Candler, M. Turk, and D. Peter, “Seismic sound lab: Sights, sounds and perception of the
723 Earth as an acoustic space,” in *Sound, Music, and Motion*, pp.161–174, Springer International Publishing
724 (2014).

- 725 [10] D. Kilb, Z. Peng, D. Simpson, A. Michael, M. Fisher, and D. Rohrlick, "Listen, Watch, Learn: SeisSound
726 Video Products," *Seismological Research Letters: Electronic Seismologist* **83**(2):281–286 (2012).
- 727 [11] Z. Peng, C. Aiken, D. Kilb, David. R. Shelly, and B. Enescu, "Listening to the 2011 magnitude 9.0 Tohoku-
728 Oki, Japan, earthquake," *Seismological Research Letters*, **83**(2):287–293 (2012).
- 729 [12] C. Hayward, "Listening to the earth sing," in Gregory Kramer, editor, *Auditory Display – Sonification,*
730 *audification, and auditory interfaces*, chap. 15, pp. 369–404. Addison–Wesley, Reading, MS (1994).
- 731 [13] M. Meier and A. Saranti, "Sonic exploration with earthquake data," in *Proceedings of the international*
732 *conference on auditory display* (2008).
- 733 [14] A. J. Michael, "Earthquake sounds," in H. Gupta, editor, *Encyclopedia of Solid Earth Geophysics*, Springer
734 (2011).
- 735 [15] N. Barrett and K. Mair, "Sonification for geoscience: Listening to faults from the inside," in *EGU General*
736 *Assembly Conference Abstracts 16*, p. 4489 (2014).
- 737 [16] N. Barrett and K. Mair, "Aftershock: A science–art collaboration through sonification", *Organised Sound*
738 **19**(1):4–16 (2014).
- 739 [17] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonification handbook*, Logos Verlag, Berlin (2011).
- 740 [18] E. Landi, R. L. Alexander, J. R. Gruesbeck, J. A. Gilbert, S. T. Lepri, W. B. Manchester, and T. H. Zur-
741 buchen, "Carbon ionization stages as a diagnostic of the solar wind," *The Astrophysical Journal* **744**(2):100
742 (2012).
- 743 [19] D. Vicinanza, R. Stables, G. Clemens, and M. Baker, "Assisted differentiated stem cell classification in
744 infrared spectroscopy using auditory feedback," in *Proceedings of the international conference on auditory*
745 *display* (2014).
- 746 [20] K. V. Nesbitt and S. Barrass "Finding trading patterns in stock market data," *IEEE transactions on*
747 *Computer graphics and applications* 24 (2004).
- 748 [21] M. Maeder and R. Zweifel "Downy oak: rendering ecophysiological processes in plants audible," in *Proceed-*
749 *ings of Sound and Music Computing* (2013).
- 750 [22] A. Pat e, L. Boschi, B. Holtzman, and J.-L. Le Carrou, "Categorization of seismic sources by auditory
751 display: a blind test," *International journal of human-computer studies* **85**:57–67 (2016).
- 752 [23] E. Bigand and B. Poulin-Charronnat, "Are we "experienced listeners"? A review of the musical capacities
753 that do not depend on formal musical training," *Cognition* **100**:100–130 (2006).

- 754 [24] S. McAdams, S. Winsberg, S. Donadieu, G. de Soete, and J. Krimphoff, “Perceptual scaling of synthesized
755 musical timbres: Common dimensions, specificities, and latent subject classes,” *Psychological Research*
756 **58**:177–192 (1995).
- 757 [25] G. Lemaitre, O. Houix, N. Misdariis, and P. Susini, “Listener expertise and sound identification influence
758 the categorization of environmental sounds,” *Journal of Experimental Psychology: applied* **16**(1):16–32
759 (2010).
- 760 [26] J. W. Tanaka and M. Taylor, “Object categories and expertise: Is the basic level in the eye of the beholder?,”
761 *Cognitive Psychology* **23**:457–482 (1991).
- 762 [27] D. Dubois, “Categories as acts of meaning: the case of categories in olfaction and audition,” *Cognitive*
763 *science Quaterly* **1**:35–68 (2000).
- 764 [28] D. Dubois, “From psychophysics to semiophysics: Categories as Acts of Meaning : a case study from
765 olfaction and audition, back to colors,” in *Speaking of colors and odors. An interdisciplinary approach*
766 *to cognitive and linguistic categorization of color vision and olfaction*, M. Plumacher and P. Holz (Eds.),
767 Benjamins, Amsterdam, the Netherlands, pp. 45–119 (2006).
- 768 [29] J. Bruner, *Acts of meanings*, Harvard University press, Cambridge, MA (1990).
- 769 [30] D. Dubois, M. Coler, and H. Wörtche, “Knowledge, sensory experience and sensor technology,” *WSPC*
770 *Proceedings of the CHESS II interaction Conference*, Saskatoon, Canada (2013).
- 771 [31] L. W. Barsalou, “Ad hoc categories,” *Memory and Cognition* **11**:211–27 (1983).
- 772 [32] L. W. Barsalou, “Grounded cognition,” *Annual Review of Psychology* **59**:617–645.
- 773 [33] V. Maffiolo, D. Dubois, S. David, M. Castellengo, and J.-D. Polack, “Loudness and pleasantness in struc-
774 turation of urban soundscapes,” *Proceedings of InterNoise*, Christchurch, New Zealand (1998).
- 775 [34] D. Dubois, C. Guastavino, and M. Raimbault, “A cognitive approach to soundscape: Using verbal data to
776 access everyday life auditory categories,” *Acta Acustica united with Acustica*, **92**(6):865-874 (2006).
- 777 [35] G. Kramer, *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Perseus Publishing
778 (1993).
- 779 [36] R. Kerr, “Geophysical exploration linking deep earth and backyard geology,” *Science* **340**:1283–1285 (2013).
- 780 [37] K. Keranen, M. Weingarten, G. A. Abers, B. Bekins, and S. Ge, “Sharp increase in central oklahoma
781 seismicity since 2008 induced by massive wastewater injection,” *Science* **345**:448–451 (2014).
- 782 [38] N. J. van der Elst, H. M. Savage, K. M. Keranen, and G. A. Abers, “Enhanced remote earthquake triggering
783 at fluid-injection sites in the midwestern united states,” *Science*, **341**(6142):164–167 (2013).

- 784 [39] G. Ekström, “Love and Rayleigh phase-velocity maps, 5-40 s, of the western and central USA from USArray
785 data,” *Earth and Planetary Science Letters* (2013).
- 786 [40] D. Dubois, *Le Sentir et le Dire : concepts et méthodes en psychologie et linguistique cognitives (Feel and
787 Say: concepts and methods in cognitive psychology and linguistics)*, L’Harmattan, Paris, France (2009).
- 788 [41] P. Gaillard, “Etude de la perception des transitoires d’attaques des sons de steeldrums : particularités
789 acoustiques, transformation par synthèse et catégorisation (Study of the perception of attack transients of
790 steeldrums’ sounds: acoustic characteristics, synthesis and categorization),” Ph.D. Thesis, Université de
791 Toulouse II - Le Mirail (2000).
- 792 [42] E. Parizet and V. Koehl, “Application of free sorting tasks to sound quality experiments,” *Applied Acoustics*,
793 **73**:61–65 (2011).
- 794 [43] C. Guastavino, “Categorisation of environmental sounds,” *Canadian journal of experimental psychology*,
795 **61**(1):54–63 (2007).
- 796 [44] J. Morel, C. Marquis-Favre, D. Dubois, and M. Pierrette, “Road traffic in urban areas: a perceptual and
797 cognitive typology of pass-by noise,” *Acta acustica united with acustica*, **98**:166–178 (2012).
- 798 [45] A. Paté, J.-L. Le Carrou, B. Navarret, D. Dubois, and B. Fabre, “Influence of the electric guitar’s fingerboard
799 wood on guitarist’s perception,” *Acta Acustica united with Acustica*, **101**(2):347–359 (2014).
- 800 [46] E. Rosch and B. B. Lloyd, *Cognition and categorization*, Lawrence Erlbaum Associates, Hillsdale, NJ
801 (1978).
- 802 [47] D. Dubois, *Sémantique et cognition (Semantics and cognition)*, Éditions du CNRS, Paris, France (1991).
- 803 [48] J. Hampton and D. Dubois, “Psychological Models of Concepts,” in *Categories and Concepts, Theoretical
804 views and inductive data analysis*, Van Mechelen *et al.* (Eds), Academic Press, London, UK, pp. 11–33
805 (1993).
- 806 [49] P. Gaillard, Website for “TCL-LabX”: <http://petra.univ-tlse2.fr/tcl-labx/>, retrieved on December 22nd
807 2016.
- 808 [50] J.-P. Barthélémy and A. Guénoche, *Trees and proximity representations*, J. Wiley, Chichester (New York)
809 (1991).
- 810 [51] D. Dubois, D. Fleury, “From classification to cognitive categorization : the road lexicon”, in *New approaches
811 in classification and data analysis*, E. Diday *et al.* (Eds), Springer Verlag, Berlin, Germany, pp. 25–35 (1994).
- 812 [52] A. Guénoche, “Consensus of partitions: a constructive approach”, *Advances in Data Analysis and Classifi-
813 cation* **5**(3):215–229 (2010).

- 814 [53] C. Guastavino, *Etude sémantique et acoustique de la perception des basses fréquences dans l'environnement*
815 *urbain (Semantic and acoustical study of the perception of low frequencies in the urban environment)*, Ph.D.
816 Thesis, Université Pierre et Marie Curie (2003).
- 817 [54] F. Guyot, M. Castellengo, and B. Fabre, "Étude de la catégorisation d'un corpus de bruits domestiques
818 (Study of the categorization of a household noises corpus)," in *Catégorisation et cognition : de la perception*
819 *au discours (Categorization and cognition: from perception to discourse)*, D. Dubois (ed), Kimé Paris,
820 France, pp. 41–58 (1997).
- 821 [55] F. Guyot, *Etude de la perception sonore en termes de reconnaissance et d'appréhension qualitative : une*
822 *approche par la catégorisation (Study of the sonic perception in terms of recognition and qualitative assess-*
823 *ment: a categorization approach)*, Ph.D. Thesis, Université du Maine (1996).
- 824 [56] J. Poitevineau, Website for "Addtree": <http://petra.univ-tlse2.fr/tcl-labx/>, retrieved on December 22nd
825 2016.
- 826 [57] A. Paté, *Lutherie de la guitare électrique solid body: aspects mécanique et perceptif (Lutherie of the solid*
827 *body electric guitar: mechanical and perceptual aspects)*, Ph.D. Thesis, Université Pierre et Marie Curie
828 (2014).
- 829 [58] J. Morel, C. Marquis-Favre, D. Dubois, and M. Pierrette, "A perceptive and cognitive typology of vehicle
830 pass-by noises," in *Proceedings of Internoise 2010*, Lisbon, Portugal (2010).
- 831 [59] M. Niessen, T. Cruys, C. Cance, and D. Dubois, "Sound and Noise in Sonic Environmental Studies: Com-
832 paring Word Meaning in Discourses of Community Noise and Soundscape Research," *Acta Acustica united*
833 *with Acustica* **99**(6):853-862, (2013).
- 834 [60] C. Cance and D. Dubois, "Dire notre expérience du sonore: nomination et dénomination (To say our sonic
835 experience: designation and denomination)," *Langue Française* **188**:15–32 (2015).
- 836 [61] D. Dubois and A. Giboreau, "Descriptors: attributes? labels? terms? names? a contribution of psycholin-
837 guistics to sensory evaluation," in H. MacFie and C. Kuesten *Developing, comparing and using consumer*
838 *and sensory vocabularies*, *Food Quality and Preference* **17**:669–672 (2006).
- 839 [62] D. D'Orazio, S. De Cesaris, and M. Garai, "A comparison of methods to compute the effective duration
840 of the autocorrelation function and an alternative proposal," *Journal of the acoustical society of America*
841 **130**(4):1954–1961 (2011).
- 842 [63] J. M. Grey, "Multidimensional scaling of musical timbres," *Journal of the Acoustical Society of America*
843 **61**(5):1270–1277 (1977).

844 [64] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, “The Timbre Toolbox: Extracting
 845 audio descriptors from musical signals,” Journal of the Acoustical Society of America **130**(5):2902–2016
 846 (2011).

Variable / test	T1	T2	T3a	T3b
Station location	variable	fixed	variable	variable
Magnitude	fixed	variable	fixed	fixed
Speed factor	fixed (150)	fixed (150)	fixed (250)	fixed (350)
Expertise of listeners	variable	variable	variable	variable

Table 1: Summary of the variable and fixed effects over the 4 tests.

	acousticians		geoscientists	
IMPACTS	<i>2, 1, only one, impact, strokes, waves, shots, temporal distance, close, near, separated by an intermediate duration, distinct</i>	<u>2, 1, 1 seul, impact, coups, ondes, détonations, distance temporelle, proches, rapprochés, séparés par une durée intermédiaire, distincts</u>	<i>2, 1, impact, near, separated</i>	<u>2, 1, impact, rapprochés, séparés</u>
FREQUENCIES	<i>high frequencies, treble, clear, high medium, treble medium, medium, low medium, low frequencies, low, dull, sub bass</i> ----- <i>balanced</i>	<u>hautes fréquences, aigus, clair, haut médium, médiums aigus, médiums, bas médium, basses fréquences, graves, sourds, sub basses</u> ----- <u>équilibré</u>	<i>high frequencies, treble, clear, medium, low frequencies, dull</i>	<u>hautes fréquences, aigu, clair, médium, basses fréquences, sourd</u>
DURATION	<u>short</u>	<u>court</u>	<u>short</u>	<u>court</u>
/ SPEED	<i>accelerated, quick, slowness, slow</i>	<u>accélééré, rapide, lenteur, lent</u>		
DISTANCE	<i>proximity, close, far, remoteness, remote, away</i>	<u>proximité, proche, lointain, éloignement, loin, distants</u>		
REVERBERATION	<i>delay, reverberation, echo</i>	<u>delay, réverbération, écho</u>	<i>echo, resonant</i>	<u>écho, résonant</u>
PERCUSSIVE	<i>percussive, abrupt, sharp, dry</i>	<u>percussif, cassant, pointu, sec</u>	<i>net, cutting</i>	<u>net, tranchant</u>
BOUNCING SOUND	<i>bouncing sound</i>	<u>rebonds</u>		
BASS	<i>rumble, low-frequency noise, sub</i>	<u>grondement, bruit basses</u>		
BACKGROUND	<i>bass</i>	<u>fréquences, infra-basse</u>		
DEEP	<i>deep</i>	<u>profond</u>		
AGGRESSIVE	<i>aggressive</i>	<u>agressif</u>	<i>aggressive</i>	<u>agressif</u>

Table 2: Words used during T1 (variable: event/station relative location, speed factor: 150), sorted by semantic classes and listener ensemble. In normal font the French original words, in italic font the English translation. Words used by the two ensembles of subjects are underlined.

	acousticians		geoscientists	
IMPACTS	<i>2, double, 1, impact, stroke, separated, near</i>	<i>2, double, 1, impact, coup, séparés, rapprochés</i>	<i>2 or 3, 2, double, 1st, impact, separated</i>	<i>2 voire 3, 2, double, 1er, impact, séparés</i>
FREQUENCIES	<i>high frequencies, high medium, medium, low frequencies, low, bass</i>	<i>hautes fréquences, haut-médium, médiums, basses fréquences, graves, basses</i>	<i>high frequencies, medium, low frequencies, dull</i>	<i>hautes fréquences, médiums, basses fréquences, sourd</i>
DISTANCE	<i>far</i>	<i>lointain</i>		
REVERBERATION	<i>echo, delay</i>	<i>écho, delay</i>		
PERCUSSIVE	<i>clear</i>	<i>clair</i>		
BOUNCING SOUND	<i>bouncing sound, sweep, swept sine</i>	<i>rebond, sweep, sinus glissant</i>		
BASS BACKGROUND	<i>sub bass background, continuous low-frequency sound, constant bass, humming noise, wobbly noise, <100Hz, low-frequency throbbing, bass background, extreme low-frequency, low-frequency background, rumble, rolling</i>	<i>fond sonore sub, son basse fréquence continu, constante grave, bourdonnement, basses tremlantes, <100Hz, vrombissement basses fréquences, bruit de fond grave, extrême grave, bruit de fond basses fréquences, grondement, roulements</i>	<i>low-frequency background, vibrations, rumble</i>	<i>bruit de fond basses fréquences, vibrations, grondement</i>
AGGRESSIVE	<i>aggressive</i>	<i>agressif</i>		

Table 3: Words used during T2 (variable: magnitude, speed factor: 150), sorted by semantic classes and listener ensemble. In normal font the French original words, in italic font the English translation. Words used by the two ensembles of subjects are underlined.

	acousticians		geoscientists	
IMPACTS	<i>2, 1, only 1, impact, attack, distinct, spaced, near, close</i>	<i>2, 1, 1 seul, impact, attaque, distinct, espacés, rapprochés, proches</i>	<i>2, 1, only 1, impact, sound, near</i>	<i>2, 1, 1 seul, impact, son, rapprochés</i>
FREQUENCIES	<i>high frequencies, low frequencies, pastel, soft</i>	<i>hautes fréquences, basses fréquences, pastel, doux</i>	<i>high frequencies, medium, low frequencies</i>	<i>hautes fréquences, médium, basses fréquences</i>
DURATION / SPEED	<i>short</i>	<i>courts</i>		
DISTANCE	<i>remote</i>	<i>lointains</i>		
REVERBERATION	<i>reverberating, resonance</i>	<i>réverbérant, résonance</i>		
PERCUSSIVE	<i>twangy, percussive, dry, abrupt, sharp</i>	<i>claquant, percussif, secs, cassant, pointu</i>	<i>net</i>	<i>net</i>
BOUNCING SOUND	<i>sweep, bouncing noise</i>	<i>sweep, rebond</i>		
BASS BACKGROUND	<i>bass / low / deep/ dull background</i>	<i>bruit de fond basses fréquences / grave/ profond / sourd</i>		
DEEP	<i>deep</i>	<i>profond</i>		
AGGRESSIVE	<i>aggressive, hissing</i>	<i>agressif, stridents</i>		
CRACKLING NOISE	<i>crackling noise</i>	<i>grésillement</i>		

Table 4: Words used during T3a (variable: event/station relative location, speed factor: 250), sorted by semantic classes and listener ensemble. In normal font the French original words, in italic font the English translation. Words used by the two ensembles of subjects are underlined.

	acousticians		geoscientists	
IMPACTS	<u>2, only 1, impact, near</u>	<u>2, 1 seul, impact, rapprochés</u>	<u>2, double, 1, impact, separated, near, close</u>	<u>2, double, 1, impact, séparés, rapprochés, proches</u>
FREQUENCIES	<i>treble, high medium, low frequencies, dull, low</i> balanced – équilibré	aigu, médium aigu, médium, basses fréquences, <u>sourd</u> , bas	<i>high frequencies, low frequencies, dull</i>	hautes fréquences, basses fréquences, <u>sourd</u>
DURATION / SPEED			<i>short</i> <i>quick</i>	court rapide
REVERBERATION	<i>echo, delay</i>	écho, delay	<i>tail</i>	tail
BOUNCING SOUND	<i>sweep</i>	sweep		
BASS	<i>extremely low frequencies</i>	composantes extrêmement graves	<i>low-frequency background</i>	bruit de fond basses fréquences
BACKGROUND				
CRACKLING	<i>small high-pitched cracklings</i>	petits claquements aigus	<i>cracklings</i>	crépitements
NOISE				
VOLUME			<i>amplitude, low, volume, high</i>	amplitude, faible, volume, fort

Table 5: Words used during T3b (variable: event/station relative location, speed factor: 350), sorted by semantic classes and listener ensemble. In normal font the French original words, in italic font the English translation. Words used by the two ensembles of subjects are underlined.

		acousticians					geoscientists					ALL			
		V35A	TUL1	W38A	U37A	W37B	V35A	V36A	W38A	U37A	W37B	1C1	1C2	1C3	1C4
				X36A	X37A				X38A	X39A		X37A	X38A	X39A	X37A
		W35A	U36A	X38A	V37A	X35A	W35A	W36A	X38A	V37A	X35A	V35A	U36A	X38A	V37A
		W36A	V36A	X39A	V38A	X35A	W35A	W36A	X39A	V38A	X36A	W35A	V36A	X39A	V38A
IMPACTS	2 clearly/well separated			2	5				2	1	1			7	6
	2 impacts		4	5	4	2	1	2	4	5	5	1	7	9	9
	2 close to one another	1	1		1				3			1	3		1
	2 very close to one another	1	4			2						1	4		
	1 or 2 very close to one another		1				2	2				2	3		
	1 impact	9	1				4					10	1		
FREQUENCIES	have different pitch				1										1
	2nd impact louder									1	1				1
	treble		5	-1	2	1	1	3	-1	3	1	1	8	-2	5
	high medium		2		1								2		1
	medium	3	1		3	1		2		2	1	3	3		5
	low medium	1										1			
DURATION / SPEED	bass	-1	2	11	-2				5		2	-1	2	16	-2
	balanced		1								1		1		
	short	1								1		1			1
DISTANCE	fast		2	-2	2	1							2	-2	2
	far			-1	3	-1							-1	3	-1
OTHERS	Reverberation	-1		1		3	1	2	1	2	1			2	2
	Percussive	2	2	1	3	2				2		1	2		5
	Bouncing sound	1	-1			1						1	-1		
	Aggressive		2										2		
	Bass background		2	2									2	2	
	Deep			1										1	

Table 6: T1 (variable: event/station relative location): Summary of the verbal description of the consensual categories of stimuli, grouped by semantic classes. The numbers indicate the cumulative number of evaluation for each verbal descriptor (line). A positive (resp. negative) number indicates the presence (resp. absence) of the sound aspect. The categories resulting from the comments of all subjects are numbered from 1C1 to 1C4.

		acousticians					geoscientists					ALL					
		E7			E16		E7					2C1	2C2	2C3	2C4	2C5	2C6
		E8			E21		E8					E7					
		E31		E12	E21		E31			E2		E8					
		E33		E17	E22	E29	E33			E19		E31					
		E39	E3	E19	E27	E4	E39	E3	E22	E34	E21	E33					E16
		E40	E34	E36	E35	E37	E40	E16	E36	E37	E35	E39	E3	E22	E12	E4	E21
												E40	E34	E27	E17	E29	E35
IMPACTS	2 clearly/well separated	1		1	1		2			1		3		1			1
	2 impacts					2	3	3	2	3	3	3	1	2		1	3
	2 very close to one another		1		1								1	1			1
	1 impact			1									1				
FREQUENCIES	1st impact shorter								1	1	1						1
	1st impact weaker							1									
	treble	1			1	1	2		2	2	1	3		1			2
	high medium					1											
	medium			1						1							
	bass	2		2		-2	1	1				3		-1	-1	-1	1
DISTANCE	Far					1										1	
OTHERS	Bouncing sound	-1	1	5	7	-1						-1	2	6	5		7
	Percussive	2				-1						2				1	
	Reverberation				1	-1								1		-1	1
	Bass background	13	-5	-6	-6	-5	9	-1		-3	-1	22	-7	-4	-5	-7	-4
	Aggressive					1											

Table 7: T2 (variable: magnitude): Summary of the verbal description of the consensual categories of stimuli, grouped by semantic classes. The numbers indicate the cumulative number of evaluation for each verbal descriptor (line). A positive (resp. negative) number indicates the presence (resp. absence) of the sound aspect. The categories resulting from the comments of all subjects are numbered from 2C1 to 2C6.

		acousticians					geoscientists				ALL			
		TUL1	U37A U38A	V35A V37A V38A	W36A W37B X35A X36A	W38B X37A X38A X39A	TUL1	U37A U38A	V35A V37A V38A	W38A X36A X37A X38A X39A	3aC1	3aC2	3aC3	3aC4
										TUL1	U37A	W35A W36A	V35A W38A	
IMPACTS	2 clearly separated		1			2								2
	2 impacts		1	2	2	1		2	5	3	1	3	7	4
	2 close one to another			2	1						1		2	
	2 very close one to another		3		2			1			1	4		
	1 or 2 very close one to another			1			2				2		1	
	1 impact	2		1	1	1	1			1	1	1	1	
FREQUENCIES	treble		4	4	2	1		1	3		1	5	7	1
	medium							1	1			1	1	
	bass		-2	-2		3	1		1	3	1	-2	-1	7
DURATION / SPEED	short				1								1	
DISTANCE	far			1									1	
OTHERS	bass background			-2	3								-2	
	deep			-1		1					1		-1	1
	aggressive			2									2	
	crackling noise	1	1	1	1	-1					1	1	1	-1
	bouncing sound			1	2	1					1		1	1
	Percussive reverberation		2	5					1	1		2	6	1
					2									2

Table 8: T3a (variable: event/station relative location, speed factor: 250): Summary of the verbal description of the consensual categories of stimuli, grouped by semantic classes. The numbers indicate the cumulative number of evaluation for each verbal descriptor (line). A positive (resp. negative) number indicates the presence (resp. absence) of the sound aspect. The categories resulting from the comments of all subjects are numbered from 3aC1 to 3aC4.

		acousticians						geoscientists						ALL				
		V35A	TUL1			W38A X37A	W37B	V35A	TUL1 V36A	U36A			W38A X37A	3bC1	3bC2	3bC3	3bC4	3bC5
		W35A	U36A	V37A	U37A	X38A	X35A	V35A	W37B	V37A	U37A	X38A	V35A	TUL1 U36A V36A	U37A	X38A	W38A X37A	W36A W37B
		W36A	V36A	V38A	U38A	X39A	X36A	W35A	X35A	V38A	U38A	X39A	W35A	V38A	U38A	X39A		
IMPACTS	2 clearly separated									1	1	2		1	1	2		
	2 impacts		1	1	2	1		1	2	3	3	3	1	2	5	4	3	
	2 close to one another								1								1	
	2 very close to one another		1				1	1	1			1	1	1				2
	1 or 2 very close	1					1											1
	1 impact	2	1					2					4					
FREQUENCIES	treble		1		3			2	3	3	2	1	2	2	5	1	3	
	high medium		1	1		-1	-1							1		-1	-1	
	medium	1	3	2	1									1	1			
	bass		1	-1	-1	3						4		1	-1	7	-1	
	balanced		1				1							1				1
DURATION / SPEED	short							1	1	1	1	1	1		1	1	1	1
	fast							1	1	1			1	1				1
OTHERS	bouncing sound	1																
	reverberation	-1	3	3		3	-1	-1	1	1	1	1	-1	2	1	4		
	bass background					1			1	1	-1				-1	1	1	
	crackling noise		2	-1	2	-1	2	1	1	1	1	-1		2	3	-2	2	
	volume							-1	2	1	1	-1	-1	1	1	-1	2	

Table 9: T3b (variable: event/station relative location, speed factor: 350): Summary of the verbal description of the consensual categories of stimuli, grouped by semantic classes. The numbers indicate the cumulative number of evaluation for each verbal descriptor (line). A positive (resp. negative) number indicates the presence (resp. absence) of the sound aspect. The categories resulting from the comments of all subjects are numbered from 3bC1 to 3bC5.

U36A	T1, T3a, T3b																		
U37A																			
U38A			T1, T3a, T3b																
V35A																			
V36A	T1, T3a, T3b	T1, T3a, T3b																	
V37A	T3b	T3b	T1, T3a	T1, T3a			T3b												
V38A	T3b	T3b	T1, T3a	T1, T3a			T3b	T1, T3a, T3b											
W35A					T1, T3a, T3b														
W36A					T3a					T3a									
W37B					T3a					T3a	T3a, T3b								
W38A																			
X35A					T3a					T3a	T3a, T3b	T3a, T3b							
X36A											T3b	T3b		T3b					
X37A													T1, T3a, T3b						
X38A																	T1, T3a, T3b		
X39A																	T1, T3a, T3b	T1, T3a, T3b	
	TUL1	U36A	U37A	U38A	V35A	V36A	V37A	V38A	W35A	W36A	W37B	W38A	X35A	X36A	X37A	X38A	X39A		

Table 10: Comparison of T1, T3a and T3b. For each pair of stimuli we give in the corresponding cell the tests during which they are put in the same consensual category.