



HAL
open science

On the Perception of Audified Seismograms

Lapo Boschi, Laurianne Delcor, Jean-Loic Le Carrou, Claudia Fritz, Arthur Paté, Benjamin Holtzman

► **To cite this version:**

Lapo Boschi, Laurianne Delcor, Jean-Loic Le Carrou, Claudia Fritz, Arthur Paté, et al.. On the Perception of Audified Seismograms. *Seismological Research Letters*, 2017, 88 (5), pp.1279-1289. 10.1785/0220170077 . hal-01609011

HAL Id: hal-01609011

<https://hal.sorbonne-universite.fr/hal-01609011>

Submitted on 3 Oct 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.

1 On the perception of audified seismograms

2 Lapo Boschi^{1,5}, Laurianne Delcor^{2,6}, Jean-Loïc Le Carrou³,
3 Claudia Fritz³, Arthur Paté⁴, and Benjamin Holtzman⁴

4 ¹Sorbonne Universités, UPMC Univ Paris 06, CNRS, Institut
5 des Sciences de la Terre de Paris (iSTeP), 4 place Jussieu
6 75005 Paris, France

7 ²Laboratory of Vibration and Acoustics, 69621 Villeurbanne,
8 France

9 ³Sorbonne Universités, UPMC Univ Paris 06, CNRS, Institut
10 Jean Le Rond d’Alembert, équipe LAM, 4 place Jussieu 75005
11 Paris, France

12 ⁴Lamont-Doherty Earth Observatory, Columbia University,
13 New York, U. S. A.

14 ⁵ALSO AT: Sorbonne Universités, UPMC Univ Paris 06,
15 CNRS, Institut Jean Le Rond d’Alembert, équipe LAM, 4
16 place Jussieu 75005 Paris, France

17 ⁶ALSO AT: Airbus Helicopters, France

May 22, 2017

19 **Abstract**

20 Recordings of the Earth’s oscillations made by seismometers, following earth-
21 quakes or other geophysical phenomena, can be made audible by simply
22 accelerating and playing them through an audio reproduction system. We
23 evaluate quantitatively the possibility of using such acoustic display of seis-
24 mic data for practical applications. We first present to listeners examples
25 of two categories of data, based on geophysical parameters (the geometry of
26 the seismic fault; the terrain–oceanic or continental–sampled by the propa-
27 gating seismic wave) that are not revealed to them. The listeners are then
28 asked to associate each of a set of audified seismograms, that are presented
29 to them binaurally, to either one of the two categories. After this exercise,
30 they are asked to define the features of audified signals that helped them in
31 completing this task. A subset of the listeners undergo a training session,
32 before taking one of the tests for a second time. While the number of listen-
33 ers is too small for a definitive statistical analysis, our results suggest that
34 listeners are able, at least in some cases, to categorize signals according to all
35 the geophysical parameters we had chosen. Importantly, we clearly observe
36 that listeners’ performance can be improved by training. Our work opens
37 the way to a number of potentially fruitful applications of auditory display
38 to seismology.

39 Introduction

40 Auditory display, or “sonification” of scientific data has been applied suc-
41 cessfully to research topics in several disciplines [e.g. *Cowen*, 2015]. Seismic
42 data analysis naturally lends itself to audification: a particularly simple form
43 of sonification which consists of accelerating seismic signals (whose frequency
44 is lower than that of audible sound) before playing them through an audio
45 reproduction system. Auditory display of seismic data was first explored dur-
46 ing the Cold War, when the ability to distinguish underground nuclear tests
47 from natural earthquakes acquired a political relevance [*Speeth*, 1961; *Frantti*
48 *and Leverault*, 1965; *Volmar*, 2013]. Audification was eventually discarded,
49 in this context, in favour of seismic-array methods [*Volmar*, 2013]; in recent
50 years, however, it has been revived by seismologists, mostly for purposes of
51 teaching and dissemination [e.g. *Dombois and Eckel*, 2011; *Kilb et al.*, 2012;
52 *Peng et al.*, 2012; *Holtzman et al.*, 2014; *Tang*, 2014]. Our own experiments
53 [*Paté et al.*, 2016, 2017] have convinced us that it is a valuable and inspi-
54 rational tool for the analysis of seismic data in many contexts. We suggest
55 that it might also soon find more specific, effective research applications.

56 This study attempts to contribute to the quantitative analysis of the
57 human auditory system’s response to audified seismic data. As researchers
58 peruse data via auditory display, the implicit assumption is made that they
59 are capable of recognizing patterns and completing some related tasks by
60 hearing. We question this assumption for the case of audified seismic data,
61 and thus begin to evaluate what can be achieved by audification that is
62 not already implemented through “traditional” techniques in seismic data

63 analysis. Both the early work of *Speeth* [1961] and the recent efforts by our
64 group [*Paté et al.*, 2016] indicate that listeners can detect meaningful clues
65 in audified seismic signals, and thus categorize the signals according to such
66 clues. *Paté et al.* [2016] showed that the categories formed by the listeners
67 can be associated with several geophysical parameters, but could not entirely
68 distinguish the effects of individual parameters (e.g., source-receiver distance,
69 geological properties of the terrain at the receiver and between source and
70 receiver, etc.) from one another. We present here a different approach to
71 the analysis of audified data: listeners are asked to complete a constrained-,
72 rather than free-categorization task, on two sets of data, each controlled by a
73 single geophysical parameter (Earth structure in the area where the recorded
74 seismic waves propagate; focal mechanism of the source). The listeners’
75 performance in auditory analysis is compared with their performance in a
76 similar task, completed via visual analysis of analogous data. We consider
77 the visual analysis of a plot to be a “traditional” task that most individuals
78 with some scientific background are, to some extent, familiar with. Visual
79 analysis serves here as a reference against which results of auditory tests
80 can be compared, and, accordingly, its results are not analyzed in as much
81 detail. Listeners are then briefly trained, and the auditory test repeated
82 after training, with a general improvement of test scores. Finally, listeners
83 are asked to explain the criteria they followed to categorize the data, and
84 their description is compared with quantitative parameters computed from
85 the data.

86 Database

87 The work of *Paté et al.* [2016] evidenced the difficulty of disentangling the
88 influences of different physical parameters on the seismic signal (e.g., source-
89 receiver distance, properties of the source, geology at the receiver location,
90 geology between source and receiver). We compiled two new audified seismic
91 data sets, each designed to emphasize the role of one specific parameter.
92 Both data sets only included events of magnitude between 6 and 8, with
93 focal depths estimated by IRIS between 20 and 40 km, and recorded at
94 epicentral distances between 4000 and 6000 km. The scale lengths under
95 consideration are therefore different from those of *Paté et al.* [2016], who
96 used recordings of a magnitude-5.5 event made no more than a few hundred
97 km from the epicenter. All events contributing to either data set occurred
98 between August 9, 2000, and April 18, 2014.

99 The first data set (DS1) is limited to source mechanisms of the strike-slip
100 type, with magnitude between 6 and 7, and the propagation path (approx-
101 imated by an arc of great circle) is required to lie entirely within either a
102 continental or oceanic region. Fig. 1 shows that events in DS1 are located
103 along the Pacific coast of Mexico and in California, while stations can be in
104 North America (continental paths), on ocean islands throughout the Pacific
105 ocean, in Chile or on the Alaskan coast (oceanic paths). It is well known
106 that a seismic waveform is affected in many ways by the properties of the
107 medium through which the wave propagates before being recorded. Based,
108 e.g., on recent work by *Kennett and Furumura* [2013] and *Kennett et al.*
109 [2014] on waveform differences across the Pacific Ocean, we anticipated that

110 the bulk properties of oceanic vs. continental crust and lithosphere would
111 result in profoundly different seismograms and audified signals. We expected
112 this ocean/continent dichotomy to be far more important than other parame-
113 ters in characterising traces in DS1, and we assumed that it would also guide
114 the subjects' response to the corresponding audified signals.

115 The second data set (DS2) is limited to continental propagation paths,
116 but includes both strike-slip and thrust events of magnitude between 6 and
117 8 (Fig. 2). We expected differences between signals generated by strike-slip
118 and thrust events to be more subtle, and harder to detect, whether visually or
119 aurally. Again, all sources contributing to DS2 are in South-Western North
120 America; stations are distributed throughout Canada and the United States,
121 and, in one case, in the Caribbean. Earthquake mechanisms were obtained
122 from the Global Centroid Moment Tensor Project (see "Data and resources"
123 section).

124 Approximately 500 seismograms meeting the requirements of DS1 and
125 DS2 were downloaded from the IRIS database (see "Data and resources"
126 section) but only traces showing, at a visual analysis, a relatively high signal-
127 to-noise ratio were kept. As a result, DS1 includes 23 "continental" and 23
128 "oceanic" signals, while DS2 includes 52 strike-slip and 52 thrust signals. No
129 filtering or instrument-response correction was applied to the data.

130 The sampling rate of all downloaded seismic traces is 50 Hz. The duration
131 of traces to be audified is 8000 s, starting 1800 s before the P -wave arrival as
132 found in the IRIS catalog, and including the most significant seismic phases
133 and most or all of the coda. Time is sped up by a factor of 1200, selected
134 so that all frequencies present in the seismic traces are mapped into the

135 audible range [*Holtzman et al.*, 2014]. Each sonified signal was normalized
136 with respect to its maximal value. The resulting, “audified,” 6-s-long signals
137 are turned into Waveform Audio File Format (WAV) files via the Matlab
138 function `audiowrite`. Their spectra show most energy between 20 and 600
139 Hz.

140 Experiments

141 All experiments (table 1) were conducted in an acoustically dry room (i.e.,
142 not entirely anechoic, but with very little reverberation of sound). The sub-
143 jects played audified seismic signals on a laptop computer via a Matlab-based
144 software interface, and listened to them through an audio card and closed
145 headphones with adjustable volume. Some tests involved the visual, rather
146 than acoustic display of the signals, which was also implemented with the
147 same interface: seismograms were plotted in the time domain as in Fig. 3 (al-
148 beit with a longer time window, extending from ~ 0 to ~ 20000 s) and subjects
149 had no way to modify the plots’ size or format. We provided each subject
150 with all necessary instructions at the beginning of the test, so that the sub-
151 ject would be able to take the test autonomously. The subjects knew that the
152 signals were originated from seismograms; at the beginning of the test, they
153 were told that all signals would belong to one and only one out of two possible
154 “families,” named A and B. By assigning “neutral” names to data families,
155 and providing no information as to their nature, we minimize the bias that
156 might be caused by a specialized (geophysical) knowledge/understanding of
157 the data. After each test, subjects were asked to briefly explain the crite-

158 ria they had followed in responding to it. They typed their answers on the
159 computer used for the test.

160 All subjects were researchers, faculty, and graduate and undergraduate
161 students with backgrounds in Earth sciences (referred to in the following
162 as “geoscientists”), room or musical acoustics (“acousticians”) or applied
163 physics/engineering (“physicists”).

164 **Constrained categorization without training**

165 In a first suite of experiments, families A and B were each defined by three
166 examples, that subjects listened to or looked at before starting the test. Each
167 of the three example audified signals could be listened to three times at most.
168 Visual examples were plotted on the screen, and could be looked at for no
169 more than three minutes before starting the test. All subjects were given
170 the same examples. The subjects were then exposed to 40 unknown signals;
171 after listening to/looking at each signal, they selected whether it belonged to
172 family A or B; no other answer was possible. Each auditory signals could be
173 listened to three times at most; plots were visible on the screen for 5 seconds.
174 The subjects’ selections were recorded by the software interface.

175 Importantly, this approach is profoundly different from that of *Paté et al.*
176 [2016], who asked subjects to form as many categories as they wanted ac-
177 cording to their own criteria [Gaillard, 2009]. It is also different from “paired
178 comparison,” where a subject is presented with two stimuli, and must choose
179 which one belongs to which of two categories. We have explored the latter
180 approach in preliminary tests with few subjects, who all obtained extremely

181 high scores: this strengthened our hypothesis that the geophysical parame-
182 ters we had selected (propagation path and orientation of the fault) do map
183 into audible acoustic properties of the corresponding audified signals. We
184 considered, however, that a paired-comparison test does not resemble any
185 real task in seismic data analysis, and discarded this approach in our subse-
186 quent experiments.

187 **Auditory and visual display of DS1 (oceanic vs. continental paths)**

188 In a first experimental session, 35 subjects (13 women, 22 men), aged between
189 18 and 61, took two tests involving data from DS1. The group included 18
190 acousticians, 9 geoscientists and 8 physicists. 40 signals were evaluated visu-
191 ally in one test, and their audified counterparts were listened to in another.

192 As explained in the “Database” section above, we made the hypothesis
193 that data belonging to DS1 would tend to be categorized according to the
194 terrain sampled by the propagation paths. Signals corresponding to oceanic
195 propagation paths were presented as examples of family A, and “continental”
196 signals as examples of B. In the following, we loosely speak of “correct”
197 answer whenever a subject associates to family A an “oceanic” signal, or to
198 family B a “continental” one. Exactly half of the signals in this experiment
199 correspond to oceanic propagation paths, the other half to continental ones.
200 The signals were the same for all subjects, but their order was random,
201 changing at each realization of the experiment.

202 The average percentage of correct answers (average “score”) in this first
203 experiment amounts to 78% for the visual test, and 63% for the auditory
204 one. All scores are summarized in the histograms of Fig. 4a and b. We

205 suspect the very low scores of two outliers (one per test) to have been caused
206 by a misunderstanding of the instructions which resulted in the subjects
207 swapping families A and B.

208 For the sake of comparison, we consider the case of entirely random an-
209 swers, i.e., the human subject is replaced by an algorithm that generates
210 random yes/no answers, or answers are given by tossing a coin. In this “null
211 hypothesis,” test scores are controlled by the cumulative binomial distribu-
212 tion [*Press et al.*, 1992, e.g.]: each signal listened to can be treated as an
213 independent “trial”, with a success probability of 50%. Fig. 4a shows that in
214 the first auditory test about one out of three subjects scored above the 99%
215 confidence level as defined through the cumulative binomial distribution: in
216 other words, the probability that a subject would obtain (at least) such score
217 by giving random answers is less than 1%. It is thus probable that some of
218 the best-scoring subjects have identified a real difference between signals that
219 they classified as belonging to families A and B. Given how we constructed
220 the two families (see “Database” section), it is also reasonable to infer that
221 the auditory clues identified by the subjects are directly related to the effects,
222 on seismic waveforms, of wave propagation through oceanic vs. continental
223 crust.

224 Our data are not numerous enough for the histograms in Fig. 4a,b to
225 clearly suggest specific statistical distributions. By visual inspection of Fig. 4a
226 one might speculate that the distribution of auditory test scores is bimodal,
227 with one peak around 50% corresponding to the null hypothesis, and an-
228 other peak around 70% reflecting the performance of subjects who did find
229 meaningful clues in the signals.

230 Scores in the visual test (Fig. 4b) were generally quite high, and higher
231 than for the auditory test. This indicates that, at this point, visual analysis
232 of the data might be a more effective way to complete the task of categorizing
233 DS1 data.

234 **Auditory and visual display of DS2 (thrust vs. strike-slip faults)**

235 Of the subjects who took part in the experiment described in the previous
236 section, 27 (15 acousticians, 7 physicists, 5 geoscientists; 7 women, 20 men)
237 also participated in a second session, involving 40 signals from DS2. Half of
238 the signals were originated from the strike-slip faults, the other half from the
239 thrust faults shown in Fig. 2b. Again, each subject took an auditory and a
240 visual test, with average scores of 52% and 62%, respectively. The results of
241 both auditory and visual tests are illustrated in Fig. 4c,d.

242 Comparison with the null hypothesis shows that the probability of achiev-
243 ing (at least) the average score associated with the visual test by selecting
244 the answers randomly was relatively low ($<10\%$); we infer that at least some
245 subjects are likely to have found visual clues in plotted seismograms. Con-
246 versely, the probability of achieving (at least) the average score obtained in
247 the auditory test by giving random answers was about 40%. Too high for
248 the average observed score to be considered significant. It might be guessed
249 that the one subject who achieved a score of 75% might have found auditory
250 clues in the signals, but overall the test cannot be considered a success.

251 **Constrained categorization with training**

252 17 subjects (10 acousticians, 4 physicists, 3 geoscientists; 4 women, 13 men),
253 who had already participated in both the constrained-categorization experi-
254 ments, accepted to undergo a training session, followed by an auditory test
255 analogous to those described above. The new exercise was conducted on
256 data from DS2, only half of which were employed in our previous experi-
257 ments. Data included in the final test had not been listened to in the course
258 of the training session. The goal of this experiment is to determine whether
259 performance in auditory analysis of seismic data can in principle be improved
260 by training: this is determined below by comparison with performance in a
261 similar task before training. It is therefore not strictly necessary to compare
262 the results against those of visual analysis of the same data, and accordingly
263 the visual test was not repeated.

264 **Training**

265 Subjects were trained [e.g. *Thorndike*, 1931; *Speeth*, 1961] by means of a
266 software interface similar to that used in the actual tests. They first listened
267 to three examples of each family, as before the previous test. They were then
268 presented with up to 24 audified signals in the same way as previously. Half
269 of these signals originated from thrust, the other half from strike-slip faults.
270 Half had been listened to during the previous experiment, half were entirely
271 new. The order in which the signals were presented was random. Upon
272 hearing each signal, subjects were asked by our software interface to evaluate
273 whether it belonged to family A or B. After giving an answer, they were

274 immediately notified whether or not it was “correct” (i.e., consistent with
275 our hypothesis), by the on-screen messages “you have identified the right
276 seismological family” (*“vous avez identifié la bonne famille sismologique”*)
277 and “that is not the right seismological family” (*“ce n’est pas la bonne famille*
278 *sismologique”*), respectively. If a subject had a perfect score after listening
279 to the first 16 sample signals, the training session would end.

280 **Auditory display of DS2 after training**

281 After a brief pause, all subjects who undertook the training session stayed
282 for a final test. 36 signals were randomly picked from DS2. Half of the picked
283 signals had to be from thrust, half from strike-slip faults. Half had to belong
284 to the pool of signals listened to in the test of section “Auditory and visual
285 display of DS2”.

286 The histogram in Fig. 5 shows that scores are generally higher now than
287 when categorizing signals from DS2 before training (Fig. 4c). Only 4 out of
288 17 subjects did not improve their score at all. In the null hypothesis, with
289 36 trials, the probability of achieving a score of at least 69.4% (24 correct
290 answers out of 36) is about 1%: 6 out of 17 subjects scored 70% or more, and
291 we infer that at least some of those 6 learned to recognize relevant auditory
292 clues in the data. Albeit small, these figures appear more significant if one
293 considers that only one brief training session was undertaken.

294 Identifying audio features relevant to catego- 295 rization

296 At the end of a test, the subject was asked to briefly explain the criteria
297 followed to categorize the signals, via the on-screen message: “according to
298 what criteria have you associated family A and B to the signals?” (“*sur*
299 *quel(s) critère(s) avez vous attribué la famille A ou B aux signaux ?*”). The
300 subject could answer by typing some comments through our software inter-
301 face.

302 Given the difficulty of an exhaustive semantic study of the resulting data
303 [Paté *et al.*, 2017], we only give here a preliminary, simplistic analysis of a
304 subset of the recorded comments. Our goal in this endeavour is to identify
305 some of the auditory clues that lead subjects to make their choices. We focus
306 on the subjects whose scores were highest, as the criteria that guided them
307 are probably related to the geophysical parameters that defined our families
308 of signals.

309 Comments on DS1

310 We first analyze the comments made by 5 subjects (2 acousticians, 2 geosci-
311 entists, and one physicist) who all achieved scores $\geq 80\%$ in discriminating
312 audified seismograms corresponding to oceanic vs. continental paths (DS1).

313 Table 2 shows a number of reoccurring suggested clues, namely: the pres-
314 ence of what the subjects identify as “background noise,” and its timbre;
315 the duration of what is considered by the subjects to be meaningful signal;

316 the identification of “echos” in the signal. These features can in principle be
317 associated to quantities calculated by seismic data analysis.

318 First of all, it is relatively easy to identify the onset of an earthquake
319 recording on a seismogram (i.e., the P -wave arrival), and it is then reasonable
320 to define as background noise the signal recorded before such arrival. In all
321 our recordings, the first 500 samples clearly precede the arrival of the main
322 signal and we accordingly identify them as noise. We define the beginning
323 of the seismic signal as the first recorded sample whose amplitude is at least
324 three times larger than the largest amplitude found within the 500 noise
325 samples. Let n_S denote its index. The signal-to-noise ratio (SNR) in decibels
326 can then be estimated, based on the mean amplitudes of signal and noise, by
327 the formula

$$\text{SNR} = 10 \log_{10} \left[\frac{\sum_{i=n_S}^N s^2[i]}{\sum_{i=1}^{500} s^2[i]} \right], \quad (1)$$

328 where $s[i]$ is the amplitude of the i -th sample, in a recording that consists of N
329 samples total. We compute the SNR of all signals in DS1, and find (Fig. 6a)
330 that continental paths tend to be associated with higher SNR values than
331 oceanic paths. This statistical result is in qualitative agreement with the
332 subjects comments.

333 We evaluate the “timbre” of background noise by taking the Fourier trans-
334 form of the first 500 samples only. Fig. 6b shows the distribution of frequency
335 values corresponding to the highest spectral peak in the resulting Fourier
336 spectrum: whether the terrain traversed by the propagating seismic wave is
337 oceanic or continental does not appear to affect significantly the frequency
338 content of noise.

339 We next attempt to quantify the duration of meaningful seismic signal
340 which the subjects believe to have recognized in their listening experiences:
341 after the main, high-amplitude interval that includes body- and surface-wave
342 arrivals, the peak amplitude of all our signals decreases until it becomes as
343 low as the peak amplitude of background noise. For each seismogram, we find
344 the latest sample whose peak amplitude is as large as 10% of its maximum
345 recorded value for that seismogram; we then measure the length of the time
346 interval that separates it from the maximum-amplitude sample, and define
347 it as the duration of seismologically meaningful signal. Fig. 6c shows how
348 such values are distributed for signals associated with oceanic vs. continental
349 propagation paths, and indicates that oceanic signals are, according to our
350 definition, longer than continental ones.

351 Finally, echos can be identified by visual analysis of a seismogram's en-
352 velope. We calculate the envelopes of all our audified seismograms, and take
353 the averages of all oceanic-path and all continental-path envelopes. In anal-
354 ogy with *Paté et al.* [2017], the envelope is defined as suggested by *D'Orazio*
355 *et al.* [2011]: starting with i coinciding with the index of the last sample in
356 a signal, if sample $i - 1$ exceeds sample i , then the value of sample $i - 1$
357 is saved as the i -th entry of the envelope; the procedure is iterated for the
358 preceding sample, until the entire trace is processed [*D'Orazio et al.*, 2011,
359 figure 5]. The results of this exercise, illustrated in Fig. 7, show that (i) the
360 amplitude of oceanic-path signal is generally larger than that of continental-
361 path signal; (ii) the oceanic-path signal is characterized by a number of
362 high-amplitude peaks that are not visible in the continental-path one; (iii)
363 the large-amplitude portion of the signal lasts longer in oceanic-path than

364 continental-path signal. While the standard deviations of both envelopes
365 are not shown in Fig. 7 in the interest of readability, these inferences are
366 confirmed even if the standard deviation is taken into account. We note
367 that the standard deviation of the oceanic-path envelope is larger than the
368 continental-path one. Observation (ii) reflects several comments made by the
369 subjects (Table 2).

370 **Comments on DS2**

371 The four subjects who achieved the highest scores ($> 55\%$) without training
372 are combined with the four who achieved the highest scores ($> 72\%$) after
373 training, resulting in a group of eight subjects whose verbal comments are
374 summarized in table 3. The group includes three geoscientists, three acous-
375 ticians, and two physicists. Two of the subjects in this group were also in
376 the group discussed in the previous section.

377 Table 3 shows that, despite some contradictory comments, most subjects
378 find strike-slip-fault signals to be characterized by a relatively weak “first
379 arrival” followed by a high-energy coda, while on the contrary they associate
380 thrust events with a strong first arrival followed by a weaker coda. This seems
381 to be consistent with the average envelopes of Fig. 8, where (i) the initial
382 peak is clearly identifiable for both families and is roughly twice as high
383 in the inverse-fault case, with respect to the strike-slip-fault one, while (ii)
384 the later inverse-fault signal is of higher amplitude than its strike-slip-fault
385 counterpart, with a $\sim 20\%$ difference in their main peaks. If the envelopes’
386 standard deviations (not shown in Fig. 8 for clarity) are taken into account,

387 however, this observation cannot be confirmed; more tests, with a broader
388 data set, need to be conducted to come to a definitive conclusion.

389 **Influence of subjects' background on the re-** 390 **sults**

391 Fig. 9 shows that test scores are not strongly affected by the background of
392 subjects. The average score achieved by geoscientists is always (except for
393 the auditory categorization of DS2) slightly higher than that of the other two
394 groups, but the subjects are not numerous enough for this small difference
395 to be considered significant.

396 On the other hand, our analysis of the subjects' recorded descriptions
397 of their categorization strategy shows that acousticians have used about 20
398 more words than both other groups to qualify sounds. We interpret this result
399 as a natural consequence of the acousticians' specific expertise in describing
400 sounds, while geoscientists and physicists usually represent their data only
401 visually. This speculation is confirmed by the study of *Paté et al.* [2017],
402 who conducted a thorough, quantitative analysis of verbal data collected in
403 a similar experiment (also involving audified seismic data, and subjects with
404 similar backgrounds).

405 Discussion

406 Conclusions

407 In our experiments, listeners were exposed to two audified seismic data sets,
408 each characterized by a single, binary control factor: the orientation of the
409 fault (strike-slip or thrust) in one case, the nature of the tectonic plate
410 through which the recorded signal had traveled prior to recording (oceanic
411 or continental) in the other. They were then asked to split each data set into
412 two categories, based on examples of signals associated with different values
413 of the control factor. Purely auditory tests were compared with similar tests,
414 where data were displayed visually rather than acoustically. Overall, listen-
415 ers were able to categorize data based on audition alone. Their performance
416 in visual tests was better, but performance in auditory categorization was
417 significantly improved by a brief training session.

418 Asked to comment on the criteria they had chosen to categorize, listeners
419 most often pointed to perception-based physical features that can be summa-
420 rized as: signal-to-noise ratio (SNR); the duration of what they interpreted
421 to be meaningful signal as opposed to background noise; the frequency con-
422 tent of background noise; the relative amplitude of first seismic “arrivals”
423 with respect to coda. At least two of these features (SNR and meaningful
424 signal duration) do correspond to quantitative parameters that we have been
425 able to define and calculate by simple data processing; we show in Fig. 6a,c
426 that those parameters are differently distributed depending on the value of
427 the relevant control parameter.

428 In summary, human listeners are able to identify geophysically relevant

429 features of audified seismic data, and can be trained to improve their per-
430 formance at such tasks. We cannot yet predict the extent to which training
431 can refine our skills at interpreting the data by listening, but we surmise
432 that auditory display can be useful to a variety of endeavors in seismic data
433 analysis.

434 Outlook

435 While the resolution and pattern recognition capabilities of the human au-
436 ditory system are generally well known [e.g., *Hartmann, 1999; Wang and*
437 *Brown, 2006a*], the seismology community does not entirely appreciate the
438 potential of auditory display as a tool for seismic data analysis. A case in
439 point is the interesting work of *Moni et al. [2012]* and *Moni et al. [2013]*,
440 where an algorithm designed to mimic the human auditory system (in the
441 words of the authors, “to solve the ‘cocktail party problem,’ i.e., separating
442 individual speakers in a room with multiple speakers”) was successfully ap-
443 plied to the problem of identifying different simultaneous microseisms, and
444 yet no attempt was made to use the human auditory system itself, simply
445 listening to the audified data.

446 Besides the benefits derived from exploiting the natural skills of our audi-
447 tory system, audification typically involves the acceleration of a seismic signal
448 by a factor between $\sim 10^2$ and $\sim 10^3$, depending on the frequency content
449 of the original data, which means that an entire day of seismic recording can
450 be listened to in a few minutes with little or no loss of information. Being
451 able to rapidly analyze large sets of data is important, as seismologists are

452 faced today with large and rapidly growing databases. For instance, precisely
453 locating the epicenters of seismic aftershocks requires solving an enormous
454 number of inverse problems [e.g., *Valoroso et al.*, 2013]. This cannot be en-
455 tirely automated if reliable results are to be obtained. All signals recorded by
456 a seismic network can however be listened to simultaneously, by the princi-
457 ples of sound spatialization [e.g. *Peters et al.*, 2011], in an anechoic chamber
458 equipped with a dense speaker network or, more simply, binaurally. The
459 human auditory system is naturally equipped to locate the source of a sound
460 [e.g. *Hartmann*, 1999; *Wang and Brown*, 2006b], and, through this setup, it
461 is reasonable to hypothesize that one might be able to learn to roughly but
462 quickly locate earthquake epicenters (global, regional or local) by listening
463 to sets of audified seismograms. This approach would involve some impor-
464 tant approximations (neglect of dispersion and of Earth lateral heterogeneity
465 effects, etc.), but could be very practical because of its speed and simplicity.

466 The auditory properties of audified seismograms have also been shown to
467 be indicative of several specific seismic processes, including mainshock/aftershock
468 sequences, earthquake swarms that accompany volcanic eruptions, or deep
469 non-volcanic tremors [*Kilb et al.*, 2012; *Peng et al.*, 2012]. Audification is
470 likely to find other potentially important applications in seismology, wher-
471 ever large datasets are to be investigated, and unknown/unexpected patterns
472 recognized. Examples include the analysis of the Earth’s seismic background
473 signal [e.g. *Boschi and Weemstra*, 2015] with implications for monitoring of
474 natural hazards [e.g. *Wegler and Sens-Schonfelder*, 2007; *Brenquier et al.*,
475 2008], and the problem of determining the evolution of a seismic rupture
476 in space and time from the analysis of seismic data [e.g., *Ide*, 2007; *Mai*

477 *et al.*, 2016]. The study of large sets of audified data can further benefit
478 from the possibilities offered by crowd-sourcing platforms: if the sounds are
479 short and meaningful enough, if the listeners’ task is simple enough, and if
480 the data set is correctly distributed among listeners (each sound is given to
481 at least one listener, some are given to several listeners for verification and
482 variability assessment), then a large data set can be effectively explored by
483 the “collaborative” work of a number of listeners.

484 **Data and resources**

485 The image and audio files that were presented to subjects in all the experi-
486 ments described here are available online at <http://hestia.lgs.jussieu.fr/boschil/downloads.html>.

487 The Global Centroid Moment Tensor Project database was searched using
488 www.globalcmt.org/CMTsearch.html (last accessed September 2016).

489 The IRIS database was searched via the Wilber interface at http://ds.iris.edu/wilber3/find_level.html
490 (last accessed September 2016).

491 Figs. 1 and 2 were made using the Generic Mapping Tools version 5.2.1
492 [*Wessel and Smith*, 1991, www.soest.hawaii.edu/gmt/].

493 **Acknowledgements**

494 We acknowledge financial support from INSU-CNRS, and from the European
495 Union’s Horizon 2020 research and innovation programme under the Marie
496 Sklodowska-Curie grant agreement No. 641943 (ITN WAVES). Danièle Dubois
497 and Piero Poli have contributed to our work through many fruitful discus-

498 sions. Zhigang Peng, Debi Kilb and one anonymous reviewer provided careful
499 reviews of our original manuscript. We are grateful to the volunteers who
500 kindly accepted to take our tests.

501 **References**

502 Boschi, L., and C. Weemstra, Stationary-phase integrals in
503 the cross-correlation of ambient noise, *Rev. Geophys.*, *53*,
504 doi:10.1002/2014RG000455, 2015.

505 Brenguier, F., N. M. Shapiro, M. Campillo, V. Ferrazzini, Z. Duputel,
506 O. Coutant, and A. Nercessian, Towards forecasting volcanic eruptions
507 using seismic noise, *Nat. Geosci.*, *1*, 126–130, 2008.

508 Cowen, R., Sound bytes, *Scientific American*, *312*, 4447, 2015.

509 Dombois, F., and G. Eckel, Audification, in *The Sonification Handbook*,
510 edited by T. Hermann, A. Hunt, and J. G. Neuhoff, pp. 301–324, Berlin:
511 Logos Publishing House, 2011.

512 D’Orazio, D., S. De Cesaris, and M. Garai, A comparison of methods to
513 compute the effective duration of the autocorrelation function and an al-
514 ternative proposal, *J. Acoust. Soc. Am.*, *130*, 19541961, 2011.

515 Ekström, G., M. Nettles, and A. M. Dziewoński, The global CMT project
516 2004-2010: Centroid-moment tensors for 13,017 earthquakes, *Phys. Earth*
517 *Planet. Inter.*, *200-201*, 1–9, 2012. doi:10.1016/j.pepi.2012.04.002 , 2012.

- 518 Frantti, G. E., and L. A. Leverault, Auditory discrimination of seismic signals
519 from earthquakes and explosions, *Bull. Seism. Soc. Am.*, *55*, 1–25, 1965.
- 520 Gaillard, P., Laissez-nous trier ! TCL-LabX et les tâches de catégorisation
521 libre de sons, in *Le Sentir et le Dire*, edited by D. Dubois, pp. 189–210,
522 L’Harmattan, Paris, France, 2009.
- 523 Hartmann, W. M., How we localize sound, *Phys. Today*, *11*, 24–29, 1999.
- 524 Holtzman, B., J. Candler, M. Turk, and D. Peter, Seismic sound lab: Sights,
525 sounds and perception of the earth as an acoustic space, in *Sound, Music,
526 and Motion*, edited by M. Aramaki, O. Derrien, R. Kronland-Martinet,
527 and S. Ystad, pp. 161–174, Springer International Publishing, 2014.
- 528 Ide, S., Slip inversion, in *Treatise of Geophysics, Vol. 4*, edited by
529 H. Kanamori, pp. 193–223, Elsevier, Amsterdam, 2007.
- 530 Kennett, B. L. N., and T. Furumura, High-frequency *Po/So* guided waves
531 in the oceanic lithosphere: I-long distance propagation, *Geophys. J. Int.*,
532 *195*, 1862–1877, doi: 10.1093/gji/ggt344, 2013.
- 533 Kennett, B. L. N., T. Furumura, and Y. Zhao, High-frequency *Po/So* guided
534 waves in the oceanic lithosphere: II-heterogeneity and attenuation, *Geo-
535 phys. J. Int.*, *199*, 614–630, doi: 10.1093/gji/ggu286, 2014.
- 536 Kilb, D., Z. Peng, D. Simpson, A. Michael, and M. Fisher, Listen,
537 watch, learn: SeisSound video products, *Seismol. Res. Lett.*, *83*, 281–286,
538 doi:10.1785/gssrl.83.2.281, 2012.

- 539 Mai, P. M., et al., The earthquake-source inversion validation (SIV) project,
540 *Seism. Res. Lett.*, *87*, 690–708, doi:10.1785/0220150,231, 2016.
- 541 Moni, A., , C. J. Bean, I. Lokmer, and S. Rickard, Source sep-
542 aration on seismic data, *IEEE Signal Process Mag.*, *29*, 16–28,
543 doi:10.1109/MSP.2012.2184,229, 2012.
- 544 Moni, A., D. Craig, and C. J. Bean, Separation and location of microseism
545 sources, *Geophys. Res. Lett.*, *40*, 3118–3122, doi:10.1002/grl.50,566, 2013.
- 546 Paté, A., L. Boschi, J. L. le Carrou, and B. Holtzman, Categorization of
547 seismic sources by auditory display: a blind test, *International Journal*
548 *of Human-Computer Studies*, *85*, 57–67, doi:10.1016/j.ijhcs.2015.08.002,
549 2016.
- 550 Paté, A., L. Boschi, D. Dubois, B. Holtzman, and J. L. le Carrou, Auditory
551 display of seismic data: Expert categorization and verbal description as
552 heuristics for geoscience, *J. Acoust. Soc. Am.*, *accepted*, 2017.
- 553 Peng, Z., C. Aiken, D. Kilb, D. Shelly, and B. Enescu, Listening to the 2011
554 magnitude 9.0 Tohoku-Oku, Japan earthquake, *Seismol. Res. Lett.*, *83*,
555 287–293, doi:10.1785/gssrl.83.2.287, 2012.
- 556 Peters, N., G. Marentakis, and S. McAdams, Current technologies and com-
557 positional practices for spatialization: A qualitative and quantitative anal-
558 ysis, *Computer Music Journal*, *35*, 10–27, 2011.
- 559 Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numer-*
560 *ical Recipes in Fortran 77*, Cambridge University Press, 1992.

- 561 Speeth, S. D., Seismometer sounds, *J. Acoust. Soc. Am.*, *33*, 909–916,
562 doi:10.1121/1.1908,843, 1961.
- 563 Tang, Y., Data sonification with the seismic signature of ocean surf, *The*
564 *Leading Edge*, *33*, doi:10.1190/tle33101,128.1, 2014.
- 565 Thorndike, E. L., *Human Learning*, Century, New York, 1931.
- 566 Valoroso, L., L. Chiaraluce, D. Piccinini, R. D. Stefano, D. Schaff, and
567 F. Waldhauser, Radiography of a normal fault system by 64,000 high-
568 precision earthquake locations: The 2009L’Aquila (central Italy) case
569 study, *J. Geophys. Res.*, *118*, 1156–1176 doi:10.1002/jgrb.50,130, 2013.
- 570 Volmar, A., Listening to the Cold War: The Nuclear Test Ban negotiations,
571 seismology, and psychoacoustics, 1958-1963, *Osiris*, *28*, 80–102, 2013.
- 572 Wang, D., and G. J. Brown, Fundamentals of computational auditory scene
573 analysis, in *Computational Auditory Scene Analysis*, edited by D. Wang
574 and G. J. Brown, pp. 1–44, Wiley & Sons, Hoboken, N. J., 2006a.
- 575 Wang, D., and G. J. Brown, *Computational Auditory Scene Analysis*, Wiley
576 & Sons, Hoboken, N. J., 2006b.
- 577 Wegler, U., and C. Sens-Schonfelder, Fault zone monitoring with passive
578 image interferometry, *Geophys. J. Int.*, *168*, 1029–1033, 2007.
- 579 Wessel, P., and W. H. F. Smith, Free software helps map and display data,
580 *EOS Trans. Am. Geophys. Union*, *72*, 445–446, 1991.

Table 1: Summary of listening experiments.

data set	number of waveforms	subjects			audio	visual	training
		A	G	P			
DS1	40	18	9	8	yes	yes	no
DS2	40	15	5	7	yes	yes	no
DS2	36	10	3	4	yes	no	yes

The first two columns to the left indicate how many signals from which data set were presented to the subjects. The letters A, G and P stand for “acousticians,” “geoscientists” and “physicists,” respectively; “audio” and “visual” indicate which type(s) of data were provided to the subjects; “training” refers to whether subjects were trained before taking the test.

Table 2: Listeners’ comments on DS1.

Family A (oceanic paths)	Family B (continental paths)
second shock very close to the first with an echo / rebound a lot of background noise high-pitched background noise background noise longer signal	echo of the first impact’s sound small rebounds little background noise low-pitched background noise shorter and duller sound sharper and shorter rising perceived frequency faster arrival buzz or intense reverberation after the explosion

Summary of written, verbal explanations given by 5 subjects (scoring $\geq 80\%$) concerning their auditory categorization of DS1. All text was originally in French and has been translated into English as literally as possible.

Table 3: Listeners' comments on DS2.

Family A (strike-slip events)	Family B (thrust events)
first shock weaker than second one wave of rising frequency louder than the first heard shock after the detonation, sound decays more slowly faster attack and decay significant intensity even after a long time lower-frequency shock duller signal higher frequencies	louder low frequencies first shock louder than the second one first shock louder than the wave more powerful and present sound sound decays quickly after the detonation slower decay

Summary of written, verbal explanations given by 8 subjects (scoring $\geq 80\%$) for the auditory categorization of DS2 before (4 subjects scoring $> 55\%$) and after training (4 subjects scoring $> 72\%$). Again, the original French text was translated into English.

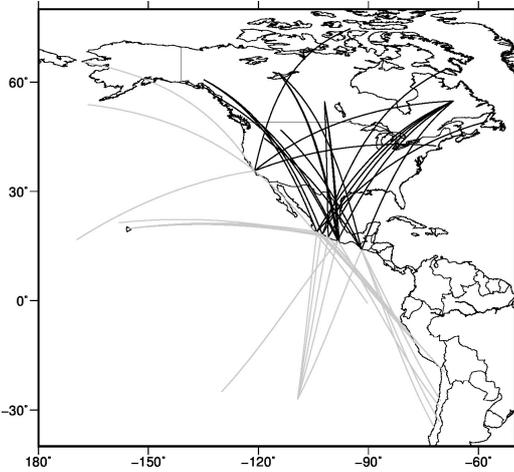


Figure 1: Surface projections of ray paths associated with audified data set DS1. DS1 consists of recordings of events occurring along the west coast of Mexico, made at stations at epicentral distances of ~ 4000 to 6000 km; recordings made at north American stations correspond to ray paths only traversing continental terrain (black lines), while stations along the Pacific coast or on ocean islands result in purely oceanic paths (grey lines).

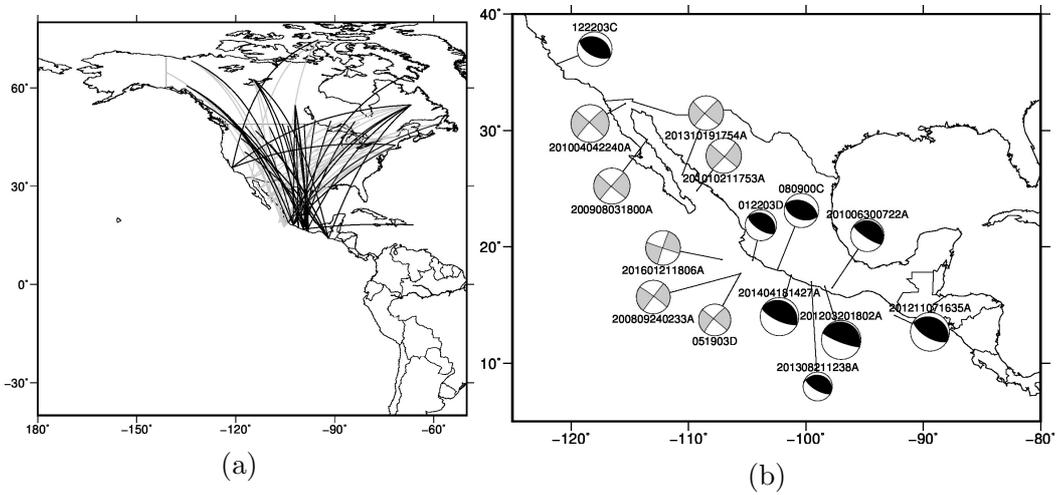


Figure 2: (a) Same as Fig. 1, but for data set DS2, which only includes recordings made at stations within the north American continent, of either strike-slip (grey ray path curves) or thrust (black) events. Their epicenters and focal mechanisms [Ekström *et al.*, 2012] are shown in (b) using the same color code.

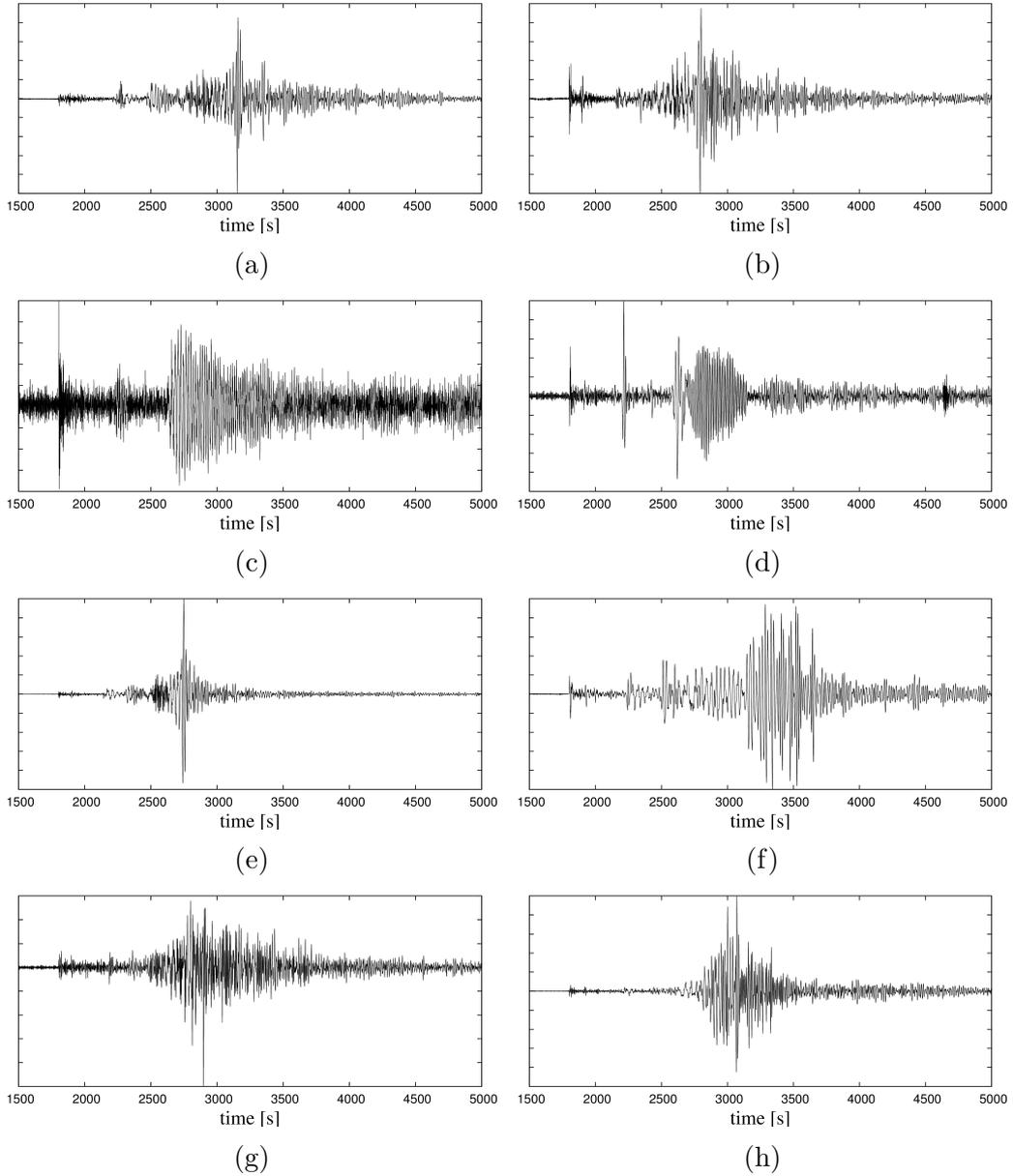
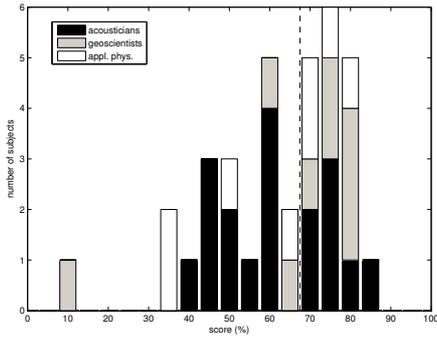
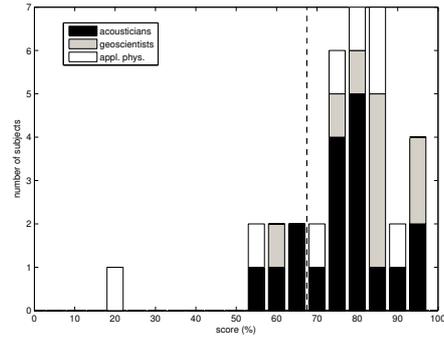


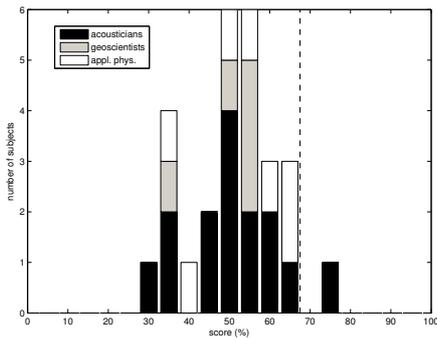
Figure 3: Examples of seismograms used in our study. (a) and (b): DS1, continental paths. (c) and (d) DS1: oceanic paths. (e) and (f): DS2, thrust faults. (g) and (h) DS2, strike-slip faults. The vertical axis is not labeled as we systematically normalize all seismograms (both visual and audio). In our visualization experiments, the horizontal axis was less exaggerated and the time span much longer, so that in principle the exact same information was provided to subjects in visualization and listening tests. The images files used in experiments are available online (see “Data and resources” section).



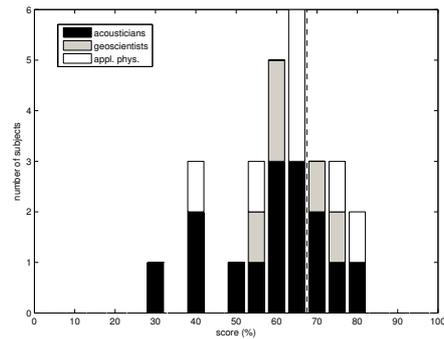
(a) DS1, auditory



(b) DS1, visual



(c) DS2, auditory



(d) DS2, visual

Figure 4: These histograms summarize the results of the constrained categorization experiments conducted before training on audified seismograms from data sets DS1 (panels (a) and (b)) and DS2 ((c) and (d)). Scores achieved in auditory tests are shown in panels (a) and (c); scores achieved in visual tests are shown in (b) and (d). The vertical dashed line marks the “99% confidence level,” i.e. the probability of achieving at least that score by categorizing the signal at random is less than 1%. Colors correspond to the different background of subjects, as explained in the inset.

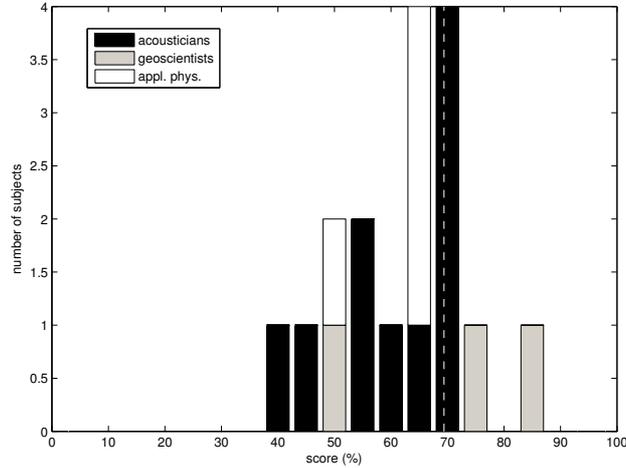


Figure 5: Histogram summarizing the results of the constrained categorization experiment conducted (on DS2) after training. The vertical dotted line, corresponding to a score of 69.4%, marks the 99% confidence level; all scores in the 67.5%-to-72.5% bin actually fall to its right.

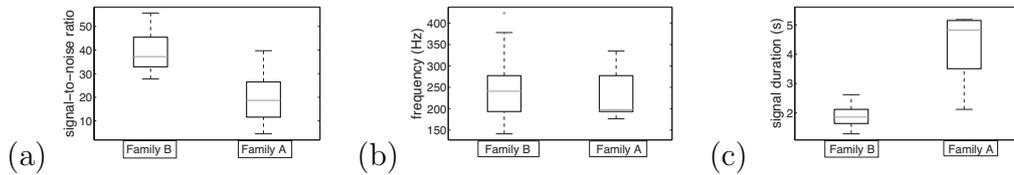


Figure 6: Distributions, shown as box-plots, of three physical parameters, corresponding to properties of the signal that subjects tend to describe as important: (a) SNR; (b) dominant frequency of background noise; (c) duration of meaningful signal. For each parameter, the distributions of parameter values for oceanic-path (“Family A) and continental-path (“Family B) signal are shown separately. Distributions are summarized by their median (thick grey segments), first and third quartiles (upper and lower sides of boxes), and minimum and maximum values (endpoints of dashed lines). Values that we neglect as outliers (their absolute value is more than 1.5 times the interquartile distance) are denoted by grey crosses.

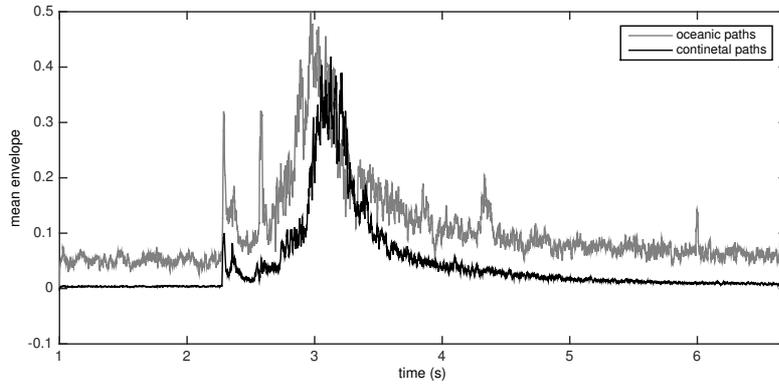


Figure 7: Signal envelope averaged over all DS1 audified seismograms corresponding to continental (black line) vs. oceanic (grey) paths.

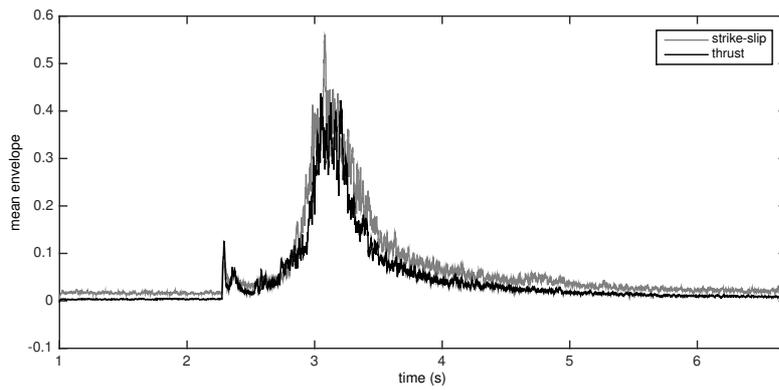


Figure 8: Same as Fig. 7, but envelopes are averaged over all DS2 signals originated from thrust (black line) vs. strike-slip (grey) events.

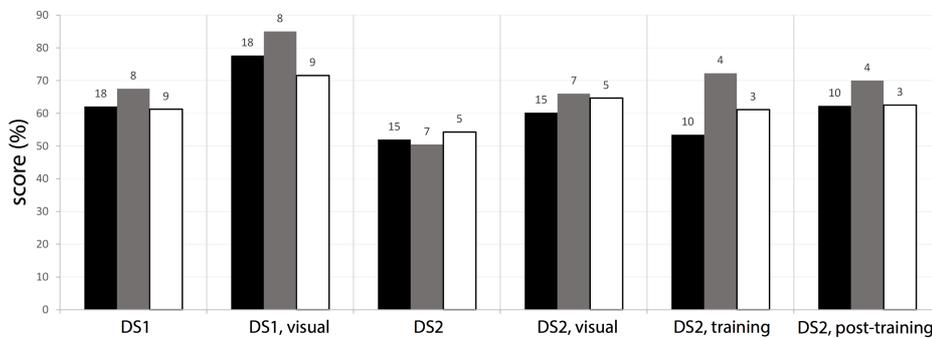


Figure 9: Average scores by test (left to right, as indicated under each bar) and by subjects' background group. Black, grey and white bars are associated with acousticians, geoscientists and physicists, respectively. The number of subjects participating to a test is shown above the corresponding bar. The label "post-training" refers to the auditory test of DS2 conducted after the training session; "training" refers to answers that were given *during* the aforementioned training session.