



HAL
open science

Comparing the Auditory Distance and Externalization of Virtual Sound Sources Simulated Using Nonindividualized Stimuli

Mathieu Lavandier, Lizette Heine, Fabien Perrin

► **To cite this version:**

Mathieu Lavandier, Lizette Heine, Fabien Perrin. Comparing the Auditory Distance and Externalization of Virtual Sound Sources Simulated Using Nonindividualized Stimuli. *Trends in Hearing*, 2024, 28, 10.1177/23312165241285695 . hal-04817987

HAL Id: hal-04817987

<https://hal.science/hal-04817987v1>

Submitted on 4 Dec 2024

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 **Comparing the auditory distance and externalization of virtual sound sources simulated using**
2
3 **non-individualized stimuli**

4 Mathieu Lavandier,^{1, a} Lizette Heine,^{1,2} and Fabien Perrin²

5
6 ¹*ENTPE, Ecole Centrale de Lyon, CNRS, LTDS, UMR5513, 69518 Vaulx-en-Velin,*
7
8 *France*

9 ²*Audition Cognition and Psychoacoustics Team, Lyon Neuroscience Research Center,*
10 *INSERM, U1028, CNRS, UMR5292, Lyon, France*

11 When reproducing sounds over headphones, the simulated source can be
12 externalized (i.e. perceived outside the head) or internalized (i.e. perceived within
13 the head). Is it because it is perceived as more or less distant? To investigate this
14 question, 18 participants evaluated distance and externalization for three types of
15 sound (speech, piano, helicopter) in 27 conditions using non-individualized stimuli.
16
17 Distance and externalization ratings were significantly correlated across conditions
18 and listeners, and when averaged across listeners or conditions. However, they were
19 also decoupled in some circumstances: (1) Sound type had different effects on distance
20 and externalization: the helicopter was evaluated as more distant, while speech was
21 judged as less externalized. (2) Distance estimations increased with simulated distan-
22 ces even for stimuli judged as internalized. (3) Diotic reverberation influenced distance
23 but not externalization. Overall, a source was not rated as externalized as soon as and
24 only if its perceived distance exceeded a threshold (e.g. the head radius). These results
25 suggest that distance and externalization are correlated but might not be aspects of a
26 single perceptual continuum. In particular, a virtual source might be judged as both
27 internalized and with a distance. Hence, it could be important to avoid using a scale
28 related to distance when evaluating externalization.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 **Keywords:** spatial perception; hearing; virtual acoustics; localization
48
49
50
51
52
53
54

INTRODUCTION

Usually, a sound is experienced as coming from a specific point in the space around the listener (the source position), with a direction (azimuth and elevation) and a distance. The use of headphones often gives the impression that the simulated sound source is “internalized”, perceived within the head (Jeffress and Taylor, 1961). Using binaural synthesis, it is possible to create an externalized percept, simulating a source perceived outside of the head (Blauert, 1997; Durlach *et al.*, 1992). A recent review discussed the subjective experience of sound externalization, its definition and measurements (Best *et al.*, 2020). The authors review a wealth of knowledge on sound externalization, but acknowledge that the link between auditory distance and externalization is not clearly established yet. Specifically, it is not clear whether externalization arises simply because the sound source has a (simulated) distance from the listener. The aim of the present study was to further investigate this link.

Some authors assume that the percepts of distance and externalization are aspects of a single perceptual continuum (Durlach *et al.*, 1992; Hartmann and Wittenberg, 1996), where the center of the head is internalized and is the 0-m reference used to evaluate distance (Best *et al.*, 2020), and where a source is externalized as soon (and only if) its perceived distance exceeds the head radius. The present study aimed to challenge this idea by systematically comparing ratings of perceived distance and externalization. The conditions tested here were not used to highlight any new effect on distance or externalization. They were chosen based on the effects previously highlighted in the literature (which could reveal differences in behavior between distance and externalization), in studies that measured either distance or externalization. Here, they were measured in parallel by the same listeners on the same stimuli, so that they could be

102 directly compared.
103

104 Overall, we expected that distance and externalization ratings would be correlated, as similar
105 activation patterns in the temporal gyri have been reported for distance and externalization tasks
106 (Callan *et al.*, 2013; Kopčo *et al.*, 2012), even if the corresponding brain activations do not completely
107 overlap. However, we also expected that distance and externalization would not be perfectly
108 correlated, and that externalization would not just be a binary version of distance either, that is to
109 say that a sound source would not be rated as externalized as soon as and only if its perceived
110 distance exceeded a threshold (e.g. the head radius). This would indicate that, while distance and
111 externalization are not independent percepts, they might not be aspects of a single perceptual
112 continuum either. In accordance with this hypothesis, distance judgments seem possible when
113 the percept of externalization is weak, such as with diotic sounds or frontal sources (e.g., Bidart
114 and Lavandier, 2016; Kopčo *et al.*, 2020; Prud'homme and Lavandier, 2020). Moreover,
115 it has been established that lateralizing the source enhances its externalization (Brimijoin
116 *et al.*, 2013; Kates *et al.*, 2018; Leclère *et al.*, 2019), while distance perception appears
117 rather similar for frontal and lateral sources (Zahorik, 2002), except at short distance (<
118 1 m; Brungart *et al.*, 1999; Kopčo *et al.*, 2020). On the one hand, externalization ratings
119 seem to depend mainly on binaural cues (Best *et al.*, 2020), in particular their reverberation-
120 induced variations (Catic *et al.*, 2013; Leclère *et al.*, 2019; Li *et al.*, 2019). Leclère *et al.* (2019)
121 showed that reverberation enhances externalization only if it creates signal differences across
122 the two ears, so that diotic reverberation does not improve externalization. On the other
123 hand, distance perception is dominated by monaural cues (Kolarik *et al.*, 2016; Zahorik *et al.*,
124 2005), in particular sound level and direct-to-reverberant ratio (DRR). While reverberation
125 produces monaural cues crucial for distance perception, it seems that the reverberation-

148 induced variations in binaural cues do not influence distance judgments for frontal sources
149 (Bidart and Lavandier, 2016; Prud'homme and Lavandier, 2020). Interaural differences
150 seem to influence distance only for very close lateral sources (< 1 m; Brungart *et al.*, 1999).
151 Finally, the vastly different scales used to evaluate externalization further complicate the
152 issue, in particular when distance labels are used for the externalization scale (for a review
153 of externalization evaluations see Best *et al.*, 2020).
154
155
156
157
158
159
160

161 While sound level is a strong distance cue, not much is known about its effect on exter-
162 nalization. Most externalization studies have used stimuli that did not vary in level, because
163 the sources were simulated at a unique distance from the listener (e.g., Brimijoin *et al.*, 2013;
164 Hendrickx *et al.*, 2017; Kates *et al.*, 2018) or because stimuli were equalized in level (e.g.,
165 Catic *et al.*, 2015; Hartmann and Wittenberg, 1996; Leclère *et al.*, 2019). Catic *et al.* (2013)
166 tested stimuli that varied in broadband level but also in bandwidth, so that it is not possible
167 to isolate the potential effect of level on externalization ratings. Hartmann and Wittenberg
168 (1996) mentioned a problem with varying level when evaluating externalization, at least for
169 random level variations that were informally reported as making the sounds “jump around
170 the room, sometimes jumping into the head”. When considering the influence of level on
171 distance, it is also important to keep in mind its potential interaction with the listeners’
172 experience and their level expectations depending on the type of source considered. Com-
173 paring whispered, conversational and shouted speech, Brungart and Scott (2001) showed
174 that speech type can have more influence on distance estimates than level (which had no
175 influence for whispered speech, perceived at a fixed close distance). Leclère *et al.* (2019) did
176 not highlight such strong effect of sound type on externalization ratings when comparing
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192

193 noise, speech, music and environmental sounds (clinking bottles).

194
195
196
197 The aim of the present study was to better understand the link between the percepts of
198 distance and externalization by comparing distance and externalization ratings on the same
199 set of stimuli. Participants were asked to judge the same 27 simulated auditory conditions on
200 distance and externalization tasks. These conditions varied in terms of source azimuth and
201 (simulated) distance, level of reverberation, and amount of binaural information present in the
202 stimuli. Three types of sound (speech, piano, helicopter) differing in terms of level expectations
203 were convolved with binaural room impulse responses (BRIRs) and anechoic head-related
204 transfer functions (HRTFs) corresponding to different simulated azimuths (and distances in
205 the room). Some BRIR-convolved stimuli were also averaged across ears to investigate diotic
206 reverberation. Finally, diotic anechoic stimuli were tested as reference of internalized sound
207 sources.
208
209
210
211
212
213
214
215
216
217

218
219 As the influence of sound level on externalization was not known, all the stimuli were
220 equalized in level. [Arend *et al.* \(2021\)](#) showed that level or loudness equalization does not
221 fully remove all level cues for distance. These cues could mask the influence of more subtle
222 cues, such as ILDs for nearby lateral sources, and even lead to distance ratings varying
223 inversely to the simulated distance in anechoic conditions where few other distance cues
224 were available. To eliminate the use of level cues, researchers can choose to rove the sound
225 level that thus varies randomly from trial to trial ([Brungart *et al.*, 1999](#); [Kopčo and Shinn-
226 Cunningham, 2011](#)). Here, the aim was not to remove level cues for distance, but to avoid
227 level variations that could severely impair the evaluation of externalization ([Hartmann and
228 Wittenberg, 1996](#)), thus roving level was not an option. We chose to equalize the overall
229
230
231
232
233
234
235
236
237

238 broadband level of the stimuli, as previous studies have shown that it can still allow for
239 reliable distance ratings especially when reverberation provides for additional distance cues
240
241 ([Akeroyd et al., 2007](#); [Bidart and Lavandier, 2016](#); [Mershon and Bowers, 1979](#); [Prud'homme
242 and Lavandier, 2020](#)). While overall level was equalized across conditions, this was not the
243 case for the direct and reverberated sound levels and the DRR that could still be used as
244
245 cues.
246
247
248
249

250
251 Even if early studies on externalization pointed to the importance of reproducing with
252 headphones the ear signals as they would be produced by real sources, the magnitude of the
253 effect of the individualization of the stimuli on externalization is still unclear. Some studies
254 mentioned some improvements in externalization when stimuli were individualized (using
255 the listeners own HRTFs) compare to generic (e.g. using measurements from a manikin),
256
257 in particular for less externalized frontal sources ([Kim and Choi, 2005](#); [Werner et al., 2016](#)).
258
259 However, the individualization of stimuli improved only marginally the externalization
260 ratings measured by [Cubick et al. \(2015\)](#) and [Leclère et al. \(2019\)](#). [Begault et al. \(2001\)](#) did not
261 observe any significant effect of individualization when measuring externalization in anechoic
262 and reverberant conditions. This was also the case for [Kates et al. \(2018\)](#), who showed that a
263 generic binaural model was as effective as individualized stimuli to produce an externalized
264 image in reverberant conditions. Reviewing recent literature on externalization, [Best et al.
265 \(2020\)](#) concluded that individual spectral cues may not be critical for externalization in realistic
266 listening conditions. Concerning distance perception, [Zahorik \(2002b\)](#) showed that listeners'
267 performance in judging distance was not impaired (in terms of both the individual estimates
268
269

280 and their standard deviations) by using the HRTFs of another listener; while [Prudhomme and](#)
281 [Lavandier \(2020\)](#) did not find any significant difference in distance estimates for individualized
282 stimuli compared to stimuli based on manikin measurements. The present study only used such
283 non-individualized stimuli. Individualized measurements are not yet available to everybody,
284 and certainly not on a large scale, so any applications associated with externalization and virtual
285 distance could benefit from knowledge on the externalization and perceived distance of
286 non-individualized stimuli.

294
295 The distance and externalization ratings for all the stimuli were systematically compared
296 by computing their correlation. The hypothesis was also tested that externalization ratings
297 could result directly from the distance ratings, the source being rated as externalized as
298 soon as the distance rating exceeds a threshold (e.g. the head radius). Finally, distance
299 and externalization were compared by investigating their potential differences in behavior in
300 specific conditions: (1) One could expect an effect of sound type on distance estimates due
301 to a priori level expectations from the listeners ([Brungart and Scott, 2001](#)), e.g. the helicopter
302 might be judged at further distances to produce the same sound level as speech at the ears. No
303 such effect was expected on externalization ratings ([Leclère et al., 2019](#)). (2) We expected higher
304 externalization ratings for lateralized sources compared to frontal sources ([Brimijoin et al.,](#)
305 [2013](#); [Kates et al., 2018](#)), while the perception of distance was not expected to change much
306 with azimuth for the distances tested (above 1 m; [Zahorik, 2002](#)). (3) BRIRs measured
307 at different distances in the same room were expected to trigger differences in perceived
308 distance ([Prud'homme and Lavandier, 2020](#); [Zahorik, 2002](#)), but a relatively fixed high level
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323

324 of externalization compared to sounds convolved with anechoic HRTFs (in particular for
325 lateral sources; [Leclère et al., 2019](#)). (4) For the diotic versions of the BRIR stimuli, perceived
326 distance was still expected to vary with simulated distance ([Bidart and Lavandier, 2016](#)), but
327 externalization was not expected to be enhanced compared to the anechoic stimuli ([Leclère](#)
328 [et al., 2019](#)).

339 METHODS

342 Stimuli

344 Non-individualized BRIRs measured by [Leclère et al. \(2019\)](#) in a gym (33.7 m x 44.5
345 m x 10.5 m) at three azimuths (0°, -30°, and 60°) and three distances (1, 3, and 5 m) from a
346 manikin were used. They were measured using a log sine sweep technique ([Farina, 2000](#)),
347 with a 15-s sweep duration and 20Hz-20kHz frequency range. The signal was played through
348 a loudspeaker (Tannoy System 8 NFM 2) at the desired location and recorded at the simu-
349 lated listener position using a MK2/NCF1 dummy head (Neutrik Cortex Instrument). The
350 average broadband reverberation time was 1.35 s. Room acoustical characteristics further
351 describing the BRIRs can be found in the supplementary material (Suppl. table 1), while
352 the room layout and measurement details were presented by [Leclère et al. \(2019\)](#).

362 To simulate anechoic stimuli, HRTFs were used. They were measured with a KEMAR
363 manikin and a loudspeaker at 1.4 m by [Gardner and Martin \(1994\)](#). Different manikins and
364 loudspeakers were used for the BRIR and HRTF measurements, leading to differences in
365
366
367
368

369 the spectrum of the corresponding direct sounds, while the spectrum of the whole *BRIR*
370 stimuli was further influenced by reverberation (Suppl. fig. 2). Due to a programming
371 error, the HRTFs used here were measured at different azimuths than the BRIRs: 0°, -60°, and
372 30°. The HRTFs are left/right symmetric, so that the lateral sources based on the
373 HRTFs were mirror images of what was intended. There is no reason to expect a
374 left/right asymmetry in the distance and externalization tasks (Begault *et al.*, 2001; Parseihian *et*
375 *al.*, 2014; Arend *et al.*, 2021; Best and Roverud, 2024), as will be further discussed below. The
376 corresponding data were thus mirror-imaged (the data for the -60° stimuli were assigned to the
377 60° condition, and the data for the 30° stimuli were assigned to the -30° condition) to allow for a
378 comparison between the anechoic and reverberant conditions.
379
380
381
382
383
384
385
386
387
388
389

390 Three types of sound were considered: a short speech excerpt also used in previous
391 studies (“Toute la nuit” meaning “All night long”, duration 0.9 s; Bidart and Lavandier,
392 2016; Leclère *et al.*, 2019), the sound of a piano from the NESSTI database (1 s; Hocking
393 *et al.*, 2013), and the sound of a helicopter (1 s; Sound-Ideas-Series, 1992).
394
395
396

397
398
399 For each of these sounds, 27 processing conditions were considered, as summarized in
400 Table I. The original diotic anechoic signal was used as a reference for a source that should
401 be very internalized. Twelve signal versions were created through convolution with the nine
402 BRIRs (3 distances x 3 azimuths) and the three HRTFs (3 azimuths). Six diotic versions of
403 the *BRIR* signals were also created by averaging the left and right channels of the sound produced
404 by the lateral sources at the three distances. Finally, 8 lateralized versions of the diotic stimuli
405 were created by adding broadband interaural level/time differences (ILD/ITD) into the diotic
406 *BRIR* signals, as well as into diotic versions of the *HRTF* signals (lateral sources only). The
407
408
409
410
411
412
413
414

415 broadband ILD/ITD values used in both cases were measured in the corresponding HRTFs.

416 The *ITDILD* stimuli contain coarse binaural cues that are broadband and constant over time.

417 They were included to test whether lateral sources could be more externalized only because they

418 are simulated on the side, even in the absence of the time-varying frequency-dependent binaural

419 cues associated with reverberation (Catic *et al.*, 2013; Leclère *et al.*, 2019; Li *et al.*, 2019).

420 Best *et al.* (2020) hypothesized such a lateralization bias when mentioning that “a listener may not

421 be inclined to give a rating of zero for the lateral sounds, which may introduce a bias toward

422 higher externalization ratings”.

423 All stimuli were equalized in overall level such that the average of the root-mean-square

424 (RMS) power of the left and right ear signals was set to the same level. The specific gains used

425 as a function of simulated distance for the BRIR stimuli are provided in Suppl. table 2. The

426 virtual source produced the same broadband (averaged across ears) level independently of

427 its distance and of the signal it reproduced. This equalization did not affect the variations in

428 DRR and source spectra associated with simulated distance, also preserving any ILD present

429 in the stimuli. Moreover, while the overall level was constant with simulated distance, the

430 direct and reverberated sound levels varied and could also be used as distance cues (Suppl.

431 fig. 1). Note that the variations in direct sound level were reduced by the equalization,

432 but the direct sound level of the equalized stimuli still decreased with simulated distance,

433 so it would not mislead the listeners by providing a cue varying inversely to what they

434 would expect (Arend *et al.*, 2021). All stimuli were sampled at 44.1 kHz, D/A converted

435 and amplified using a Lynx TWO sound card and delivered through headphones (HD 650

436 Sennheiser; Wedemark, Germany) at 60 dB SPL (calibrated using the MK2 dummy head).

TABLE I. Details of the 27 conditions tested for the 3 sound types (speech, piano, helicopter)

label	simulated azimuth	reverberation level/ simulated distance	binaural information/ listening mode
<i>REF</i>		anechoic, non-spatialized	diotic
<i>HRTF_0</i>	0°	anechoic at 1.4 m	all binaural info
<i>HRTE_30</i>	-30°	anechoic at 1.4 m	all binaural info
<i>HRTF_60</i>	60°	anechoic at 1.4 m	all binaural info
<i>ITDILD_HRTF_30</i>	-30°	anechoic at 1.4 m	broadband ITD & ILD
<i>ITDILD_HRTF_60</i>	60°	anechoic at 1.4 m	broadband ITD & ILD
<i>BRIR1m/3m/5m_0</i>	0°	room at 1, 3, or 5 m	all binaural info
<i>BRIR1m/3m/5m_30</i>	-30°	room at 1, 3, or 5 m	all binaural info
<i>BRIR1m/3m/5m_60</i>	60°	room at 1, 3, or 5 m	all binaural info
<i>ITDILD_BRIR1m/3m/5m_30</i>	-30°	room at 1, 3, or 5 m	broadband ITD & ILD
<i>ITDILD_BRIR1m/3m/5m_60</i>	60°	room at 1, 3, or 5 m	broadband ITD & ILD
<i>Dio_BRIR1m/3m/5m_30</i>		room at 1, 3, or 5 m	diotic
<i>Dio_BRIR1m/3m/5m_60</i>		room at 1, 3, or 5 m	diotic

464 **Procedure**
465
466

467 Two experimental tasks were used, each in a different experimental session. In each
468 task, the participants were asked to imagine a loudspeaker that produces different types of
469 sound (speech, helicopter, and piano), to close their eyes while listening to the sound played,
470 and then open their eyes to respond to the question displayed on a screen in front of them.
471 The participants were told that each sound could be replayed as much as they liked. Only the
472 question changed between the two tasks.
473
474
475
476
477
478
479

480 For the distance task, listeners were asked to indicate the distance of the virtual source in
481 meters by entering a number (using decimals if needed). No reference was provided, except
482 that they were told to indicate a distance of 0 m when they perceived the sound to originate
483 from within their head. For the externalization task, they were asked to evaluate whether
484 the sound seemed to be originating from inside or outside their head (binary response, coded
485 as 0 for internalized and 1 for externalized), and were also asked to indicate their degree of
486 confidence in this answer using a slider on a horizontal line going from “not at all confident”
487 to “very confident”. The analyses associated with this degree of confidence are not presented
488 below, because confidence ratings were generally high and constant across conditions so that
489 their analysis did not add anything to the analysis of the binary ratings.
490
491
492
493
494
495
496
497
498
499

500 A short practice session using six stimuli (the 3 *REF* and the 3 *BRIR5m_60*) was
501 included to familiarize the participants with the testing environment. All sessions were per-
502 formed in a double-walled soundproof booth, with a screen visible through a window and
503 access to a mouse and keyboard within the booth. The participant was always facing the
504 screen placed about 1 m away. The order of sessions was randomized across participants,
505 who did the two sessions the same day. For each session, the 27 processing conditions
506 were repeated twice for each sound type. Thus, participants responded to one question
507 for 162 sounds (27x3x2) during each of the two sessions. The experiment lasted on av-
508 erage 1.5 hours per participant (including breaks, practices, instructions and audiogram
509 measurement).

527 **Listeners**

531 Eighteen university students (mean age=22 years old, SD=3 years; 10 female) partici-
532 pated in this study. They had audiometric thresholds <20 dB hearing level (HL) at octave
533 frequencies between 125 and 8000 Hz. Note that 6000 Hz was also tested and a threshold of
534 25 dB HL only at this frequency for one listener in one ear was observed. All participants
535 signed a written informed consent and were compensated for their participation.

541 **Data analyses**
542

543 Because perceived distance varies as a power function of simulated distance (Zahorik,
544 2002; Zahorik *et al.*, 2005), distance ratings d were log-transformed before all analyses. The
545 transformation $\ln(1 + d)$ was used as 0 m was a valid response (Prud'homme and Lavandier,
546 2020). Statistical analysis and calculations of mean and standard error were done using
547 $\ln(1 + d)$. The values of means and error bars presented in the figures correspond to the
548 inverse transform of $\ln(1 + d)$ applied to the mean, the mean plus the standard error and
549 the mean minus the standard error calculated with $\ln(1 + d)$.
550
551
552
553
554
555
556

557 The externalization and (log-transformed) distance ratings were systematically compared.
558 First, the raw data were analyzed. A logistic regression was used to evaluate the extent to which the
559 (binary) externalization ratings could be predicted from the distance ratings, while controlling
560 for the various effects of the experimental conditions. The externalization data were fitted with
561 a generalized linear mixed-effects model having the (log-transformed) perceived distance and the
562 experimental factors as fixed effects and the participants as a random effect (using *lme4::glmer* in
563 R version 4.2.1). The point-biserial correlation between externalization and distance ratings was
564 also computed.
565
566
567
568
569
570
571
572

573 Correlation analyses were then performed on the data averaged across repetitions and
574 sound types, because the statistical analyses did not reveal any interaction between the effects of
575 sound type and processing conditions. This averaging allowed for the externalization ratings to be
576 more “continuous” rather than limited to 3 discrete values when averaged only across the 2
577 repetitions (0, 0.5 and 1), thus more suitable for Pearson correlation computation. The mean
578 externalization rating is informative of the proportion of externalized responses in a given
579 condition (Brimijoin *et al.*, 2013). The mean distance and externalization ratings were compared
580 by computing their Pearson and Spearman correlation coefficients (r_P and r_S) across processing
581
582
583
584
585
586
587
588
589

590 conditions and listeners, or averaged across listeners or processing conditions. Because three
591 correlations were considered, the alpha value used to evaluate their significance was Bonferroni
592 corrected to $0.05/3 = 0.0166$.
593
594

595 To test whether a source was rated as externalized as soon as and only if its distance
596 rating exceeded a distance threshold (e.g. the head radius), the distance ratings were also
597 used to create binary distance ratings. For each distance rating, the binary rating was set to
598 0 (inside the head) when the distance rating was below 10 cm, and 1 (outside the head) when
599 the distance rating was above 10 cm. The 10-cm threshold was chosen as an approximation
600 of a head radius and also because it corresponds to a dip in the distribution of the distance
601 ratings (see $\ln(1 + d) = 0.15$ in Suppl. fig. 3). After averaging across repetitions, the
602 binary distance ratings were directly compared to the externalization ratings by computing
603 their correlation (across listeners, conditions and types of sound, $n = 18 * 27 * 3 = 1458$),
604 and analyzing the percentage of the data for which they are 0, 0.5, and 1. To evaluate the
605 sensitivity of this analysis to the choice of distance threshold, it was also performed with
606 two other thresholds set at 20 or 1 cm, this latter arising from the distance task in which
607 listeners were told that 0 m should correspond to a source perceived inside their head.
608
609
610

611 To test for potential differences in behavior between externalization and (continuous)
612 distance ratings in specific conditions, different statistical analyses were used. The (binary)
613 externalization ratings were analyzed using logistic regressions, fitting the data with
614 generalized linear mixed-effects models having the experimental factors and their interactions as
615 fixed effects and the participants as a random effect. The significant effects were then assessed
616 using analyses of deviance and Tukey pairwise comparisons of estimated marginal means (with
617 *car::Anova* and *emmeans*). The (continuous log-transformed) distance ratings (averaged across
618 repetitions) were analyzed using within-subject two-way analyses of variance and Tukey pairwise
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637

638 comparisons (with *aov* and *TukeyHSD*). The hypotheses stated at the end of the Introduction were
639 tested by applying the statistical analyses to three specific subsets of the data depending on the
640 particular experimental factors considered. Because the three analyses were conducted on
641
642 partially overlapping data, the alpha value used to evaluate significance in these analyses was
643
644 Bonferroni corrected to 0.0166.
645
646

647 Hierarchical clustering was performed for each session in order to assess the homogeneity
648
649 of the listeners ratings (Gordon, 1999; Prud'homme and Lavandier, 2020). This was done using
650
651 a matrix of dissimilarities across participants, with dissimilarity between two participants
652
653 being calculated as 1 minus the Pearson's correlation of their distance/externalization
654
655 estimates across all conditions in the session (after averaging across repetitions). The par-
656
657 ticipants' ratings were found to be homogeneous for the two sessions.
658
659
660

661 RESULTS

662 Overall comparisons of distance and externalization estimates

663
664
665
666
667
668 Figure 1 presents the scatter plot of the raw externalization and distance ratings. The
669
670 logistic regression controlling for the effects of sound type and processing conditions indicated a
671
672 significant effect of perceived distance on externalization ($\chi^2(1) = 80.9, p < 0.001$). The point-
673
674 biserial correlation between externalization and distance ratings is 0.36. When averaging across
675
676 repetitions, among the 112 distance ratings at 0 m, 86.6% have an externalization rating at 0,
677
678 9.8% at 0.5 and 3.6% at 1. Among the 567 sources rated as internalized for the two repetitions
679
680 (externalization ratings at 0), 17.1% have a distance at 0 m while the remaining 82.9% have a
681
682 distance different from 0 (up to 125 m): 72.8% above 10 cm, 70.9% above 20 cm, 59.8% above
683
684 50 cm, 48.3% above 1 m.
685

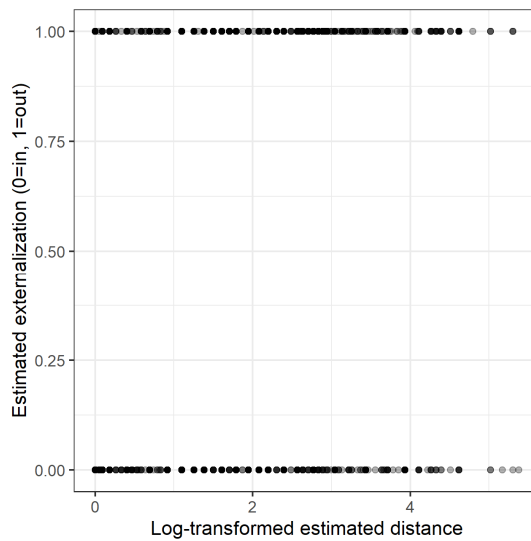
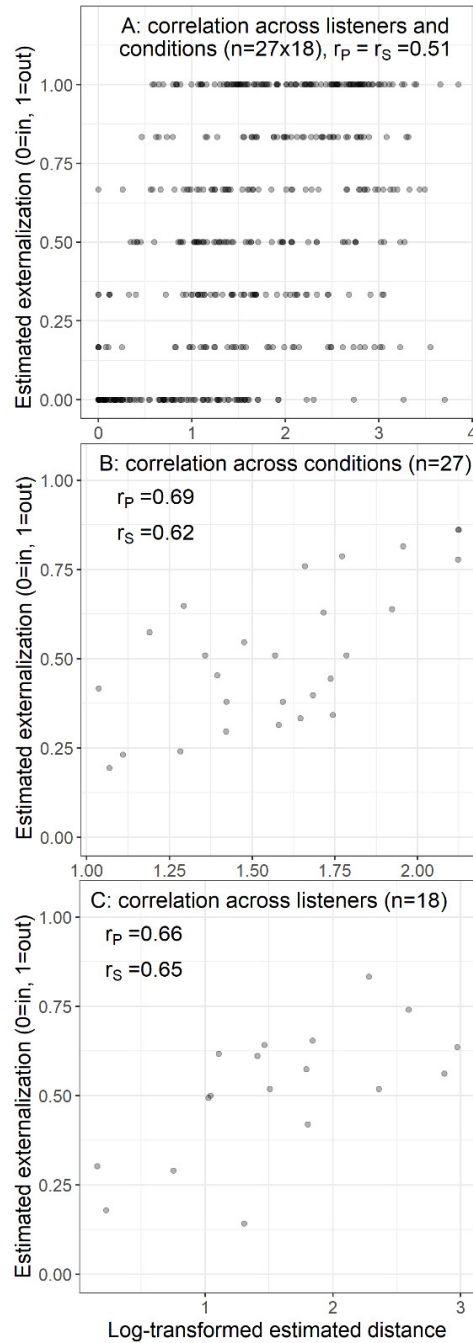


FIG. 1. Scatter plots of the raw externalization and log-transformed distance ratings. The data symbols are partly transparent, so that they appear darker when data overlap.

Figure 2 presents scatter plots of the data averaged across repetitions and sound types, displaying the externalization and distance ratings across processing conditions and listeners (panel A), or averaged across listeners (panel B) or processing conditions (panel C).

Distance and externalization ratings are significantly correlated in these three cases. The Pearson and Spearman correlation coefficients are $r_P = r_S = 0.51$ across processing conditions and listeners, $r_P = 0.69$ and $r_S = 0.62$ across processing conditions (averaging across listeners), $r_P = 0.66$ and $r_S = 0.65$ across listeners (averaging across processing conditions).



707

708

709

710

711

712

713

714

FIG. 2. Scatter plots of the mean (across repetitions and sound types) externalization and log-transformed distance ratings across the 27 processing conditions and 18 listeners (panel A), or averaged across listeners (panel B) or processing conditions (panel C). The data symbols are partly transparent, so that they appear darker when data overlap increases, highlighting the data distribution. The corresponding Pearson and Spearman correlation coefficients (r_P and r_S) are presented in each panel. All correlations are significant.

715 To test whether a source was rated as externalized as soon as its distance rating exceeded
716 10 cm, the binary distance ratings were compared to the externalization ratings (after aver-
717 aging across repetitions but not sound types). Their correlation across listeners, conditions
718 and types of sound is significant with $r_P = 0.44$ and $r_S = 0.46$. There is a match between
719 binary distance and externalization for 51.3% of responses: for 40.3% of responses the source
720 was rated as externalized and with a “distance outside the head”; for 10% of responses the
721 source was rated as internalized and with a “distance inside the head”; for 1% of responses
722 the ratings are both 0.5. There is a first type of mismatch between binary distance and ex-
723 ternalization for 22.1% of responses: 0.4% of the sources were externalized with a distance
724 inside the head, while 21.7% were internalized with a distance outside the head. There is a
725 second type of mismatch for 26.6% of responses: one rating is at 0.5 while the other is at
726 0 or 1 (mostly the externalization at 0.5 and binary distance at 1, in 17.8% of responses,
727 or the externalization at 0 and binary distance at 0.5, in 7.1% of responses), indicating a
728 mismatch at one of the two sound presentations.
729
730
731
732
733
734
735
736
737
738
739
740
741
742

743 These analyses detailed by processing conditions are presented in Suppl. table 3. The
744 mismatch where an internalized source is rated with a distance outside the head is observed
745 more in the diotic conditions (*Dio*, *REF*), with proportions ranging from 25.9% to 37%
746 of the responses, and also for the frontal sources (*HRTF_0*, *BRIR_0*), with proportions
747 between 16.7% and 38.9% of the responses. The proportions and correlations remain broadly
748 unchanged with a distance threshold set at 1 or 20 cm (Suppl. table 4; $r_P = 0.40$ and
749 $r_S = 0.43$ for 1 cm, $r_P = 0.45$ and $r_S = 0.46$ for 20 cm).
750
751
752
753
754
755
756
757
758
759

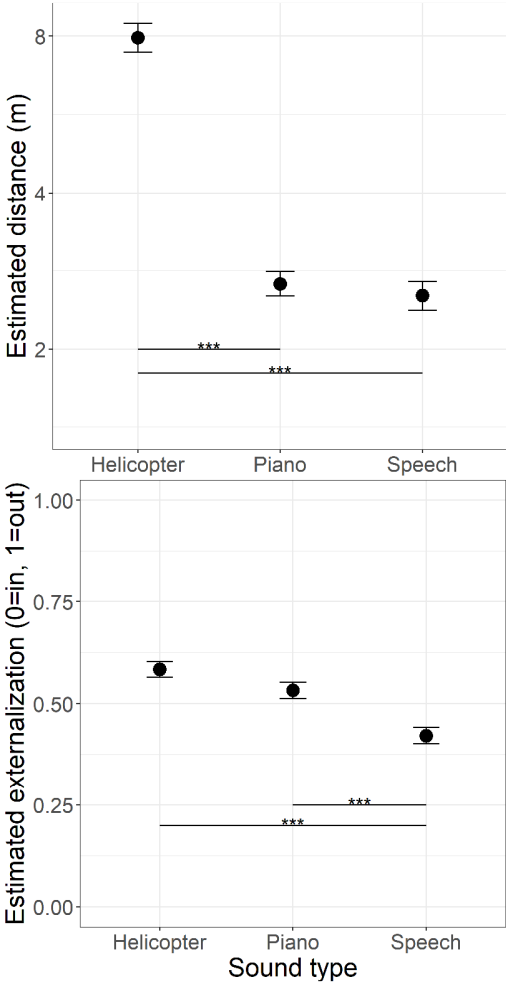
760
761 **Specific comparisons of distance and externalization estimates**
762

763
764 Distance and externalization were also compared by investigating potential differences in
765 behavior in specific conditions. Our hypotheses concerned the effects of (1) sound type, (2)
766 simulated azimuth, (3) simulated distance and reverberation, (4) listening mode and diotic
767 reverberation, (5) simulated distance and reverberation, (6) listening mode and diotic
768 reverberation, (7) simulated distance and reverberation, (8) listening mode and diotic
769 reverberation.
770

771
772
773 **Sound type**
774

775
776 To investigate the effect of sound type on distance and externalization estimates, a main
777 analysis was applied to the whole dataset considering the factors sound type (3 levels)
778 and processing condition (27 levels). It revealed significant main effects of sound type
779 (distance: $F(2,1377) = 129.1, p < 0.001$; externalization: $\chi^2(2) = 9.7, p < 0.01$) and processing
780 condition (distance: $F(26,1377) = 5.1, p < 0.001$; externalization: $\chi^2(26) = 130.2, p < 0.001$).
781 To highlight the effect of sound type, Figure 3 presents the externalization (bottom panel)
782 and distance (top panel) estimates averaged across processing types and listeners. Pairwise
783 comparisons on externalization indicated that the speech was perceived less externalized
784 than both the piano and helicopter. These comparisons on distance estimates indicated
785 that the helicopter was perceived at a significantly larger distance than both the piano
786 and speech. Distance ratings (averaged across repetitions) varied between 0 and 200 m; 90
787 ratings (6.1% of ratings) were at 35 m or more, 8 were speech (from 2 listeners), 1 was piano
788
789
790
791
792
793
794
795
796
797
798
799

800



801
802

803
804

805

806

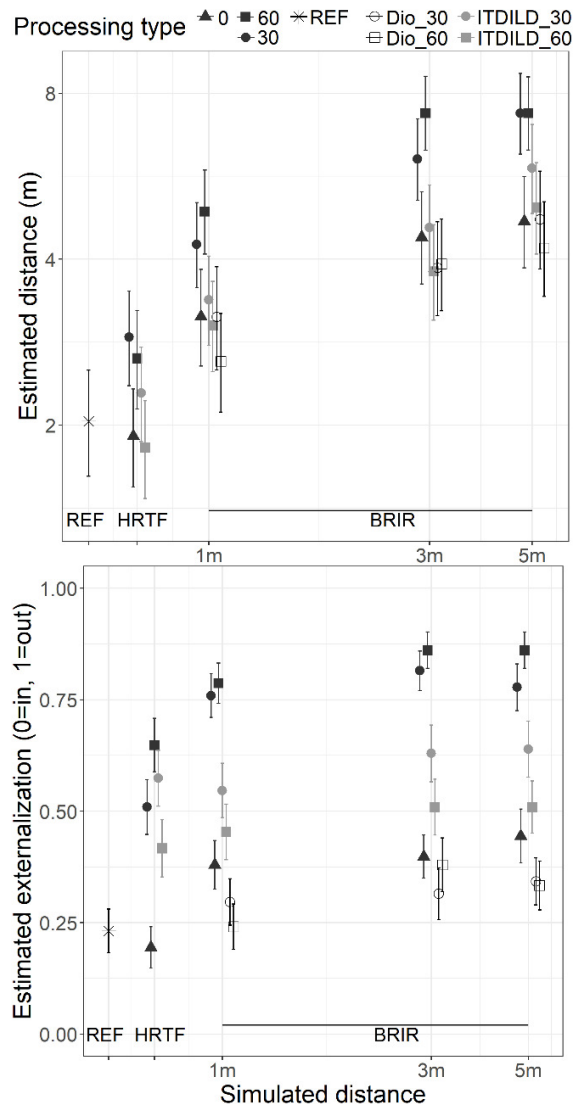
FIG. 3. Mean distance (top) and externalization (bottom) estimates with standard errors across listeners and processing types plotted as a function of sound type. The significant differences are highlighted with *** ($p < 0.0001$).

807
808
809
810
811
812

(from a listener who judged 7 speech at 35 m or more) and 81 were helicopters (from 6 listeners); 29 ratings (2% of ratings) were at 70 m or more, they were all helicopters (from 4 listeners).

Simulated azimuth, distance and reverberation

Figure 4 presents the distance (top panel) and externalization (bottom panel) estimates averaged across sound types and listeners. To investigate the effects of simulated azimuth and distance/reverberation, sub-analysis 1 considered a subset of 12 processing conditions (black filled symbols): the *HRTF* conditions at the 3 simulated azimuths and the *BRIR* conditions at the 3 simulated azimuths and distances. They involved four reverberation levels (one anechoic and one for each simulated distance). Sub-analysis 1 thus tested for three experimental factors: source azimuth (3 levels), reverberation level (4 levels), and sound type (3 levels). On distance estimates, it revealed significant main effects of source azimuth ($F(2,612) = 7.8, p < 0.001$), reverberation level ($F(3,612) = 19.5, p < 0.001$), and sound type ($F(2,612) = 62, p < 0.001$). Pairwise comparisons indicated that the frontal sources (black triangles, top panel of Fig. 4) were judged significantly closer than the two lateral sources (black squares and circles). All reverberation levels led to significant differences in distance estimates, apart from the BRIRs at 3 and 5 m. Perceived distance increased when going from anechoic (*HRTF*) to reverberant (*BRIR*), and further increased when simulating the source further away in the room at 3 m compared to 1 m. The analysis on externalization indicated significant main effects of source azimuth ($\chi^2(2) = 12.4, p < 0.01$) and sound type ($\chi^2(2) = 9.3, p < 0.01$). Pairwise comparisons revealed that the frontal sources (black triangles, bottom panel of Fig. 4) were judged significantly less externalized than the two lateral sources (black squares and circles). The effects of sound type were identical to those already described (Fig. 3).



858
859
860

861 FIG. 4. Mean distance (top) and externalization (bottom) estimates with standard errors across
 862 listeners and sound types in the 27 tested conditions grouped by processing type: the categories
 863 *REF* (original unprocessed diotic signal), *HRTF* (anechoic signals), and *BRIR* (reverberated sig-
 864 nals convolved with a *BRIR* at one of the 3 simulated distances: 1, 3 or 5 m) are displayed on the
 865 x-axis, while the symbols code for the categories *0/30/60* (signals convolved with the *BRIR/HRTF*
 866 at 0° , -30° , and 60° , respectively), *Dio_30/60* (diotic versions of the *BRIR* signals at -30° and 60°), and
 867 *ITDILD_30/60* (broadband ITD and ILD added to the diotic versions of the *BRIR/HRTF* signals at
 868 -30° and 60°). A small horizontal offset has been added to the data to reduce symbol overlap. For
 869 figure readability, significant differences are reported only in the text.

Diotic reverberation and listening mode

To investigate the effects of listening mode and diotic reverberation, sub-analysis 2 considered a subset of 18 processing conditions (all the BRIR data in Fig. 4 apart from the black triangles of the frontal conditions): the BRIR conditions at the 3 simulated distances (1, 3 and 5 m), for the lateral sources, in 3 listening modes (signals convolved with the BRIRs 30/60, diotic versions of these signals *Dio_30/60*, and signals with broadband ITD and ILD added to the diotic versions of the BRIR signals *ITDILD_30/60*). Sub-analysis 2 thus tested for four experimental factors: listening mode (3 levels), as well as sound type (3 levels), source azimuth (2 levels), and reverberation level/simulated distance (3 levels). On distance estimates, it revealed significant main effects of listening mode ($F(2,918) = 13.8, p < 0.001$), sound type ($F(2,918) = 71.3, p < 0.001$) and simulated distance ($F(2,918) = 11.6, p < 0.001$). Pairwise comparisons indicated that the BRIR sources (black filled circles and squares in the top panel of Fig. 4) were perceived further away than their *Dio* and *ITDILD* versions (open and gray symbols). The sources simulated at 1 m were perceived significantly closer than the sources simulated at 3 and 5 m. The effect of sound type was identical to the one already described (Fig. 3). The analysis on externalization indicated a significant main effect of listening mode ($\chi^2(2) = 14.1, p < 0.001$). Pairwise comparisons revealed that the 3 listening modes led to externalization estimates that were all significantly different from each other, in particular the *ITDILD* signals (gray filled symbols in the bottom panel of Fig. 4) led to more externalized sources than the *Dio* signals (open symbols).

914 The original diotic anechoic signals (*REF*) were used as a reference for sources that
915 should be very internalized. Figure 4 confirms that this condition provided for the lowest
916 externalization ratings, along with the frontal anechoic source (*HRTF_0*, which was diotic
917 and anechoic like *REF*).
918
919
920
921
922
923
924
925
926

927 **DISCUSSION**

928
929

930 The present study aimed at comparing evaluations of perceived distance and externalization
931 by the same listeners for the same source simulations. A logistic regression indicated that
932 externalization ratings could be partly predicted from the distance ratings. The results confirmed
933 that ratings averaged across repetitions and sound types were significantly correlated, r_P and r_S
934 varying between 0.51 and 0.69 when considering comparisons across listeners and/or conditions.
935 They also indicated that a source was not rated as externalized as soon as and only if its
936 perceived distance exceeded a threshold: using a 10-cm threshold, only 51.3% of responses are
937 consistent with this hypothesis. The ratings were also compared in specific subsets of conditions
938 chosen to test our hypotheses concerning potential differences in behavior between distance and
939 externalization. These conditions were not used to highlight any new effect. Most of
940 the effects previously highlighted in the literature were replicated here. As discussed below,
941 while the present study cannot conclude concerning a potential difference for the effect of
942 source azimuth on distance and externalization, it highlights differences for the effects of
943 sound type, simulated distance, and diotic reverberation.
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958

Overall comparisons of distance and externalization

The 0.36 point-biserial correlation between the raw externalization and distance ratings indicates that the variance in the distance ratings can account for 13% of the variance in the externalization ratings. When averaging ratings across repetitions and sound types, Figure 2 confirms that distance and externalization are not independent percepts: the corresponding ratings are correlated across processing conditions and listeners (panel A), with $r_P = r_S = 0.51$. These correlations coefficients indicate that the variance in distance ratings/rankings explains 26% of the variance in externalization ratings/rankings (and reciprocally). The ratings are also correlated when averaged across listeners (panel B, $r_P = 0.69$ and $r_S = 0.62$), so that conditions in which sources were perceived as far/near tended to be conditions in which the sources were perceived as externalized/internalized, with the variance in distance ratings/rankings explaining 47.6% and 38.4% of the variance in the externalization ratings/rankings, respectively. The ratings are correlated when averaged across conditions (panel C, $r_P = 0.66$ and $r_S = 0.65$), so that listeners who judged a source as far/near also tended to judge it as externalized/internalized, with the variance in distance ratings/rankings explaining 43.6% and 42.2% of the variance in externalization ratings/rankings, respectively.

One could argue that the limited amounts of explained variance could result from the hypothesis that externalization is expected to be constant when perceived distance increases

997 above the head radius, the source being perceived inside the head when perceived distance
998 is below this threshold. The binary distance ratings (0 for $d \leq 10$ cm; 1 otherwise) are
999 significantly correlated with the externalization ratings ($r_P = 0.44$ and $r_S = 0.46$) and 51.3%
1000 of responses support this hypothesis. However, a mismatch is observed across repetitions in
1001 22.1% of responses, mostly associated with sources rated as internalized but with a distance
1002 outside the head (21.7% of responses), in particular in the diotic conditions and for the
1003 frontal sources. For the remaining 26.6% of responses, one of the two sound presentations
1004 also led to a mismatch. Overall, these results indicate that a source was not rated as
1005 externalized as soon as and only if its perceived distance exceeded a threshold. This was
1006 further confirmed by analyses setting the threshold just above 0 or at 20 cm.
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016

1017 The reference sources for internalization (*REF*) were on average rated at 2 m by the
1018 listeners (despite being instructed to rate at 0 m a source perceived in their head), while
1019 still providing among the lowest externalization ratings. This is surprising, even when taking
1020 into account that distance estimations are less reliable when measured in a listening booth
1021 ([Cubick et al., 2015](#)). Three main reasons could explain this result. First, sound level was
1022 the only distance cue available for these stimuli (while DRR was another cue for
1023 the reverberated sounds) and the listeners were faced with the difficult task of making
1024 absolute judgments for sounds equalized in level, a situation which would not be expected in
1025 natural environments. Second, spatial perception is dependent on the stimulus context, in
1026 particular the perceived location of the preceding stimuli, as highlighted for noise bursts
1027 ([Andrejková et al., 2023](#); [Carlile et al., 2001](#)), band-limited noises ([Laback, 2023](#)) and tones
1028 ([Lingner et al., 2018](#)). Here, the distance estimate in a given condition was influenced by the
1029 other distances simulated in the experiment. The non-spatialized sources *REF* would have
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042

1043 certainly been rated at much shorter distances if the other sources had been simulated only at
1044 distances between 15 and 100 cm, instead of 1 to 5 m. Third, the distance ratings have most
1045 probably been influenced also by source level expectations from the listeners (see the discussion
1046 below on the effect of sound type), such that the helicopter was perceived much farther than
1047 the speech and piano. The distance ratings for the helicopter were large in all conditions, in
1048 particular also in the *REF* condition, contributing to the average rating of this condition being at
1049 2 m. Although this 2-m value is surprising, it is reassuring to note that the REF sources still led to
1050 the lowest distance ratings in the experiment. This confirms that the interpretations of absolute
1051 distance judgments need to account for the multiple possible influences of context (task, stimuli,
1052 instructions), as suggested by [Lingner et al. \(2018\)](#) for absolute spatial perception in general.
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063

1064 **Specific comparisons of distance and externalization**

1065
1066

1067 ***Sound type***

1068
1069

1070 Similar trends across processing conditions were obtained for the three types of sound.
1071 However, the main effects of sound type on distance and externalization ratings were dif-
1072 ferent: while the helicopter was evaluated at further distances than the piano and speech,
1073 speech was judged as less externalized than the helicopter and piano. Another way of for-
1074 mulating this differential role of sound type on distance and externalization is that, while
1075 the relationship between the helicopter and speech was identical in the two measures (the
1076 helicopter was rated as more distant and more externalized compared to the speech), the
1077 piano was found more externalized but not more distant than the speech, and found less
1078 distant but not less externalized than the helicopter.
1079
1080
1081
1082
1083
1084
1085
1086
1087

1088
1089 While the stimuli equalized in level were not equalized in loudness (across sound types
1090 with different spectra), the very large distance estimates obtained for the helicopter indi-
1091 cate that factors other than loudness were probably involved, including cognitive factors.
1092
1093 [Philbeck and Mershon \(2002\)](#) suggested that, in an experimental context, distance percep-
1094 tion can be affected by past experience independently from the comparison with previous
1095 stimuli in the experiment. A helicopter is generally not a source experienced at short dis-
1096 tances by most listeners, but more critically they would probably expect it to be very loud,
1097 and thus could consider that it should be far away to produce the same sound level as speech
1098 at the ears. Equalizing the ear level removed a strong distance cue for each source type, but
1099 it potentially reinforced or introduced an across-type cue due to different level expectations
1100 from the listeners depending on the source. The present study points toward level expect-
1101 ations having a strong effect on absolute distance ratings, generalizing the conclusions of
1102 [Brungart and Scott \(2001\)](#) to stimuli other than (whispered, conversational and shouted)
1103 speech.
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115

1116
1117 It is important to note that the present study used convolutions with BRIRs and HRTFs
1118 to simulate a loudspeaker (the one used in the BRIR/HRTF measurements) reproducing
1119 different types of sound, in agreement with the instructions given to the listeners who were
1120 asked to imagine such a quite common situation in real life (e.g. when watching a movie).
1121
1122 The observed effect of level expectations could indicate that listeners might have imagined
1123 real sources rather than a loudspeaker. The convolutions used would not be appropriate to
1124 simulate such a situation: at distances of 1 to 5 m, a helicopter and even a piano cannot be
1125 approximated by a point source. This approximation would only be realistic if these sources
1126
1127
1128
1129
1130
1131
1132

1133 were simulated far away. This could have further contributed to the large distance estimates
1134 obtained for the helicopter.
1135

1136 The speech was found less externalized than the helicopter and piano, consistent with
1137 the results of [Leclère et al. \(2019\)](#) who found speech slightly less externalized than music
1138 in part of their experiments. Speech might be more internalized because internal speech is
1139 not uncommon as people can hear their own voice (at least partly) inside their head when
1140 talking and are also quite familiar with phone conversations with earphones that generally
1141 also produce internal speech. However, no strong effect of sound type on externalization
1142 was found here nor in the study of [Leclère et al. \(2019\)](#), contrary to the strong effect found
1143 here for distance.
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155

1156 ***Frontal vs. lateralized sources***

1157
1158

1159 Externalization ratings were expected to be higher for lateral sources compared to frontal
1160 sources ([Brimijoin et al., 2013](#); [Kates et al., 2018](#); [Leclère et al., 2019](#)). This was true here.
1161
1162 The lateral sources were more externalized than the frontal sources, both in
1163 anechoic and reverberant conditions. This effect can be further understood by considering
1164 the *ITDILD* sources. They tested whether lateral sources could be more externalized only
1165 because they are simulated on the side, even in the absence of the time-varying frequency-
1166 dependent binaural cues associated with reverberation that have been shown to enhance
1167 externalization ([Catic et al., 2013](#); [Leclère et al., 2019](#); [Li et al., 2019](#)). These cues explain
1168 here that the *ITDILD* sources were less externalized than the sources with full binaural cues.
1169
1170
1171
1172
1173
1174
1175
1176
1177

1178 However, the fact that the *ITDILD* sources were more externalized than the corresponding
1179 diotic sources indicates that there is something in addition to the time/frequency-dependent
1180 binaural cues that enhances externalization for lateral sources. It gives support to the
1181 existence of a bias in which listeners might not rate a source as internalized as soon as it is
1182 lateralized, even if it might still be inside the head, because it tends to be perceived on the
1183 side of the head, closer to the surface of the skull (Best *et al.*, 2020).
1184
1185
1186
1187
1188
1189

1190
1191 Contrary to externalization, perceived distance was not expected to change much with
1192 azimuth for simulated distances above 1 m. Zahorik (2002) did not find a direct influence
1193 of source direction on perceived distance while comparing the parameters of power func-
1194 tions fitted to distance estimates. Here, the direct comparison of such estimates showed
1195 that frontal sources were perceived closer than lateral sources. The discrepancy might be
1196 explained by the fact that in the present study stimuli were equalized in sound level, while
1197 level was kept as a distance cue in Zahorik's study, who also showed that the relative weights
1198 given by listeners to the level and reverberation cues depended on source direction. The
1199 present study indicates that, when overall level is kept constant, distance perception can
1200 depend on source direction even for distances above 1 m. The fact that the diotic versions
1201 of the lateral sources were perceived at the same shorter distances as the frontal sources
1202 indicates that the difference in perceived distance between lateral and frontal sources proba-
1203 bly relies on additional binaural information available for the lateral sources (this is further
1204 discussed below).
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221

1222 Even if Figure 4 indicates that the effect of source azimuth could be larger for external-
1223 ization than for distance, the magnitude of the differences in ratings between frontal and
1224 lateral sources is difficult to compare on different scales. Moreover, the statistical analyses
1225
1226 point to a significant effect of source azimuth for both externalization and distance. Thus
1227
1228 it cannot be concluded here that this effect was different for distance and externalization.
1229
1230
1231
1232
1233

1234 ***Simulated distance and reverberation***

1235

1236 Reverberation was expected to enhance externalization (Catic *et al.*, 2013; Leclère *et al.*,
1237 2019; Li *et al.*, 2019). Figure 4 shows that sources simulated in the room tended to be
1238
1239 perceived as more externalized than the anechoic sources, but this trend was not statistically
1240
1241 significant (with a p-value of 0.0447 in sub-analysis 1, the effect was above the Bonferroni-
1242
1243 corrected significance level at 0.0166). This is probably due to a lack of statistical power.
1244
1245 More listeners or repetitions would have been necessary to confirm this effect highlighted in
1246
1247 the literature.
1248
1249

1250 Perceived distance was expected to increase with simulated distance in the room (Prud'homme
1251
1252 and Lavandier, 2020; Zahorik, 2002) and this was observed when simulated distance increased
1253
1254 from 1 to 3 m. Increasing this distance to 5 m did not further increase perceived distance,
1255
1256 maybe because of the existence of an auditory “horizon” observed in several previous studies
1257
1258 (Akeroyd *et al.*, 2007; Bronkhorst and Houtgast, 1999; Mershon and Bowers, 1979; Zahorik,
1259
1260 2002). Contrary to perceived distance, externalization remained fixed while varying the
1261
1262 source distance in the room.
1263
1264
1265
1266

1267
1268 ***Diotic reverberation***
1269

1270 Perceived distance was still expected to vary with simulated distance in diotic listening
1271 (Bidart and Lavandier, 2016; Prud'homme and Lavandier, 2020) and this was verified: the
1272 effect of simulated distance was significant and did not interact with the listening mode
1273 (BRIR vs. Diotic vs. ITDILD). Significant differences in perceived distance were obtained
1274 in the diotic conditions that were among the most internalized conditions. For the lateral
1275 sources, the diotic stimuli led to shorter perceived distances than the *BRIR* stimuli with
1276 full binaural information. This contrasts with previous results obtained for frontal sources,
1277 in which diotic and dichotic listening led to very similar distance estimates (Bidart and
1278 Lavandier, 2016; Prud'homme and Lavandier, 2020). Diotic listening was not tested here
1279 for the frontal sources to limit the number of tested conditions. Nevertheless, the frontal
1280 sources were perceived at about the same distances as the diotic and *ITDILD* versions
1281 of the lateral sources, closer than the lateral sources with full binaural information. This
1282 suggests that the difference in perceived distance between lateral and frontal sources could
1283 be due to the time-varying frequency-dependent binaural differences available for the lateral
1284 sources simulated in the room. The long-term broadband approximation of these binaural
1285 differences (in the *ITDILD* stimuli) and the binaural differences created by reverberation
1286 for the frontal sources do not seem sufficient to increase perceived distance, consistent with
1287 the results of the aforementioned studies in which dichotic and diotic listening led to similar
1288 distance estimates for frontal sources.
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308

1309 Contrary to its effect on distance, diotic reverberation was not expected to improve ex-
1310
1311

ternalization (Catic *et al.*, 2015; Leclère *et al.*, 2019). Our results show a strong reduction in externalization when reverberation was diotic (*Dio_30/Dio_60*) compared to dichotic (30/60). The diotic conditions were the least externalized conditions with the anechoic frontal sources and the *REF* conditions. Adding the broadband ITD and ILD to the diotic versions of the *BRIR* and *HRTF* stimuli increased externalization (as discussed above), while this was not observed for distance.

Limitations

Many studies on externalization have based their evaluation on an underlying measure of distance (for a review see Best *et al.*, 2020). The most commonly used externalization scale is going from the center of the head towards a fixed point in the environment, usually a (silent) loudspeaker used as a reference, often using adjectives such as “close/near” and “far” (Catic *et al.*, 2013; Gil-Carvajal *et al.*, 2016; Hartmann and Wittenberg, 1996; Hassager *et al.*, 2016; Kim and Choi, 2005). Using such a scale is implicitly assuming that distance and externalization are aspects of a single perceptual continuum. Here, such a distance-related measure of externalization was replaced by a binary question, like previously used by Brimijoin *et al.* (2013). To create a continuous scale for externalization without using distance references, a question concerning the confidence in each externalization rating was included. It varied continuously between 0 (not confident at all) and 1 (very confident). Through multiplication of the externalization rating (changed to -1 for internalized, keeping 1 for externalized) with the confidence rating, one can create a continuous scale between -1 for extremely confident that the sound originated internally, to 1 for extremely confident that the sound was externalized. However, in the present study, the confidence ratings were generally high

1358 and these continuous ratings showed highly similar results as the simpler binary ratings.

1361 Stimuli were equalized in overall level in the present study, thus removing a strong dis-
1362 tance cue for each sound type (but not the across-type cue associated with level expectations
1363 discussed above, nor the variations in direct and reverberated sound levels that might have
1364 been used as distance cues). If externalization were independent of level, then, because
1365 distance is strongly dependent on level, when natural level variations are present one would
1366 expect externalization and distance ratings to behave even more differently than what has
1367 already been shown here. However, investigating the potential effect of level on external-
1368 ization might not prove as trivial as one might think, [Hartmann and Wittenberg \(1996\)](#)
1369 informally reporting sounds to “jump around” when testing roved levels.

1370 The fact that non-individualized HRTFs/BRIRs were used here for all participants, regard-
1371 less of their head size compared to the manikins used for the measurements, could have impai-
1372 red their spatial perception compared to what they would have experienced listening with their
1373 own ears. However, as pointed out in the Introduction, based on previous studies, no such im-
1374 pairment was expected for distance perception ([Zahorik, 2002b](#); [Prudhomme and Lavandier,](#)
1375 [2020](#)) and, if any, the impairment was probably limited for externalization ([Begault et al., 2001](#);
1376 [Cubick et al., 2015](#); [Kates et al., 2018](#); [Leclère et al., 2019](#)). In particular, [Leclère et al. \(2019\)](#)
1377 showed that individualization of their stimuli had a much smaller effect on externalization
1378 ratings than the effects of reverberation and source azimuth (front vs. side), also smaller than
1379 the already small effect of sound type. The effect of source azimuth on externalization was large
1380 in the present study, but the effect of sound type was small and the effect of reverberation did
1381 not reach significance. Thus, if tested, the effect of individualization would have probably been
1382 very limited. Using non-individualized stimuli could still have slightly impaired externalization

1406 for the less externalized frontal sources (Werner et al., 2016), in particular in the anechoic
1407 condition. This does not jeopardize the study conclusions, which might even generalize to
1408 individualized stimuli, but this remains to be tested.
1409
1410
1411

1412 One of the limitations of the present study is the programming error that affected the
1413 HRTF-based stimuli. The corresponding data had to be mirror-imaged assuming that the
1414 distance and externalization tasks were not affected by any left-right asymmetry. The anechoic
1415 non-individualized HRTFs used here are left/right symmetric, so any asymmetry
1416 affecting the results would have been perceptual. Arend et al. (2021) measured the
1417 perceived distance of nearby sources simulated between 0.25 m and 1.50 m, in anechoic
1418 conditions, for three source azimuths: 30°, 150°, and -90°. A left/right asymmetry cannot be
1419 evaluated directly because the sources on each side were tested at different azimuths, but already
1420 the effect of azimuth was significant essentially at very short distances below 0.5 m, or for
1421 conditions without level equalization in which the effect was interpreted as resulting from
1422 azimuth-dependent loudness differences in the stimuli. Parseihian et al. (2014) also measured
1423 the perceived distance of nearby sources, between 0.72 m and 1.08 m, but for a configuration of
1424 sources placed on a tabletop surface. Sources were tested every 30° between +/- 60° and also at
1425 +/- 90° and +/- 120°, in a low reverberant space, with stimuli equalized in loudness. Listeners had to
1426 place their hand at the perceived position of the source and results were analyzed in terms of
1427 errors in perceived localization. The error in distance was significantly smaller for the lateral
1428 sources (60°/90°) compared to the other positions, but no significant differences between +/- 30°
1429 or +/- 60° was found, indicating no sign of a left/right asymmetry despite the (non-anechoic)
1430 room potentially introducing differences in the stimuli for the left/right sources. In the study of
1431 Best and Roverud (2024), listeners wearing different types of hearing aids were seated in a large
1432 sound-treated booth and presented with sound from one of seven loudspeakers at 0°, +/- 15°, +/-
1433
1434

1455 30°, +/- 90°, at a fixed distance. Perceived distance was evaluated on a scale ranging from 10 (at
1456 the loudspeaker ring) to 0 (in the head) to -10 (the furthest behind the listener). For normal-
1457 hearing listeners, no significant effect of source azimuth was found. The authors also used the
1458 distance ratings to compute binary internalization ratings and did not find a significant effect of
1459 source azimuth. There was no evidence for a left/right asymmetry in perceived distance or
1460 externalization, despite potential differences in the stimuli due to non-symmetric reverberation.
1461 [Begault et al. \(2001\)](#) measured externalization for sources simulated in a room at 0°, +/- 45°, +/-
1462 135°, and 180°. Again, while reverberation could have introduced differences in the stimuli for the
1463 left/right sources, they did not find a significant effect of source azimuth, in particular no sign of
1464 a left/right asymmetry. In their experiment 1, [Brimijoin et al. \(2013\)](#) measured externalization
1465 while reproducing sound over loudspeakers placed at a fixed distance, every 30° symmetrically
1466 around the listener, in a hemi-anechoic room. The signal was either fixed, reproduced on a single
1467 loudspeaker, or it moved across the loudspeakers to follow the listener head movements
1468 (simulating headphone listening). The pattern of externalization ratings appears very
1469 symmetrical (see their Fig. 3), but the ratings of symmetric conditions are not identical, in
1470 particular for the signals following head movements (though error bars almost always overlap).
1471 The effect of source azimuth was significant but no post-hoc analyses were provided to test for an
1472 effect of source laterality. A few other studies measured externalization for sources simulated on
1473 both sides of the listeners, but the data were averaged across sides without mentioning/testing
1474 for any left/right difference ([Wenzel, 1995](#); experiment 2 of [Brimijoin et al., 2013](#); [Hendrickx et al., 2017](#); [Heine et al., 2021](#)). Even if these studies point towards the absence of a left/right
1475 asymmetry in perceived distance and externalization, none was specifically designed to
1476 investigate such asymmetry. Previous studies have highlighted a left/right asymmetry in
1477 perceived azimuth ([Abel et al., 1999, 2000](#); [Burke et al., 1994](#); [Savel, 2009](#)) and elevation ([Butler, 1994](#)), as well as in the build-up of the echo suppression ([Grantham, 1996](#)). Interestingly, [Abel et](#)

1504 *al.* (1999, 2000) and *Burke et al.* (1994) associated the asymmetry in azimuth perception with
1505 more front-back confusions for the right sources, and thus to a difference in the treatment of
1506 spectral cues, as it would be the case for the asymmetry in perceived elevation (*Butler, 1994*).
1507 Therefore, it could be interesting to investigate the potential existence of a left/right asymmetry
1508 for perceived distance and externalization, particularly in conditions in which they rely on
1509 spectral cues (*Zahorik et al., 2005; Kolarik et al., 2016; Best et al., 2020*), while controlling for
1510 any confounding effect that could be associated with non-symmetric reverberation.
1511
1512
1513
1514
1515
1516
1517
1518

1519 The results of the present study are probably influenced by different contextual effects
1520 associated with the tasks and stimuli. The distance and externalization tasks were done in
1521 different experimental blocks. Listeners could report the source to be at a distance in one
1522 block, and then in their head in the second block, without being challenged about that.
1523 The tasks were blocked so that their results could be compared to those of experiments that
1524 measured only distance or only externalization, but the results might have been different if
1525 the tasks were done one just after the other for each sound presentation. The contextual
1526 effects also concern the different type of stimuli that were mixed in the same experimental
1527 block. It was already mentioned that the range of simulated distances influences all distance
1528 estimates, even those of non-spatialized sounds (*REF*). There could be also contextual
1529 effects associated with presenting a reverberated sound just after an anechoic one. The
1530 effects of sound type and level expectations could have been reduced if the sound types were
1531 tested in different experimental blocks. These contextual effects are impossible to assess
1532 here, but they likely influenced the data. This question requires further investigation, e.g.
1533 by designing an experiment in which the types of sound are blocked but the tasks are not.
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549

1550
1551 The results of the present study are probably also influenced by the instructions given
1552 to the listeners. Visual information may have affected judgments in both tasks. Even if the
1553 listeners were asked to keep their eyes closed while listening, they were not monitored to
1554 check that they obeyed this instruction. This could have impaired their ability to imagine a
1555 loudspeaker at a certain distance producing sounds. If they had been explicitly told that a
1556 source perceived at a distance had to be outside their head, and that they should monitor
1557 their externalization responses as such, the data might be different. Listeners might make a
1558 distinction between making judgments in virtual space where they might vary their distance
1559 ratings (e.g. a source is rated farther away when it is more reverberated) even if the source
1560 is internalized (no matter how reverberated it is), and real space where sources have to be
1561 externalized to vary in distance.

1572
1573 In the present study, simulated distance varied from 1 to 5 m. Such distances allowed
1574 to highlight that externalization does not arise simply because the source has a (simulated)
1575 distance from the listener. However, if the aim were to get a better spatial resolution closer
1576 to the head radius, smaller increments of distance below 1 m should be considered. Note
1577 that then distance becomes more depend on binaural information ([Brungart et al., 1999](#)),
1578 like externalization is ([Catic et al., 2013](#); [Leclère et al., 2019](#); [Li et al., 2019](#)), thus the
1579 correlation between distance and externalization might be stronger. This warrants further
1580 investigations.
1581
1582
1583
1584
1585
1586
1587

1588
1589 In the distance task, listeners were told to use a source perceived in their head as the
1590 reference for 0 m, as commonly done in distance studies ([Best et al., 2020](#); [Kolarik et al.,](#)
1591 [2016](#)). An externalization reference was thus used for a distance task. In the same way

1594
1595 than distance labels might not be the most appropriate for an externalization scale, we will
1596
1597 restrain from using an externalization reference to define the origin of a distance scale in the
1598
1599 future, keeping in mind that significant differences in perceived distance have been reported
1600
1601 in internalized conditions, such as the diotic conditions here and in previous studies (Bidart
1602
1603 and Lavandier, 2016; Prud'homme and Lavandier, 2020). Note that listeners might not
1604
1605 need to be provided with a reference to evaluate distance, and that their reference might be
1606
1607 different between virtual and real spaces.

1608
1609 Finally, one criticism that could be made is that participants did not spontaneously evoke
1610
1611 their subjective experience and were guided in the experience they had to report. They might
1612
1613 spontaneously not perceive externalization but, in experimental conditions when asked to
1614
1615 choose between externalized and internalized, they could develop cognitive strategies to
1616
1617 respond in a distributed manner between the two extremes, based on acoustic cues that are
1618
1619 unrelated to the externalization percept. The same could be said about distance.

1620
1621 Despite these limitations, it is reassuring to note that the cues that vary externalization
1622
1623 and distance in our study have already been reported as such in the literature (Best *et al.*,
1624
1625 2020; Kolarik *et al.*, 2016; Zahorik *et al.*, 2005). The present study replicated several effects
1626
1627 highlighted in the literature in studies that measured either distance or externalization.
1628
1629

1630 1631 **CONCLUSION** 1632

1633
1634 Comparing evaluations of distance and externalization for the same stimuli by the same
1635
1636 listeners indicated that distance and externalization ratings were significantly correlated
1637
1638 across conditions and listeners, and when averaged across listeners or conditions. These

1639
1640 correlations indicated that the variations in perceived distance explained between 26% and
1641
1642 47.6% of the variations in externalization (and reciprocally; 13% in the raw data before
1643
1644 averaging across repetitions and sound types). It was also shown that a source was not
1645
1646 rated as externalized as soon as and only if its perceived distance exceeded a threshold,
1647
1648 e.g. the head radius. Important differences between externalization and distance were
1649
1650 highlighted. Diotic reverberation influenced distance but not externalization. Adding
1651
1652 broadband ITD and ILD to diotic sounds increased externalization but not distance. The
1653
1654 effects of sound type were different for distance and externalization: the helicopter was
1655
1656 evaluated as more distant, while speech was judged as less externalized (or the piano was
1657
1658 found more externalized but not more distant than speech, and found less distant but not
1659
1660 less externalized than the helicopter). Distance ratings increased with simulated distance
1661
1662 even for stimuli judged as internalized (diotic sounds, frontal sources). Taken together, these
1663
1664 results indicate that distance and externalization are correlated but might not be aspects of a
1665
1666 single perceptual continuum. As a consequence, future studies evaluating externalization could
1667
1668 consider not to use a scale related to distance.
1669

1670
1671
1672 **ACKNOWLEDGMENTS, AUTHOR DECLARATIONS AND DATA AVAILABIL-**
1673
1674 **ITY STATEMENT**
1675

1676
1677 The study was carried out in accordance with the guidelines of the Declaration of Helsinki.
1678
1679 The authors declare that there is no conflict of interest. They thank Dr. Nibert Kopčo,
1680
1681 Robin Duclermortier and Dr. Virginia Best for advice on a previous version of this paper.
1682
1683 This project has received funding from the Labex CeLyA (Grant No. ANR-10-LABX-0060),
1684

1685 from the PHC Danube program (Grant Nos. 45268RE, APVV DS-FR-19-0025, and WTZ
1686 MULT 07/2020), and from the European Union’s Horizon Europe research and innovation
1687 program (Grant No. 101129903). The data that support the findings of the present study
1688
1689 are available from the corresponding author upon reasonable request.
1690
1691
1692
1693

1694 Abel, S. M., Giguère, C., Consoli, A., and Papsin, B. C. (1999). “Front/back mirror image
1695 reversal errors and left/right asymmetry in sound localization,” *Acta Acustica united with*
1696 *Acustica* **85**(3), 378-386.
1697
1698
1699

1700 Abel, S. M., Giguère, C., Consoli, A., and Papsin, B. C. (2000). “The effect of aging on
1701 horizontal plane sound localization,” *J. Acoust. Soc. Am.* **108**(2), 743-752.
1702
1703
1704

1705 Akeroyd, M. A., Gatehouse, S., and Blaschke, J. (2007). “The detection of differences
1706 in the cues to distance by elderly hearing-impaired listeners,” *J. Acoust. Soc. Am.* **121**,
1707 1077–1089.
1708
1709

1710 Andrejgovà, G., Best, V., and Kopčo, N. (2023). “Time scales of adaptation to context in
1711 horizontal sound localization,” *J. Acoust. Soc. Am.* **154**, 2191–2202.
1712
1713
1714

1715 Arend, J. M., Liesefeld, H. R., and Pörschmann, C. (2021). “On the influence of non-
1716 individual binaural cues and the impact of level normalization on auditory distance esti-
1717 mation of nearby sound sources,” *Acta Acustica* **5**, 10.
1718
1719
1720

1721 Begault, D. R., Wenzel, E. M., and Anderson, M. R. (2001). “Direct comparison of the impact
1722 of head tracking, reverberation, and individualized head-related transfer functions on the
1723 spatial perception of a virtual speech source,” *J. Audio Eng. Soc.* **49**, 904–916.
1724
1725
1726

1727 Best, V., Baumgartner, R., Lavandier, M., Majdak, P., and Kopčo, N. (2020). “Sound
1728 externalization: A review of recent research,” *Trends Hear.* **24**, 2331216520948390.
1729
1730
1731

Best, V., and Roverud, E. (2024). “Externalization of speech when listening with hearing

1732 aids,” *Trends Hear.* **28**, 23312165241229572.

1733
1734 Bidart, A., and Lavandier, M. (2016). “Room-induced cues for the perception of virtual
1735 auditory distance with stimuli equalized in level,” *Acta Acustica united with Acustica*
1736 **102**(1), 159–169.

1740 Blauert, J. (1997). *Spatial Hearing. The Psychophysics of Human Sound Localization* (The
1741 MIT Press, Harvard MA, 2nd edition).

1745 Brimijoin, W. O., Boyd, A. W., and Akeroyd, M. A. (2013). “The contribution of head
1746 movement to the externalization and internalization of sounds,” *PLoS One* **8**(12), e83068.

1748 Bronkhorst, A. W., and Houtgast, T. (1999). “Auditory distance perception in rooms,”
1749 *Nature* **397**, 517–520.

1752 Brungart, D. S., Durlach, N. I., and Rabinowitz, W. M. (1999). “Auditory localization
1753 of nearby sources. II. Localization of a broadband source,” *J. Acoust. Soc. Am.* **106**(4),
1754 1956–1968.

1758 Brungart, D. S., and Scott, K. R. (2001). “The effects of production and presentation level
1759 on the auditory distance perception of speech,” *J. Acoust. Soc. Am.* **110**, 425–440.

1762 Burke, K. A., Letsos, A., and Butler, R. A. (1994). “Asymmetric performances in binaural
1763 localization of sound in space,” *Neuropsychologia* **32**(11), 1409-1417.

1766 Butler, R. A. (1994). “Asymmetric performances in monaural localization of sound in space,”
1767 *Neuropsychologia* **32**(2), 221-229.

1770 Callan, A., Callan, D. E., and Ando, H. (2013). “Neural correlates of sound externalization,”
1771 *NeuroImage* **66**, 22–27.

1774 Carlile, S., Hyams, S., and Delaney, S. (2001). “Systematic distortions of auditory space
1775 perception following prolonged exposure to broadband noise,” *J. Acoust. Soc. Am.* **110**,

1778
1779 416–424.

1780 Catic, J., Santurette, S., Buchholz, J. M., Gran, F., and Dau, T. (2013). “The effect of
1781 interaural-level-difference fluctuations on the externalization of sound,” *J. Acoust. Soc.*
1782 *Am.* **134**, 1232–1241.

1783
1784
1785
1786 Catic, J., Santurette, S., and Dau, T. (2015). “The role of reverberation-related binaural
1787 cues in the externalization of speech,” *J. Acoust. Soc. Am.* **138**(2), 1154–1167.

1788
1789
1790 Cubick, J., Rodriguez, C. S., Song, W., and MacDonald, E. N. (2015). “Comparison of
1791 binaural microphones for externalization of sounds,” in *Proc. Int. Conf. Spat. Audio*.

1792
1793
1794 Durlach, N. I., Rigopulos, A., Pang, X. D., Woods, W. S., Kulkarni, A., Colburn, H. S., and
1795 Wenzel, E. M. (1992). “On the externalization of auditory images,” *Presence Teleoperators*
1796 *Virtual Environ.* **1**, 251–257.

1797
1798
1799
1800 Farina, A. (2000). “Simultaneous measurement of impulse response and distortion with
1801 swept-sine technique,” in *AES 108th Convention*, Preprint 5093 (D-4).

1802
1803
1804 Gardner, B., and Martin, K. (1994). “HRTF measurements of a KEMAR dummy-head
1805 microphone,” *J. Acoust. Soc. Am.* **97**, 3907–3908.

1806
1807
1808 Gil-Carvajal, J. C., Cubick, J., Santurette, S., and Dau, T. (2016). “Spatial hearing with
1809 incongruent visual or auditory room cues,” *Sci. Rep.* **6**, 37342.

1810
1811
1812 Gordon, A. D. (1999). *Classification* (Chapman and Hall/CRC, Boca Raton, USA, 2nd
1813 Edition).

1814
1815
1816 Grantham, D. W. (1996). “Left-right asymmetry in the buildup of echo suppression in normal-
1817 hearing adults,” *J. Acoust. Soc. Am.* **99**(2), 1118–1123.

1818
1819
1820
1821 Hartmann, W. M., and Wittenberg, A. (1996). “On the externalization of sound images,”
1822

1823 J. Acoust. Soc. Am. **99**(6), 3678–3688.

1824 Hassager, H. G., Gran, F., and Dau, T. (2016). “The role of spectral detail in the binaural
1825 transfer function on perceived externalization in a reverberant environment,” J. Acoust.
1826 Soc. Am. **139**, 2992–3000.

1830 Heine, L., Corneyllie, A., Gobert, F., Luauté, J., Lavandier, M., and Perrin, F. (2021) “Virtually
1831 spatialized sounds enhance auditory processing in healthy participants and patients with a
1832 disorder of consciousness,” Sci. Rep. **11**, 13702.

1836 Hendrickx, E., Stitt, P., Messonnier, J., Lyzwa, J., Katz, B. F., and de Boishéraud, C.
1837 (2017). “Influence of head tracking on the externalization of speech stimuli for non-
1838 individualized binaural synthesis,” J. Acoust. Soc. Am. **141**(3), 2011–2023.

1843 Hocking, J., Dzafic, I., Kazovsky, M., and Copland, D. A. (2013). “NESSTI: Norms for
1844 environmental sound stimuli,” PLoS One **8**(9), e73382.

1847 Jeffress, L. A., and Taylor, R. W. (1961). “Lateralization vs localization,” J. Acoust. Soc.
1848 Am. **33**, 482–483.

1851 Kates, J. M., Arehart, K. H., Muralimanohar, R. K., and Sommerfeldt, K. (2018). “Exter-
1852 nalization of remote microphone signals using a structural binaural model of the head and
1853 pinna,” J. Acoust. Soc. Am. **143**, 2666–2677.

1857 Kim, S.-M., and Choi, W. (2005). “On the externalization of virtual sound images in head-
1858 phone reproduction: a wiener filter approach,” J. Acoust. Soc. Am. **117**, 3657–3665.

1861 Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., and Pardhan, S. (2016). “Auditory
1862 distance perception in humans: A review of cues, development, neuronal bases, and effects
1863 of sensory loss,” Atten. Percept. Psychophys. **2**(78), 373–395.

1867 Kopčo, N., Doreswamy, K. K., Huang, S., Rossi, S., and Ahveninen, J. (2020). “Cortical

1868
1869 auditory distance representation based on direct-to-reverberant energy ratio,” *NeuroImage*
1870
1871 **208**, 116436.

1872
1873 Kopčo, N., Huang, S., Belliveau, J. W., Raji, T., Tengshe, C., and Ahveninen, J. (2012).
1874 “Neuronal representations of distance in human auditory cortex,” *Proc Natl Acad Sci USA*
1875
1876 **109**(27), 11019–11024.

1877
1878
1879 Kopčo, N., and Shinn-Cunningham, B. G. (2011). “Effect of stimulus spectrum on distance
1880
1881 perception for nearby sources,” *J. Acoust. Soc. Am.* **130**(3), 1530–1541.

1882
1883 Laback, B. (2023). “Contextual lateralization based on interaural level differences is pre-
1884
1885 shaped by the auditory periphery and predominantly immune against sequential segrega-
1886
1887 tion,” *Trends Hear.* **27**, 23312165231171988.

1888
1889 Leclère, T., Lavandier, M., and Perrin, F. (2019). “On the externalization of sound sources
1890
1891 with headphones without reference to a real source,” *J. Acoust. Soc. Am.* **146**(4), 2309–
1892
1893 2320.

1894
1895 Li, S., Schlieper, R., and Peissig, J. (2019). “The role of reverberation and magnitude spectra
1896
1897 of direct parts in contralateral and ipsilateral ear signals on perceived externalization,”
1898
1899 *Appl. Sci.* **9**(3), 460.

1900
1901 Lingner, A., Pecka, M., Leibold, C., and Grothe, B. (2018). “A novel concept for dynamic
1902
1903 adjustment of auditory space,” *Sci Rep* **8**, 8335.

1904
1905 Mershon, D. H., and Bowers, J. N. (1979). “Absolute and relative cues for the auditory
1906
1907 perception of egocentric distance,” *Perception* **8**, 311–322.

1908
1909 Parseihian, G., Jouffrais, C., and Katz, B. F. G. (2014). “Reaching nearby sources: comparison
1910
1911 between real and virtual sound and visual targets,” *Frontiers in Neuroscience* **8**, 269.
1912

- 1913 Philbeck, J. W., and Mershon, D. H. (2002). "Knowledge about typical source output
1914 influences perceived auditory distance," *J. Acoust. Soc. Am.* **111**, 1980–1983.
1915
1916
1917 Prud'homme, L., and Lavandier, M. (2020). "Do we need two ears to perceive the distance
1918 of a virtual frontal sound source?," *J. Acoust. Soc. Am.* **148**(3), 1614–1623.
1919
1920 Savel, S. (2009). "Individual differences and left/right asymmetries in auditory space
1921 perception. I. Localization of low-frequency sounds in free field," *Hearing Research*
1922 **255**(1-2), 142-154.
1923
1924
1925
1926 Sound-Ideas-Series (1992). "Sound ideas series 6000," in *General Sound Effects Library*.
1927
1928 Wenzel, E. M. (1995). "The relative contribution of interaural time and magnitude cues to
1929 dynamic sound localization," in *IEEE Workshop on Applications of Signal Processing to*
1930 *Audio and Acoustics*, pp. 80–83.
1931
1932
1933
1934 Werner, S., Klein, F., Mayenfels, T., and Brandenburg, K. (2016). "A summary on acoustic
1935 room divergence and its effect on externalization of auditory events," 8th International
1936 Conference on Quality of Multimedia Experience, 1–6.
1937
1938
1939
1940 Zahorik, P. (2002). "Assessing auditory distance perception using virtual acoustics," *J.*
1941 *Acoust. Soc. Am.* **111**, 1832–1846.
1942
1943
1944 Zahorik, P. (2002b). "Auditory display of sound source distance," in *Proceedings of the 8th*
1945 *International Conference on Auditory Display*, edited by R. Nakatsu and H. Kawahara
1946 (Kyoto, Japan), 239–243.
1947
1948
1949
1950 Zahorik, P., Brungart, D. S., and Bronkhorst, A. W. (2005). "Auditory distance perception
1951 in humans: A summary of past and present research," *Acta Acustica united with Acustica*
1952 **91**, 409–420.
1953
1954
1955
1956