



**HAL**  
open science

## Targeted Cortical Manipulation of Auditory Perception

Sebastian Ceballo, Zuzanna Piwkowska, Jacques Bourg, Aurélie Daret, Brice Bathellier

► **To cite this version:**

Sebastian Ceballo, Zuzanna Piwkowska, Jacques Bourg, Aurélie Daret, Brice Bathellier. Targeted Cortical Manipulation of Auditory Perception. *Neuron*, 2019, 10.1016/j.neuron.2019.09.043 . hal-02379771

**HAL Id: hal-02379771**

**<https://hal.science/hal-02379771v1>**

Submitted on 21 Jul 2022

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



22 **Abstract**

23 Driving perception by direct activation of neural ensembles in cortex is a necessary step for  
24 achieving a causal understanding of the neural code for auditory perception and developing  
25 central sensory rehabilitation methods. Here, using optogenetic manipulations during an  
26 auditory discrimination task in mice, we show that auditory cortex can be short-circuited by  
27 coarser pathways for simple sound identification. Yet, when the sensory decision becomes  
28 more complex, involving temporal integration of information, auditory cortex activity is  
29 required for sound discrimination and targeted activation of specific cortical ensembles  
30 changes perceptual decisions as predicted by our readout of the cortical code. Hence, auditory  
31 cortex representations contribute to sound discriminations by refining decisions from parallel  
32 routes.

## 33 **Introduction**

34 The role of primary sensory cortical areas in perceptual decisions is complex. Primary sensory  
35 areas are often viewed as necessary links between peripheral sensory information and  
36 decision centers, but multiple observations challenge this simplified model. For example,  
37 human subjects with primary visual cortex lesions display residual visual abilities, a  
38 phenomenon termed ‘blindsight’ (Sanders et al., 1974; Schmid et al., 2010). In animals,  
39 classical associative conditioning (Lashley, 1950; LeDoux et al., 1984) or some operant  
40 behaviors (Hong et al., 2018) based on sensory stimuli can be performed in the absence of  
41 primary sensory cortex. However, several studies also report sensory-based behaviors that are  
42 abolished or severely impaired by primary sensory cortex silencing (Letzkus et al., 2011;  
43 O'Connor et al., 2010; Poort et al., 2015; Sachidhanandam et al., 2013). In addition, cortical  
44 stimulation experiments show that primary cortex for all sensory modalities can perturb  
45 perceptual decisions or initiate sensory-driven behaviors (Choi et al., 2011; Houweling and  
46 Brecht, 2008; Huber et al., 2008; Musall et al., 2014; O'Connor et al., 2013; Peng et al., 2015;  
47 Salzman et al., 1990; Yang et al., 2008; Znamenskiy and Zador, 2013), suggesting a role in  
48 perception. This apparent contradiction is particularly evident in hearing, for which  
49 involvement of auditory cortex (AC) is controversial even for discrimination of two distinct  
50 sounds. Indeed, lesions or reversible silencing of auditory cortex lead to deficits or have little  
51 effect depending on task conditions, silencing methods and animal models (Diamond and  
52 Neff, 1957; Gimenez et al., 2015; Harrington et al., 2001; Jaramillo and Zador, 2011;  
53 Kuchibhotla et al., 2017; Ohl et al., 1999; Pai et al., 2011; Rybalko et al., 2006; Talwar et al.,  
54 2001). Thus, auditory cortex does not seem to be always necessary for sound discrimination.

55

56 Moreover, requirement of auditory cortex in particular task settings points towards two  
57 alternative mechanistic implications. One alternative is that AC requirement reflects a

58 permissive role for the task (Otchy et al., 2015), for example by providing some global gating  
59 signals to other areas, which is not informative about the decision to be taken, but without  
60 which the behavioral decision process is impaired. The second alternative is that AC actually  
61 provides for each stimulus distinct pieces of information which contribute to drive the  
62 discriminative choices. Optogenetic manipulations of AC activity can modulate auditory  
63 discrimination performance (Aizenberg et al., 2015), which could be both explained by a  
64 permissive or a driving role of AC in the task. Targeted manipulation of AC outputs in the  
65 striatum can bias sound frequency discrimination towards the sound frequency corresponding  
66 to the preferred frequency of the manipulated neurons (Znamenskiy and Zador, 2013). This  
67 indicates that AC output can be sufficient to drive discrimination, but as the effect of  
68 complete AC inactivation was not tested in this task, it remains possible that this manipulation  
69 does not reflect the natural drive occurring in the unperturbed behavior. In support of this,  
70 targeted manipulation of generic AC neurons failed to drive consistent biases in this task  
71 (Znamenskiy and Zador, 2013). Thus, necessity and sufficiency of precise auditory cortex  
72 activity patterns in a sound discrimination behavior remain to be established.

73

74 Here we combine optogenetic silencing and patterned activation techniques in head-fixed  
75 mice to show that AC is not required in a simple frequency discrimination task but is  
76 necessary for a difficult discrimination involving sounds with frequency overlap. We also  
77 show based on a discrimination of distinct optogenetically driven AC activity patterns that  
78 specific AC information is sufficient for driving a discrimination, and decisions in this case  
79 also take longer than choices made in the simple task. Last, we show that focal stimulation of  
80 AC is able to modify the animal's choice in the difficult task but not in the simple one. We  
81 also show that the AC region most sensitive to focal stimulation encodes auditory features  
82 used by the mouse to discriminate the two trained sounds. These results indicate that auditory

83 cortex provides necessary and sufficient information to drive decisions in difficult sound  
84 discriminations, while other pathways bypass it in simple discriminations.

85

## 86 **Results**

### 87 **Requirement of auditory cortex depends on sound discrimination complexity.**

88 Lesion studies in gerbils, rats, cats and monkeys (Diamond and Neff, 1957; Harrington et al.,  
89 2001; Ohl et al., 1999; Rybalko et al., 2006) suggest that the involvement of auditory cortex  
90 in sound discrimination could be related to the sound features that have to be discriminated,  
91 and thus potentially to the difficulty of the discrimination. To test this idea we trained  
92 different groups of mice to perform a head-fixed Go/NoGo task, one group was trained to  
93 discriminate pure frequency features (a 4 kHz against a 16 kHz pure tone, PTvsPT task) while  
94 a second group had to integrate frequency variations over time in order to discriminate a  
95 linearly rising frequency modulated sound (4-12 kHz) against a pure tone (4 kHz, FMvsPT  
96 task), (**Figure 1A**). While the PTvsPT task was rapidly learned, the FMvsPT task was more  
97 challenging as it required much longer training ( $82 \pm 12$  trials,  $n = 34$  mice for the PTvsPT,  
98 against  $416 \pm 54$  trials,  $n = 28$  mice, for FMvsPT task to reach 80% performance; Wilcoxon  
99 rank-sum test,  $p = 2.5 \times 10^{-8}$ , **Figure 1B,C**). We thus wondered whether this difference in task  
100 difficulty could relate to differential involvement of auditory cortex (AC). To address this  
101 question, we decided to silence AC by activating parvalbumin-positive interneurons (PV)  
102 expressing Channelrhodopsin2 (hChR2-tdTomato). Electrophysiological calibration in awake,  
103 passive mice (**Figure 2A** and **S1**) showed that this manipulation strongly disrupted cortical  
104 activity (e.g. **Figure 2B**), decreasing population firing rates to target sounds by ~70%, with  
105 about 50% of putative PV-negative neurons displaying a complete suppression of their  
106 response (**Figure 2C** and **S1**). To evaluate the effect of PV-activation on task performance,  
107 we trained a linear Support Vector Machine (SVM) classifier to discriminate the two target

108 sounds based on the population activity of putative PV-negative neurons (n = 125 neurons,  
109 recorded in 14 separate populations across 4 mice). We then measured the performance of the  
110 classifier with AC responses to the same sounds during PV-activation. For the pure tones, the  
111 SVM performance dropped from almost perfect classification to less than 70% and for the  
112 pure tone versus FM sound, SVM performance dropped to chance levels (**Figure 2D**),  
113 suggesting that the optogenetic manipulation perturbs sound encoding, although, given the  
114 small size of our neuronal sample, we cannot verify to what extent the information remains at  
115 a broader scale.

116

117 When bilaterally applying this PV-activation protocol through cranial windows in behaving  
118 mice, we indeed observed a drop of performance down to chance level (50%) in the FMvsPT  
119 task, independent of whether the FM sweep was the rewarded or non-rewarded stimulus  
120 (**Figure S2**), while mice lacking Chr2 expression performed normally (**Figure 2E**).  
121 However, the same manipulation during the PTvsPT task yielded less than a 10%  
122 performance drop (**Figure 2E**), which is significantly lower than the ~30% performance drop  
123 to chance level observed for the FMvsPT task (Wilcoxon rank sum test,  $p = 0.002$ ,  $n = 6$  and  $7$   
124 mice). This mild impact was not due to the wide difference in frequency in our PTvsPT task,  
125 as mice were not more impaired by optogenetic AC silencing when discriminating 4 kHz  
126 versus 8 kHz or 4 kHz versus 6 kHz (Figure S2), but we do not exclude AC involvement for  
127 even finer discriminations. Because the same inactivation strategy applied to the inferior  
128 colliculus, an earlier stage of the auditory system, fully abolished PTvsPT performance  
129 (**Figure 2F**), we suspected that the continued performance of the PTvsPT task after AC  
130 perturbation was not due to incompleteness of the silencing (**Figure 2C and S1**), but rather to  
131 the absence of AC involvement. We thus performed bilateral lesions that covered the entire  
132 AC and observed only a ~10% drop in behavioral performance, even for a smaller frequency

133 difference, on the day following lesion (**Figure 2G** and **S2**) and (**Figure S2**). In summary, AC  
134 is dispensable for a simple pure tone discrimination while it is necessary for the more difficult  
135 FMvsPT task.

136

### 137 **Information from auditory cortex can drive slow discriminative choices**

138 Necessity alone does not prove that AC activity patterns causally drive decisions in the  
139 FMvsPT task (Hong et al., 2018; Otchy et al., 2015). Seeking to establish a proof of  
140 sufficiency for AC activity in this task, we first tested if two distinct activity patterns  
141 unilaterally delivered to AC can provide enough information to drive a discrimination task, as  
142 so far, cortical stimulation studies have only demonstrated detection of a single cortical  
143 stimulation (Houweling and Brecht, 2008; Huber et al., 2008; Musall et al., 2014; O'Connor et  
144 al., 2013; Peng et al., 2015; Salzman et al., 1990; Znamenskiy and Zador, 2013) or  
145 discrimination of two stimulations in different cortical areas or hemispheres (Choi et al.,  
146 2011; Manzur et al., 2013; Yang et al., 2008). To this end, we used a micro-mirror device  
147 (Dhawale et al., 2010; Zhu et al., 2012) to apply two-dimensional light patterns onto the AC  
148 surface through a cranial window in Emx1-Cre x Ai27 mice expressing ChR2 in pyramidal  
149 neurons (**Figure 3A**).

150 We chose to trigger the two discriminated activity patterns with light disks of 0.4 mm  
151 diameter at two distinct positions. To make sure that optogenetic patterns were qualitatively  
152 similar to sound-evoked activity, we calibrated them in awake mice by recording isolated  
153 single units with multi-electrode silicon probes while targeting the light disks to many  
154 locations within the cranial window (**Figure 3A-B**). In this procedure we observed that  
155 recorded single units tended to respond only to a subset of the tested locations (e.g. **Figure**  
156 **3C**). Pooling together the activity of all recorded neurons, we observed that optogenetically-



157 and sound-triggered responses had similar latencies and time courses above a 50 ms time  
158 scale (**Figure 3B**, note that optogenetic responses followed the 20 Hz modulation of the light  
159 stimulus below the 50 ms time scale). Moreover, response amplitudes were similar in  
160 particular for the lowest light intensity tested (12mW, **Figure 3B**), reflected by a high  
161 correlation of response amplitudes to the preferred sound and the optogenetic pattern (**Figure**  
162 **3D**). Last, the spatial extent of optogenetic responses (600-800  $\mu\text{m}$  diameter, depending on  
163 light intensity, **Figure 3E**) matched the extent of pure tone responses in primary auditory  
164 cortex (e.g. **Figure S3**) and changed little with depth (**Figure 3F-G**), although evoked  
165 population firing rates were clearly reduced in the deepest layers. Together, these  
166 measurements indicated that the chosen cortical stimulation patterns yielded response  
167 characteristics within the range of natural sound responses.

168 Using the same behavioral protocol as for the sound discrimination tasks, we trained mice to  
169 discriminate between unilateral cortical stimulations ( $32 \text{ mW}\cdot\text{mm}^{-2}$ ) in low versus high  
170 frequency tonotopic areas of the primary auditory field of AC (**Figure 4A-C**). The locations  
171 of these areas with respect to blood vessel patterns were identified using intrinsic imaging  
172 (**Figure 4C and S3**) as previously established (Bathellier et al., 2012; Deneux et al., 2016;  
173 Kalatsky et al., 2005; Nelken et al., 2004). Blood vessels were then used as landmarks to  
174 robustly target the same regions across days. We observed that mice could learn to  
175 discriminate these artificial AC patterns within hundreds of trials (**Figure 4D and S3**),  
176 showing AC activity can drive a discrimination task.

177 Strikingly however, discrimination of cortical patterns was executed about four times slower  
178 than the cortex-independent discrimination of real pure tones ( $\sim 400 \text{ ms}$  vs  $\sim 100 \text{ ms}$ , **Figure**  
179 **4E-F and S4**), even for small tone frequency differences (**Figure S4**), as measured by  
180 determining the first time point after stimulus onset at which licking behavior significantly  
181 differed for S+ and S- stimuli (**Figure S4**). In addition,  $\sim 400\text{ms}$  reaction times were observed

182 for the detection of optogenetic activation in AC, independent of pattern size and intensity  
183 (**Figure S4**), showing that the long decision latencies for AC-driven discrimination are not  
184 due to insufficient drive from the artificial cortical patterns, but eventually to a rate-limiting  
185 process in the pathway downstream of AC that triggers the animal's decisions. Alternatively,  
186 abnormal aspects of the optogenetic stimulation such as weaker activation in deeper layers  
187 may explain the long latencies. Interestingly, in the cortex-dependent auditory task (FMvsPT),  
188 discrimination times were also ~400 ms (**Figure 4E-F**), but contrary to direct cortical  
189 stimulation, discrimination was preceded by an early non-discriminative response whose  
190 onset timing was similar to the discriminative lick response in the PTvsPT task (**Figure 4F**,  
191 arrowhead), which is likely a response to the 4 kHz frequency range, present at the onset of  
192 both S+ and S- stimuli.

193 These observations suggest that sound discrimination activates two pathways. (i) An AC-  
194 independent pathway producing an early response which can distinguish very distinct sounds  
195 but fails to discriminate more complex differences developing over time. (ii) An AC-  
196 dependent, slower pathway which can provide specific information to improve the responses  
197 initiated by the AC-independent pathway, in case these fail to discriminate the proposed  
198 sounds, in particular when critical information arrives only after an ambiguous sound onset. In  
199 fact, AC is not required for discrimination of an FM sweep that starts at a frequency very  
200 distinct from the discriminated pure tone (**Figure S2**), and in this case, non-discriminative  
201 licking at sound onset is weak, allowing short latency discrimination (**Figure S4**). This  
202 suggests that what makes the discrimination difficult in our FMvsPT task is the necessity to  
203 ignore the non-informative sound onsets to make the correct decision.

204

205 **Focal auditory cortex stimulations bias decisions in the complex discrimination**

206 If indeed auditory cortex provides specific information to drive the correct decision  
207 exclusively in the difficult task, then specific activation of neural ensembles in AC should  
208 influence behavioral decisions in the FMvsPT task but not in the PTvsPT task. To test this  
209 hypothesis, we trained two groups of Emx1-Cre x Ai27 to perform these tasks. For both  
210 groups, the non-rewarded stimulus (S-) was the 4 kHz pure tone. The rewarded stimulus was  
211 the 16 kHz pure tone for one group and the 4-12 kHz FM sweep for the other group. After  
212 task acquisition, the right AC of each mouse was functionally mapped using intrinsic imaging  
213 to determine its tonotopic organization (**Figure 5A**). Then a grid of optogenetic stimulus  
214 locations (400 $\mu$ m disks) was aligned to this map. Each location was stimulated during  
215 behavioral performance as occasional non-rewarded catch trials (10% occurrence) together  
216 with the non-rewarded 4 kHz pure tone (**Figure 5B**). The purpose of this design was for the 4  
217 kHz tone to drive the rapid non-specific response and test if additional specific cortical drive  
218 mimics, at least partially, the perception of the rewarded stimulus instead of the presented S-,  
219 thereby changing the late discriminative response. To mimic sound-induced activation as  
220 closely as possible with our method, we used the lower of our two calibrated light intensities  
221 (12 mW.mm<sup>-2</sup>, **Figure 3C-D**). Unlike previous studies (O'Connor et al., 2013; Znamenskiy  
222 and Zador, 2013), these catch trials were not rewarded to measure a spontaneous  
223 interpretation of optogenetic perturbations by the brain and rule out reward-associated  
224 responses potentially generated by fast learning processes. Each of the 17 locations in the grid  
225 was tested across five catch trials and the average lick count elicited was compared to the  
226 mean expected lick count distribution in five regular S- trials to assess statistical significance  
227 (e.g. **Figure 5C**, p<0.05). We accounted for multiple testing across the 17 locations using the  
228 Benjamini-Hochberg correction. This conservative analysis showed that specific locations  
229 (e.g. **Figure 5D**) significantly increased lick counts above random fluctuations and brought  
230 the behavioral response close to the normal rewarded stimulus (S+) response for the FMvsPT

231 task (**Figure 5E** and **S5**). Moreover, the fraction of locations perturbing behavior was  
232 significantly above the false positive rate across all animals tested (**Figure 5F**). This was not  
233 the case for the easier PTvsPT task (**Figure 5F**), in which no stimulus location could raise the  
234 lick count close to S+ level (**Figure 5E** and **S5**). Together, this experiment showed that  
235 targeted, physiologically realistic AC stimulation is sufficient to change the decisions of mice  
236 in the difficult auditory task but not in the simple task, corroborating the idea that AC is not  
237 involved in the former but contributes decisive information in the latter. This idea was also  
238 reinforced by the observation that locations significantly perturbing behavior clustered around  
239 a specific area of AC when mapped across all animals in the FMvsPT task but not in PTvsPT  
240 task (**Figure 5G-H**). This area spanned the mid-frequency range of the primary auditory field  
241 (A1) and to a lesser extent the central region of AC (**Figure 5H**) as defined from intrinsic  
242 imaging maps (**Figures 5A** and **6A**).

243

#### 244 **Tonotopic location of effective focal stimulations matches frequency cues used for** 245 **discrimination**

246 The central, non-tonotopic region of auditory cortex, potentially part of secondary auditory  
247 areas (A2), was recently suggested to contain neurons specific for vocalizations and FM  
248 sweeps (Honma et al., 2013; Issa et al., 2014; Issa et al., 2017). This is in line with its  
249 contribution to the decisions in the FMvsPT task. We thus wondered if the mid-frequency  
250 area of A1 is also involved in coding the 4-12 kHz FM sweep. First, we observed that strong  
251 intrinsic imaging responses to the FM sweep occur in A1, particularly in the mid-frequency  
252 area (**Figure 6A**). To better quantify the information contained in this area, we performed  
253 two-photon calcium imaging in layer 2/3 of AC in awake, passively listening mice who had  
254 been injected with pAAV1.Syn.GCaMP6s virus. All imaged fields were aligned to intrinsic

255 imaging maps using blood vessel patterns (**Figure 6B**) allowing a comparison of micro- and  
256 mesoscale maps. Color coding cell locations with respect to their best frequency, we  
257 qualitatively observed a good match between tonotopic maps derived from the intrinsic and  
258 calcium imaging (**Figure 6C and S6**), even if a certain level of disorganization could be seen  
259 in the high resolution map as reported previously (Bandyopadhyay et al., 2010; Rothschild et  
260 al., 2010). For example, the mid-frequency area identified with intrinsic imaging mostly  
261 contained cells with best frequencies at 8 or 16 kHz when imaged at high resolution (e.g.  
262 **Figure 6C and S6**). Based on these data, we calculated the similarity (correlation coefficient)  
263 between population responses to FM sweeps and to different pure tones and found, both at  
264 coarse and high spatial resolution, that the response to the FM sweeps is most similar to the  
265 responses to pure tones around 8 kHz (**Figure 6D**). We reasoned that neurons coding for this  
266 frequency range in A1 could be particularly important for discriminating the FM sweep from  
267 the 4 kHz tone. To quantify this idea with our two-photon calcium imaging measurements, we  
268 calculated the fraction of neurons responding to different pure tones within the population that  
269 responds to the FM sweep but not to the 4 kHz pure tone (discriminative FM sweep neurons).  
270 This showed that neurons which discriminatively respond to the FM sweep indeed have a  
271 preference for frequencies around 8 kHz (**Figure 6E**), a result in line with the localization of  
272 the most disruptive optogenetic perturbations in the mid-frequency region of A1, at least if  
273 AC representations are similar in passive and active mice.

274 To evaluate if context-dependent changes (Francis et al., 2018; Kuchibhotla et al., 2017)  
275 could impact our conclusions, we also performed calcium imaging in mice engaged in the  
276 FMvsPT task. We found as previously reported positive and negative modulations  
277 (Kuchibhotla et al., 2017) for a fraction of neural responses during engagement, which  
278 averaged to a net decrease of population responsiveness (**Figure S7**). However, AC  
279 population activity classifiers trained in the passive state maintained good performance in

280 discriminating the sounds when tested in the active state (**Figure S7**), confirming the overall  
281 similarity of active and passive sound representations in AC.

282 Thus together, our optogenetic and imaging experiments suggest that AC contributes to the  
283 discrimination of the 4-12 kHz sweep vs 4 kHz pure tone by engaging spatially segregated  
284 neurons that preferentially code for intermediate frequencies around 8 kHz. A prediction of  
285 this model is that mice engaged in the FMvsPT discrimination should strongly associate the 4-  
286 12 kHz sweep with pure tones in the 8 kHz frequency range. To test this prediction, we  
287 presented pure tones in occasional non-rewarded catch trials during the task and measured the  
288 licking responses of the mouse (**Figure 7A**). We observed that mice responded to 5.6 kHz and  
289 8 kHz non-rewarded pure tones with a number of licks similar to the lick count observed after  
290 the rewarded 4-12 kHz sweep (**Figure 7B**), revealing the expected association of these  
291 frequencies with the FM sound. We thus wondered if this categorization behavior crucially  
292 depends on neurons sensitive to frequencies in the 8 kHz region. We trained a linear classifier  
293 (SVM) to discriminate the 4-12 kHz sweep and the 4 kHz pure tone based on the responses of  
294 neurons imaged in passive mice and significantly activated by only one of the two sounds  
295 (discriminative neurons). When testing this classifier with the single trial population  
296 responses elicited by pure tones, we could qualitatively reproduce the categorization curves  
297 observed during behavior (**Figure 7B**). This indicates a match between the structure of the  
298 coding space in AC and of the perceptual space as probed in the FMvsPT behavioral task.  
299 However, when the classifier was trained without the neurons that significantly respond to the  
300 8 kHz pure tone, then no match was observed between behavioral and cortical categorization  
301 curves, although the remaining neurons still discriminate the FM sweep from the 4 kHz pure  
302 tone (**Figure 7B**). This result corroborates the idea that activation of neurons responding in  
303 the 8 kHz frequency range critically contributes to the identification of the FM sweep against  
304 the 4 kHz tone, in line with the observation that focal optogenetic stimulation in the region

305 most sensitive to the mid-frequency ranges promotes licking responses associated with the  
306 FM sweep (**Figure 5H**).

307

## 308 **Discussion**

309 Together, our results demonstrate that, while auditory cortex is dispensable for a simple pure  
310 tone discrimination task, specific activity patterns in AC provide necessary and sufficient  
311 information for discrimination of a more complex sound pair (FMvsPT task). Here the more  
312 complex task involves two sounds which overlap in frequency at their onset but then diverge  
313 due to upward frequency modulation for one of them. The two sounds are thus theoretically  
314 distinguishable thanks to both spectral (mid- and high-frequencies of the FM sound) and  
315 temporal (continuous frequency variations) cues. It is striking that despite the large spectral  
316 differences between the two sounds of the FMvsPT task, this discrimination was much harder  
317 for mice to learn than the pure tone discrimination task. The challenge likely comes from the  
318 frequency overlap at the beginning of the two sounds as suggested by the lack of AC  
319 involvement when the FM sweep starts from a frequency very different from the pure tone  
320 (**Figure S2 and S4**). This spectral overlap implies that both sounds activate a large group of  
321 identical neurons throughout the auditory pathway, as we observe in the auditory cortex (the  
322 FM sweep also activates the 4 kHz region, **Figure 6A and S5**). Activation of the set of  
323 neurons coding for the overlapping frequency is a strong and early predictor of rewards in S+  
324 trials. Therefore, a simple reinforcement learning rule might associate activity of these  
325 neurons to the licking response, which makes it difficult for the animal to refrain from licking  
326 in S- trials, in which the same neurons are active. The challenge for the brain might be to use  
327 information from the subsequent, non-overlapping parts of the sound representations to  
328 counteract such generalization. Analysis of the licking profiles in the FMvsPT task indicates

329 that mice do not fully resolve this problem. Indeed, their initial response both to S- and S+ is  
330 to lick shortly after sound onset. It is only after this initial impulse that modulation of licking  
331 leads to the discriminative response (**Figure 4F**). The role of auditory cortex might be to  
332 solve this initial confound, while alternative pathways are able to learn the correct  
333 associations when discriminated stimuli are encoded with little overlap as expected for the  
334 distant pure tones of the PTvsPT task. Corroborating this explanation, when the direction of  
335 the FM sweep is reversed (12 to 4 kHz), thereby removing the initial frequency overlap,  
336 discrimination with the 4 kHz pure tone is much less sensitive to AC lesions (**Figure S2**) and  
337 discrimination time is similar to the PTvsPT task (**Figure S4**). Interestingly, a similar  
338 phenomenon is observed in a Pavlovian discriminative fear condition protocol, in which  
339 conditioning to a specific direction of frequency ramps cannot be achieved without AC, while  
340 AC is dispensable for fear conditioning to a specific pure tone (Dalmay et al.). Thus, the  
341 requirement of auditory cortical areas for distinguishing temporal variations in frequency  
342 overlapping stimuli might be a generic principle independent of the sound association  
343 protocol used to uncover it (Diamond and Neff, 1957; Harrington et al., 2001; Ohl et al.,  
344 1999), with possible hemispheric specializations as pointed out by a previous study (Wetzel et  
345 al., 2008).

346 Our study does not identify the alternative pathways which provide auditory information to  
347 motor centers in the absence of AC in the simple PTvsPT task. Connectivity studies suggest  
348 multiple connections between the auditory system and motor-related centers (Pai et al., 2011).  
349 Primary auditory thalamus is known to project to striatum (LeDoux et al., 1991) a brain  
350 region necessary for appetitive sound discrimination tasks (Guo et al., 2018). Thalamo-striatal  
351 projections could possibly implement the main sensory-motor association in the easy task  
352 (Gimenez et al., 2015; Guo et al., 2017). Sensory-motor associations are also possible at the  
353 mid-brain level, where inferior colliculus contacts superior colliculus which in turn projects to



354 several motor centers (Stein and Stanford, 2008). The strong impact of optogenetic silencing  
355 of the inferior colliculus on the PTvsPT discrimination suggests that the sensory-motor  
356 association does not occur at brainstem level before information reaches colliculus (**Figure**  
357 **2A**), although further investigation would be necessary to verify that our colliculus silencing  
358 experiment does not uncover a permissive effect of colliculus in the task (Otchy et al., 2015).  
359 Likewise, thalamic silencing was shown to strongly impact frequency discrimination  
360 (Gimenez et al., 2015), suggesting that the sensory-motor association in our PTvsPT task does  
361 not occur before auditory information reaches thalamus. Finally, another possibility is that  
362 decisions in the PTvsPT task are driven by alternative cortical routes excluding auditory  
363 cortex but receiving auditory information from primary or secondary thalamus. Such direct  
364 pathways, for example to more associative areas have been proposed to support “blindsight”  
365 abilities in patients lacking primary visual cortex (Cowey, 2010). In line with this possibility,  
366 recent results indicate that fear conditioning to simple sounds does not require auditory cortex  
367 but requires the neighboring temporal association area (Dalmay et al.). Also, recordings in  
368 prefrontal cortex during an auditory discrimination task show the presence of fast responses to  
369 sounds (Fritz et al., 2010), compatible with a direct pathway. Similarly, in our task, bypass  
370 pathways to temporal or prefrontal associative areas could eventually be sufficient to make  
371 simple auditory decisions, even in absence of auditory cortex.

372 Independent of the identity of the alternative pathway involved in the PTvsPT, our results  
373 suggest that it provides a faster route to generate motor responses than the route which  
374 involves auditory cortex. This idea derives from the long discrimination times (~400 ms)  
375 observed when mice discriminate two optogenetically-driven cortical activity patterns as  
376 compared to the short discrimination times (~150ms) observed for the PTvsPT task (**Figure**  
377 **4E-F**). This long discrimination time is unlikely to result from difficulties in differentiating  
378 the two optogenetic stimuli, as detection of one stimulus required also ~400 ms independent

379 of stimulus strength and size (**Figure S4**), consistent with previous measurements in another  
380 optogenetic stimulus detection task (Huber et al., 2008). Some electrical microstimulation  
381 experiments report shorter reaction times in behavior (Houweling and Brecht, 2008). This  
382 could be due to the weaker drive elicited by optogenetics in deeper layers (**Figure 3F-G**).  
383 Alternatively, microstimulation can directly activate principal neurons in subcortical areas by  
384 antidromic propagation, which does not occur in optogenetics due to specificity of expression  
385 and less efficient activation of antidromic spikes (Tye et al., 2011). Antidromic activations  
386 could recruit pathways faster than those activated by cortex alone. Calibration of our  
387 optogenetic stimuli also showed that the two chosen stimulus locations can be efficiently  
388 discriminated based on triggered cortical activity within 50 to 100 ms (**Figure S4**). However,  
389 although population firing rate and spatial extent parameters of optogenetic stimuli were  
390 similar to sound responses in AC (**Figure 3**), we cannot fully rule out that some non-  
391 controlled parameters of the artificial stimulations lead to extended reaction times. It is also  
392 possible that absence of subcortical drive in the optogenetic task leads to threshold effects that  
393 slow down the behavioral response. Yet, in support of a slow sensory cortical pathway,  
394 discrimination times observed in cortex-dependent sensory discrimination tasks in head-fixed  
395 mice using visual or tactile cues are typically above 300-400 ms (O'Connor et al., 2010; Poort  
396 et al., 2015; Sachidhanandam et al., 2013). In the cortex-dependent auditory task presented in  
397 our study (FMvsPT), discrimination times are also long (~350 ms, **Figure 4E-F**). We  
398 measured that AC activity starts to be discriminative for the FM sweep and 4 kHz pure tone  
399 only about 100 to 150 ms after sound onset while two pure tones are discriminated by AC  
400 activity already 50 ms after onset (**Figure 2D**). This leaves a supplementary delay between  
401 information arrival and decision of at least 100 ms in the cortex-dependent FMvsPT task as  
402 compared to the cortex independent PTvsPT task. This delay could be due to complex  
403 multisynaptic pathways downstream to AC (e.g. including associative cortical areas) or to

404 downstream temporal integration processes. Timing differences could also play a role in the  
405 absence of AC engagement in simple tasks. Indeed, while AC rapidly receives auditory  
406 information (**Figure 2D**), its slow impact on behavior could allow alternative pathways which  
407 provide faster responses to short-circuit information coming from AC, which would be  
408 ignored by decision areas. It is also possible that alternative pathways inhibit AC output when  
409 they provide a solution to the task.

410 This might explain the absence of strong perturbations of the easy PTvsPT task when we  
411 applied non-rewarded focal stimulations of AC, calibrated to typical sound response levels,  
412 despite specific targeting to the tonotopic fields relevant for the discrimination. We do not  
413 exclude that rewarding licking responses to the occasional focal stimulations could rapidly  
414 lead, after some learning, to a reinforced participation of stimulated cortical areas in  
415 behavioral decisions.

416 In the case of the harder FMvsPT task, focal stimulations had a significant effect on behavior  
417 despite the absence of rewards in our protocol. This indicates that, despite the coarseness of  
418 the stimulation approach which globally targets heterogeneous cell ensembles enriched with  
419 particular spectral information due to tonotopic organization (**Figure 6** and **S6**), some  
420 optogenetically-driven ensembles activate large enough parts of the stimulus representations  
421 to change behavioral decisions. Thus, our results demonstrate that manipulation of auditory  
422 cortex representations impact perception, at least, if the behavioral task for perceptual read out  
423 is appropriately chosen. This opens interesting possibilities for central auditory rehabilitation  
424 techniques. In the FMvsPT task, we took advantage of the fact that spectral cues can be used  
425 to distinguish the two auditory stimuli, such that neural ensembles relevant for discrimination  
426 in AC tend to be clustered in space and can be activated with broad light patterns, while  
427 minimizing activation of cell types carrying confounding information. However, spatial  
428 organization is not absolutely strict in mouse auditory cortex (Bandyopadhyay et al., 2010;

429 Rothschild et al., 2010), especially for more complex features, including temporal  
430 modulations (Deneux et al., 2016; Kuchibhotla and Bathellier, 2018). Thus, fine scale  
431 stimulation methods (Packer et al., 2015) implemented at a sufficiently large scale would be  
432 necessary to precisely interfere at AC level with the discrimination of sounds that only differ  
433 based on complex, non-spectral features, and thereby eventually construct more precise  
434 artificial auditory perceptions.

435 The fact that the AC pathway is dispensable for simple discriminations suggests that auditory  
436 judgments, and potentially perception, result from the interplay between a coarse description  
437 of sensory inputs and a time-integrated, more elaborate description that involves auditory  
438 cortex. As we showed by attempting to drive auditory judgments directly at the cortical level  
439 (**Figure 5**), the coexistence of these pathways is a critical issue to manipulate auditory  
440 perception. It might thus be advantageous to combine stimulations of cortical and sub-cortical  
441 (Guo et al., 2015) levels in order to improve the quality of artificially generated perception in  
442 the central auditory system.

443 **Acknowledgements**

444 We thank A. Chédotal, D. DiGregorio, Y. Frégnac, J. Letzkus and E. Harrell for comments on  
445 the manuscript, G. Hucher for histology and the GENIE Project, Janelia Farm Research  
446 Campus, Howard Hughes Medical Institute, for GCAMP6s constructs. This work was  
447 supported by the Agence Nationale pour la Recherche (ANR “SENSEMAKER”), the Fyssen  
448 foundation, the Human Brain Project (WP 3.5.2), the DIM “Region Ile de France”, the Marie  
449 Curie Program (CIG 334581), the International Human Frontier Science Program  
450 Organization (CDA-0064-2015), the Fondation pour l’Audition (Laboratory grant), the  
451 European Research Council (ERC CoG DEEPEN), the DIM Cerveau et Pensée and Ecole des  
452 Neurosciences de Paris Ile-de-France (ENP, support to SC).

453

454 **Author contributions**

455 BB, ZP and SC designed the study. BB, AD and SC performed behavioral experiments. BB  
456 and JB performed inferior colliculus surgeries. SC and ZP performed cortical activation  
457 experiments. SC and BB analyzed the data and wrote the manuscript with comments from all  
458 authors.

459 **Declaration of Interests**

460 The authors declare no competing interests.

461

## 462 **Figures**

463

464 **Figure 1 | Discriminating a frequency-modulated sound from a pure tone is more**  
465 **difficult than discriminating two distant pure tones. A.** Sketch of the head-fixed Go/NoGo  
466 discrimination task. **B.** Learning curves for all mice performing each task. **C.** Number of trials  
467 needed to reach 80% discrimination performance for PTvsPT and FMvsPT discrimination  
468 task (PTvsPT, n = 34 mice. FMvsPT, n = 28 mice. Wilcoxon rank-sum test,  $P < 0.001$ ). Open  
469 circles correspond to single mice. Bar plots show the mean and SEM.

470

471 **Figure 2 | Cortical requirement for sound discrimination depends on discrimination**  
472 **complexity. A.** Sketch of the optogenetic inactivation strategy through a cranial window  
473 covering the primary and secondary auditory cortex. For calibration, silicon probes were  
474 inserted beneath the window. **B.** Sound response PSTH of a sample PV-negative cell with and  
475 without light-driven PV-neuron activation. **C.** Mean population response to sounds with and  
476 without optogenetic activation of parvalbumin neurons ( $15 \pm 1.4$  Hz vs  $4.7 \pm 0.9$  Hz; n = 125  
477 single units,  $p = 6.2 \times 10^{-18}$ , Wilcoxon rank-sum test) and distribution of sound response  
478 reduction during light-ON trials for putative PV-negative neurons. **D.** S+ versus S-  
479 discrimination performance for linear SVM classifiers trained on a sample of normal  
480 population responses (5 repetitions, one classifier for each 50 ms time bin) and tested on  
481 either normal population responses or on population responses recorded during optogenetic  
482 inactivation (5 independent repetitions, on the same 50 ms time bin) (n = 125 single units,  
483 from 4 mice). **E.** Discrimination performance without (black) and with (blue) optogenetic  
484 inactivation of AC during PTvsPT (left,  $91 \pm 2.4\%$  vs  $82 \pm 2\%$ , n = 6 mice,  $p = 0.025$ ,  
485 Wilcoxon rank-sum test) and FMvsPT (right,  $83 \pm 1.7\%$  vs  $52 \pm 3.5\%$ , n = 7 mice,  $p = 0.002$ ,  
486 Wilcoxon rank-sum test) tasks in mice expressing ChR2 in PV interneurons. The same  
487 measurements are shown for control mice without ChR2 expression (dark blue). **F.**

488 Discrimination performance without (black) and with (blue) optogenetic inactivation of  
489 inferior colliculus during PTvsPT (left,  $85 \pm 2.3\%$  vs  $54 \pm 1.9\%$ ,  $n = 6$  mice,  $p = 0.004$ ,  
490 Wilcoxon rank-sum test) and FMvsPT (right,  $75 \pm 2.7\%$  vs  $52 \pm 5.6\%$ ,  $n = 5$  mice,  $p = 0.016$ ,  
491 Wilcoxon rank-sum test) tasks. **G.** Left; Example histological section showing the extent of  
492 bilateral AC lesions. Right; Mean discrimination performance during PTvsPT before and after  
493 AC lesion ( $97 \pm 1.1\%$  vs  $87 \pm 2.9\%$ ,  $n = 4$  mice,  $p = 0.043$ , Wilcoxon rank-sum test) or sham  
494 surgery ( $93 \pm 1.4\%$  vs  $91 \pm 4.5\%$ ,  $n = 4$  mice,  $p = 0.8$ , Wilcoxon rank-sum test). Bar plots  
495 show the mean and SEM throughout the figure.

496

497 **Figure 3 | 2D light pattern delivery triggers focal activity in auditory cortex.** **A.** Left,  
498 Photo-stimulation setup and optogenetic calibration protocol. Center, example image of a  
499  $\sim 400 \mu\text{m}$  disk projected onto the cranial window. Right, image of the AC cranial window  
500 with a sketch of the silicon probe insertion site and grid locations where light was projected  
501 (light blue). **B.** Time course of mean firing rate (50ms time bin) in response to a single pure  
502 tone of 500 ms at 70 dB SPL (black line) or in response to  $12 \text{ mW}\cdot\text{mm}^{-2}$  (left) or  $32 \text{ mW}\cdot\text{mm}^{-2}$   
503 (right) focal optogenetic stimulations (light blue line, disk diameter  $400 \mu\text{m}$ ) for  $n = 84$  single  
504 units (shadings represent SEM). Insets: magnification at stimulus onset. **C.** 2D maps of an  
505 example single unit for two light intensities, representing the mean firing rate over 10  
506 repetitions during photo-stimulation of the different locations on the grid. **D.** Best photo-  
507 stimulation response plotted against best sound response measured as the largest trial-  
508 averaged response over a 500 ms window starting at stimulus onset. For  $12 \text{ mW}\cdot\text{mm}^{-2}$  the  
509 slope of the regression line was 1.01 and the correlation coefficient 0.74,  $p = 2.4 \times 10^{-16}$ ,  $n = 84$   
510 single units recorded between 300 and 900  $\mu\text{m}$  in depth. **E.** Mean of Gaussian models that  
511 were fitted to the 2D lateral distribution profile of photo-stimulation responses (shadings  
512 represent SEM). Half-width diameter:  $560 \pm 147 \mu\text{m}$  at  $12 \text{ mW}\cdot\text{mm}^{-2}$  and  $800 \pm 235 \mu\text{m}$  at

513 32mW.mm<sup>-2</sup> (n = 84 single units). **F.** 2D maps as in **C**, together with the fitted 2D Gaussian  
514 model for three single units recorded at three different depths (L2-L4: >100 μm to < 350 μm,  
515 L4-L5: >350 μm to < 600 μm, L5-L6: >600 μm). **G.** Mean 1D profile of the fitted Gaussian  
516 models as in **E**, but for each estimated depth range defined in **F** and for each light intensity.  
517 Number of cells: 32 mW.mm<sup>-2</sup>, L2-L4: n =31, L4-L5: n = 66 and L5-L6: n = 47; 12 mW.mm<sup>-2</sup>,  
518 L2-L4: n =29, L4-L5: n = 62 and L5-L6: n = 34).

519

520 **Figure 4 | Mice can discriminate two distinct artificial activity patterns in auditory**  
521 **cortex.** **A.** Sketch of the Go/NoGo discrimination task for optogenetic stimuli with sample  
522 trials showing licking signals for each stimulus. **B.** Sample cranial window superimposed with  
523 the location of two optogenetic stimuli in AC. **C.** Localization of the two optogenetic stimuli  
524 in the AC tonotopic map obtained from intrinsic imaging. Same mouse as in **B.** **D.** Population  
525 learning curve for discrimination of two optogenetic stimulations (340±105 trials to reach  
526 80% correct performance, n = 5 mice). **E.** Mean decision times for the three different  
527 discrimination tasks (optogenetic 424 ± 58 ms, n = 5 mice; FMvsPT, 366 ± 25 ms, n = 29  
528 mice; PTvsPT, 153 ± 22 ms, n = 30 mice). PTvsPT was significantly different from the two  
529 other groups (p = 7.3×10<sup>-9</sup> for FMvsPT and p = 1.1×10<sup>-3</sup> for optogenetic, Wilcoxon rank-sum  
530 test) which were not different from each other (p = 0.2, Wilcoxon rank-sum test). **F.** Mean  
531 licking traces for S+ and S- stimuli during optogenetic (top), FMvsPT (middle) and PTvsPT  
532 (bottom) discrimination. Arrowhead: unspecific initial licks. Grey shading: typical early  
533 response delay. Green shading: typical late discrimination delay.

534



535 **Figure 5 | Focal optogenetic AC activation biases decisions only in the FMvsPT task. A.**  
536 Intrinsic imaging response maps for three pure tones aligned and averaged across 10 mice.  
537 Bottom right, map of tonotopic gradients obtained by subtracting the 4 and 32 kHz response  
538 maps. The approximate contours and direction of the main tonotopic gradients (primary  
539 auditory field A1, anterior auditory field AAF, secondary auditory field A2) are  
540 superimposed. **B.** Schematic of the focal AC perturbation experiment: in occasional non-  
541 rewarded trials, the S- sound (4 kHz) was played synchronously with a local optogenetic  
542 stimulation in one out of 17 predefined locations aligned to the tonotopic map. **C.** Probability  
543 distribution of S- lick counts averaged over three random trials in one sample mouse. The  
544 superimposed red circles represent the mean lick counts over three trials for each optogenetic  
545 stimulus location. Open circles: non-significant responses. Filled circles: lick count is larger  
546 than 95% of S- lick counts. Filled circles with black contours: significant locations after  
547 Benjamini-Hochberg correction for multiple testing ( $P < 0.05$ ). **D.** Map of normalized lick  
548 counts for optogenetic perturbations obtained by summing the estimated 2D neuronal  
549 population response profile for one light disk, multiplied by observed lick count and  
550 positioned at all optogenetic locations with a significant response. **E.** Mean lick counts for all  
551 optogenetic stimulus locations and all mice involved in the PTvsPT or FMvsPT tasks. Red  
552 filled circles correspond to significant locations ( $P < 0.05$ , corrected for multiple testing). **F.**  
553 Percentage of significant locations found in the PTvsPT and FMvsPT tasks across all mice.  
554 Data are represented as mean  $\pm$  SEM. Dashed line: expected false discovery rate. **G.** Maps of  
555 normalized lick counts for significant optogenetic locations averaged across all mice  
556 performing the PTvsPT task. Black circles: centers of optogenetic stimuli. **H.** Same as G. but  
557 for the FMvsPT task.  
558

559 **Figure 6 | AC neurons coding for intermediate frequencies distinguish sounds of the**  
560 **difficult task.** **A.** Sketch of intrinsic imaging setup and averaged re-aligned intrinsic imaging  
561 maps sampled over 12 mice for 4, 8 and 16 kHz pure tones and FM sweeps. White lines and  
562 arrows indicate the main tonotopic fields as in **Figure 5A.** **B.** Sketch of 2-photon setup and  
563 picture of the AC surface with the localization of calcium imaging fields-of-view (FoV) for a  
564 sample mouse. A magnification of FoV 5 is shown on the right. Raw calcium traces (scale  
565 bar: 50 %  $\Delta f/f_0$ ) and a pure tone tuning curve (70 dB SPL) for a sample neuron are shown on  
566 the bottom. The calcium imaging setup is sketched on the left. **C.** Magnification of cranial  
567 window shown in B with the contours of intrinsic imaging responses to 4 and 8 kHz pure  
568 tones (left) and the locations of all significantly responding neurons which do not respond to  
569 both 4 kHz and 4-12 kHz (discriminative neurons see *methods*) recorded with 2 photon  
570 calcium imaging (non-discriminative neurons removed). Color-code corresponds to their  
571 preferred frequency over the same sounds used during intrinsic imaging experiments. Circle  
572 diameters are proportional to the mean deconvolved calcium response for the best frequency  
573 sound over 20 repetitions (right). The dashed line indicates the position of the main tonotopic  
574 fields. **D.** *Red dots:* Correlation of the pooled intrinsic response maps of the FM sweeps with  
575 pooled maps of four different pure tones (n = 12 mice for the aligned maps). *Black dots:*  
576 correlation between the population vector responses of the FM sweeps and of various pure  
577 tones (time bin 0 to 1s after sound onset, from 5157 significantly responding neurons over 30  
578 sessions in 5 mice). Intrinsic: max=1; min 0.75. Calcium imaging: max = 0.25; min =0. **E.**  
579 Fraction of neurons that respond significantly to the FM sweep (n = 520 neurons, Wilcoxon  
580 rank sum test,  $\alpha = 0.01$ ) from the pool of significantly responding neurons (discriminative  
581 neurons, n = 4692 neurons).  
582

583 **Figure 7 | Mice use intermediate frequencies as a cue to discriminate between FM**  
584 **sweeps and low frequency pure tones. A.** Sketch of the auditory categorization assay. **B.**  
585 Mean response probabilities in response to unrewarded pure tones occasionally replacing  
586 trained stimuli in the FMvsPT discrimination task (mean  $\pm$  SEM, n = 6 mice). Circles in red  
587 are the probabilities of an SVM classifier, trained to discriminate between FM sweeps and 4  
588 kHz pure tones, to classify different pure tones as an FM sweep. Population vectors used to  
589 train the classifier correspond to the mean deconvolved calcium responses of 4692  
590 significantly responding and discriminative neurons in a [220 ms ; 448 ms] time bin with 20  
591 repetitions for each sound). Circles in grey correspond to the same analysis but after exclusion  
592 of neurons that respond significantly to an 8 kHz pure tone.  
593  
594

## 595 **STAR Methods**

### 596 **CONTACT AND MATERIALS AVAILABILITY**

597 Further information and requests for resources and reagents should be directed to and will be  
598 fulfilled by the Lead Contact, Brice Bathellier ([brice.bathellier@cnrs.fr](mailto:brice.bathellier@cnrs.fr)). This study did not  
599 generate new unique reagents.

600

### 601 **EXPERIMENTAL MODEL AND SUBJECT DETAILS.**

602 We used the following mouse lines: PV-Cre (Jax #008069) x Ai27 (flex-CAG-hChR2-  
603 tdTomato; Jax # 012567), for optogenetic inactivation, Emx1-IRES-Cre (Jax #005628) x Ai27  
604 (Jax # 012567) for optogenetic activation, and C57Bl6J for lesion experiments. In all  
605 experiments, young adult females and males between 8 to 16 weeks old were used. Animals  
606 were housed 1–4 animals per cage, in normal light/dark cycle (12 h/12 h). All procedures  
607 were in accordance with protocols approved by the French Ethical Committee (authorization  
608 00275.01).

609

### 610 **METHOD DETAILS**

611 **Behavior.** Behavioral experiments were monitored and controlled using homemade software  
612 (Elphy, G. Sadoc, UNIC, France) coupled to a National Instruments card (PCIe-6351).  
613 Sounds were amplified (SA1 Stereo power amp, Tucker-Davis Technologies) and delivered  
614 through high frequency loudspeakers (MF1-S, Tucker-Davis Technologies) in a pseudo-  
615 random sequence. Water delivery (5–6  $\mu$ l) was controlled with a solenoid valve (LVM10R1-  
616 6B-1-Q, SMC). A voltage of 5V was applied through an electric circuit joining the lick tube  
617 and aluminum foil on which the mouse was sitting, so that lick events could be monitored by  
618 measuring voltage through a series resistor in this circuit (**Figure 2F**). Before starting the  
619 training procedure, mice were water restricted for two consecutive days. The first day of

620 training consisted of a habituation period for head fixation and to reliably receive water by  
621 licking the lick port without any sound. After this period, S+ training was conducted for 2 or 3  
622 days where S+ trials were presented with 80-90% probability, while the remaining trials were  
623 blank trials (no stimulus). A trial consisted of a random inter-trial interval (ITI) between 6 and  
624 8 seconds to avoid prediction of stimulus appearance, a random 'no lick' period between 3  
625 and 5 seconds and a fixed response window of 1.5 seconds. Licking during the response  
626 window on an S+ trial above lick threshold (3-5 consecutive licks) was scored as a 'hit' and  
627 triggered immediate water delivery. No licks was scored as a 'miss' and the next trial  
628 immediately followed. Each behavioral session contained ~150 rewarded trials allowing mice  
629 to obtain their daily water supply of ~800  $\mu$ l. At the beginning of each session, ~20 trials with  
630 'free rewards' were given independent of licking to motivate the mice. When animals reached  
631 more than 80% 'hits' for the S+ stimulus, the second sound (S-) was introduced, and the lick  
632 count threshold was set to 5 licks (in a few PTvsPT experiments, only 3 licks). During  
633 presentation of the S- sound, licking below threshold was considered as a 'correct rejection'  
634 (CR) and the next trial immediately followed, licking above threshold on S- trials was  
635 considered as a 'false alarm' (FA), no water reward was given, and the animal was punished  
636 with a random time out period between 5 and 7 seconds. Each session then contained 300  
637 trials with 50% probability for each trial type. Sounds of 0.2 seconds at 60 dB were used for  
638 the PT vs PT tasks and sounds of 0.5 seconds at 70 dB SPL for the FMvsPT tasks. During the  
639 entire behavioral training period, food was available ad libitum and animal weight was  
640 monitored daily. Water restriction (0.8mL/day) was interleaved with a 12h ad libitum supply  
641 overnight every Friday. Mice performed behavior five days per week (Monday to Friday).  
642 With this schedule, mice initially showed a weight drop < 20% and progressively regained  
643 their initial weight over about 30 days.

644

645 **Behavior analysis.** Learning curves were obtained by calculating the fraction of correct  
646 responses over blocks of 10 trials. Discrimination performance over one session was  
647 calculated as (hits + correct rejections)/total trials. Discrimination time was calculated on  
648 collections of 10 trial blocks in which the discrimination performance was greater than or  
649 equal to 80%, containing a total of at least 100 trials. The licking signal for each stimulus was  
650 binned in 10 ms bins, and for every time bin, we used the non-parametric Wilcoxon rank sum  
651 test to obtain the p-value for the null hypothesis that the licking signal was coming from the  
652 same distribution for S+ and S-. The discrimination time was determined as the first time bin  
653 above a p-value threshold of 0.01 while applying the Benjamini-Hochberg correction for  
654 multiple testing over all time bins.

655

656 **Cranial window implantation.** To allow chronic unilateral access to AC (for 2-photon  
657 calcium imaging or electrophysiology), a cranial window was incorporated into the skull and  
658 a metal post for head fixation was implanted on the contralateral side of the craniotomy.  
659 Surgery was performed in 4 to 6 week old mice placed on a thermal blanket under anesthesia  
660 using a mix of ketamine (Ketasol) and medetomidine (Domitor) (antagonized with  
661 atipamezole - Antisedan, Orion pharma – at the end of the surgery). The eyes were covered  
662 using Ocry gel (TVM Lab) and Xylocaine 20mg/ml (Aspen Pharma) was injected locally at  
663 the site where the incision was made. The right masseter was partially removed and a large  
664 craniotomy (~5 mm diameter) was performed above the auditory cortex (AC) using bone  
665 sutures of the skull as a landmark. For 2P calcium imaging, we did 3 to 5 injections (~300  $\mu$ m  
666 separation) of 200nL (35 nl/min) using pulled glass pipettes of rAAV1.syn.GCamP6s.WPRE  
667 virus ( $10^{13}$  virus particles per ml) diluted 10 times (Vector Core, Philadelphia, PA, USA).  
668 After this, the craniotomy was immediately sealed with a 5 mm circular cover slip using  
669 cyanolite glue and dental cement (Ortho-Jet, Lang). For AC inactivation experiments with

670 optogenetics, the same procedure was repeated on both brain hemispheres in a single surgery.  
671 For inferior colliculus inactivation experiments, the 5mm cranial window was placed on the  
672 midline such that it covered the dorsal part of both hemispheres of the inferior colliculus. In  
673 all cases, mice were subsequently housed for at least one week without any manipulation.

674

675 **Optogenetics inactivation of AC during behavior.** For optogenetic inactivation  
676 experiments, mice were first trained to respond to the S+ stimulus alone (see Behavior), then  
677 when the S- stimulus was introduced for discrimination training, a pair of blue LEDs (1.1W,  
678 PowerStar OSOLON Square 1+, ILS, wavelength 455nm) was placed 2 cm above their head,  
679 facing each other at 1 cm distance, and were flashed 1 over 5 trials for visual habituation of  
680 the animal to light flashes (initially inhibiting licking). When discrimination performance was  
681 above 80% and similar both with and without light flashes, one to three test sessions were  
682 performed in which the LEDs were either placed on the bilateral AC cranial windows, or on  
683 the inferior colliculus window (1 LED only). Stimulus presentation was pseudo-randomized  
684 over blocks of 100 trials and light flashes appeared in 1 out of 5 trials for each stimulus with  
685 the same reward or punishment conditions as regular trials. Blue light delivery followed a  
686 square wave (20Hz) time course and started 100ms before onset lasting for 700ms (sound  
687 duration was no longer than 500ms). Light intensity at the brain surface was measured to be  
688 36 mW/mm<sup>2</sup>. Discrimination performance was computed as the number of hits and correct  
689 rejections over the total number of trials, separating trials in which the LED was turned ON or  
690 OFF.

691

692 **Patterned optogenetics during behavior.** For patterned optogenetic activation in the mouse  
693 AC, we used a LED-based video projector (DLP LightCrafter, Texas Instruments) tuned at  
694 460 nm. To project a two-dimensional image onto the auditory cortex surface (**Figures 2** and

695 4), the image of the micromirror chip was collimated through a 150 mm cylindrical lens  
696 (Thorlabs, diameter: 2 inches) and focused through a 50 mm objective (NIKKOR, Nikon).  
697 Light collected by the objective passes through a dichroic beam splitter (long pass, >640nm,  
698 FF640-FDi01, Semrock) and is collected by a CCD camera (GC651MP, Smartek Vision)  
699 equipped with a 50 mm objective (Fujinon, HF50HA-1B, Fujifilm). For discrimination of  
700 artificial AC patterns, two disks of 400  $\mu\text{m}$  were placed at two different locations in AC. The  
701 first disk was defined as the S+ stimulus and systematically located at the center of the low  
702 frequency domain in A1. The second disk, defined as the S- stimulus was placed at the center  
703 of the high frequency domain in the UF obtained with intrinsic imaging (**Figure S4**, see  
704 Intrinsic Imaging below). Alignment of optogenetic stimulus locations across days was done  
705 using blood vessel patterns at the surface of the brain with a custom made GUI in Matlab. In  
706 short, a reference blood vessel image was taken at the beginning of the experiment. In  
707 subsequent days, a new blood vessel image was taken and aligned to the reference image by  
708 optimizing the image cross- correlation to obtain the appropriate rotation and translation  
709 matrix. Behavioral training with optogenetic stimuli was done with the same protocol as for  
710 sounds.

711

712 For focal AC activation during sound discriminations, a grid of 15 to 17 locations was  
713 constructed using disks of  $\sim 400 \mu\text{m}$  distributed in 5 rows. The grid was aligned to auditory  
714 cortex for each animal using the three regions of maximal responses in the intrinsic imaging  
715 response map to a 4 kHz pure tone (see Intrinsic Imaging) as a reference for horizontal  
716 position and orientation. Each grid was constructed to maximize the coverage of the different  
717 frequency domains and tonotopic subfields seen in intrinsic imaging (see **Figure 4**). To probe  
718 responses to optogenetic perturbation during discrimination behavior we performed test  
719 sessions of 300 pseudo-randomized trials in which 45 to 54 trials were non-rewarded catch



720 trials with stimulation of a grid location together with the S- sound. Sound (4 kHz, 500 ms at  
721 70 dB SPL) and light (12 mW/mm<sup>2</sup>, 1 s duration, square wave intensity profile at 20 Hz)  
722 started at the same time. Responses to focal optogenetic perturbations were computed over 3  
723 catch trial repetitions. The significance of these responses was assessed by computing the  
724 distribution of lick count responses over 10,000 random triplets of responses to S- alone. The  
725 p-value of the response to a particular optogenetic stimulation location was taken as the  
726 percentile of this distribution corresponding to the trial-averaged lick count for the  
727 optogenetic stimulation. Based on the p-values computed for all grid locations, we performed  
728 a Benjamini-Hochberg correction for multiple testing to identify the locations with a  
729 significant response at a false positive rate of 0.05. For display, lick count responses (L) to  
730 optogenetic perturbations were normalized for each mouse as  $(L-s-)/(s+ - s-)$ , where s+ and s-  
731 are the mean lick counts observed for S+ and S- stimuli. Maps of response to optogenetic  
732 perturbation were aligned across different mice, by matching the grid locations placed on the  
733 4 kHz landmarks, calculating the best rotation and translation matrix as for alignment of the  
734 intrinsic imaging maps (see Intrinsic Imaging). To account for the actual spread of the  
735 optogenetic stimulus in the AC network, the response at each location was represented by a  
736 two-dimensional spatial profile identical to the estimated profile shown in **Figure 2A** for 12  
737 mW/mm<sup>2</sup> focal stimuli. Profiles of all significant locations were summed to construct the  
738 maps (**Figure 4F**).

739

740 **In vivo electrophysiology.** Recordings for calibration of optogenetics were done in mice  
741 implanted already with a cranial window above AC for at least 2 or 3 weeks. On the day of  
742 the recording, the mouse was briefly anesthetized (~30 min, ~1% isoflurane delivered with  
743 SomnoSuite, Kent Scientific) to remove some of the cement seal and a piece of the cover slip  
744 of the cranial window was cut using a diamond drill bit. The dura was resected ventral to AC

745 as previously delimited using intrinsic imaging (see methods below). The area was covered  
746 with Kwik-Cast<sup>TM</sup> silicon (World precision instruments) and the animal was placed in his  
747 home cage to recover from anesthesia for at least 1 hour before head-fixation in the recording  
748 setup (same as used for behavior). 14 recordings in 4 mice (84 single units) were performed  
749 using four shank Buzsaki32 silicon probes (Neuronexus). For estimation of optogenetic  
750 responses at different depths, 3 recordings in 3 mice (144 single units) were performed using a  
751 single shank 64 contacts across 1.4 mm silicon probes (A1x64-Poly2-6mm-23s-160,  
752 Neuronexus). Before each recording, the tips of the probe were covered with DiI (Sigma). The  
753 silicon was gently removed and the area cleaned using warm Ringer's buffer. The probe was  
754 inserted at a  $\sim 30^\circ$  angle with respect to brain surface with a micromanipulator (MP-225,  
755 Butter Instrument) at 1-2  $\mu\text{m}$  per second, with pauses of 1-2 minutes every 50  $\mu\text{m}$ . Each  
756 recording session time, including probe insertion and removal, lasted no more than 3 hours.  
757 After the experiment, the animal was deeply anesthetized (isoflurane) and euthanized by  
758 cervical dislocation. Brains were fixed overnight by 4% paraformaldehyde (PFA) in 0.1 M  
759 phosphate buffer (PB). Coronal brain slides of 80  $\mu\text{m}$  were prepared and imaged (Nikon  
760 eclipse 90i, Intensilight, Nikon) to identify the electrode track tagged with DiI and determine  
761 recording depth. For calibration of optogenetic inactivation experiments, sounds played  
762 during the recordings included a blank, 4, 4.5, 4.7, 5, 6, 8, 16 kHz pure tones of 200 ms  
763 duration at 60 dB SPL, a 4 kHz 500 ms pure tone at 70 dB SPL, 4 to 12 kHz 500 ms  
764 frequency modulated sound at 70 dB SPL and white noise ramps of 1000 ms from 60 to 85  
765 dB SPL and 85 to 60 dB SPL. The same blue directional LED ( $24^\circ$  light cone) as used for  
766 inactivation during behavior was placed above the cranial window, however at a slightly  
767 higher distance (about 0.75 cm) of the window to leave space for the electrode. Light was  
768 delivered for 700 ms, starting 100 ms before sound onset. Silicon probe voltage traces were  
769 recorded at 20 kHz and stored using RHD2000 USB interface board (Intan Technologies).

770 Raw voltage traces were filtered using a Butterworth high-pass filter with a 100 Hz cutoff  
771 (Python). Electro-magnetic artefacts from the LED driving current were removed by  
772 subtracting a template calculated across all LED ON trials. For calibration of focal  
773 optogenetic activations, we used a grid of contiguous 8x5 400  $\mu\text{m}$  circles. Each location was  
774 played randomly and repeated 10 times. Also, a 4 kHz pure tone of 500 ms at 70 dB SPL, 4-  
775 12 kHz FM sound of 500 ms at 70 dB SPL and white noise ramps of 1000 ms from 60 to 85  
776 dB SPL and 85 to 60 dB SPL were played to compare later with optogenetic responses.

777

778 **Analysis of electrophysiology.** Spikes were detected and sorted using the KlustaKwik spike  
779 sorting algorithm (Harris et al., 2000) (Klusta, <https://github.com/kwikteam/klusta>) with a  
780 strong and weak threshold of 6 and 3 respectively. Each shank was sorted separately where  
781 putative single units were visualized and sorted manually using KlustaViewa. Data analysis  
782 was done using custom Python scripts. Spikes were binned in 25 ms bins. Firing rates during  
783 stimulation periods were calculated by averaging across trials. For AC inactivation,  
784 significant responses to at least one sound were identified using the Kruskal-Wallis H-test for  
785 independent samples and the Benjamini-Hochberg procedure was applied to correct for  
786 multiple testing across units ( $p < 0.05$ ). Percentage of inhibition was calculated as follows:  
787  $100 - (L * 100 / S)$ , where S corresponds to best sound response and L to the response to the  
788 same sound during optogenetic activation of PV interneurons. To estimate the robustness of  
789 sound representations in our recordings, we trained a linear Support Vector Machine classifier  
790 to discriminate two sounds based on single unit responses binned in 50 ms bins. Only units  
791 that were inhibited by light (putative PV-negative neurons) were used to train the classifier.  
792 To evaluate classification performance without optogenetics, the classifier was trained over 5  
793 trials for each pair of sounds and tested with other 5 trials. To test the effect of light on sound  
794 responses, the classifier was trained with 10 unperturbed sound delivery trials and tested with

795 10 sounds and light delivery trials. For focal optogenetic activations, we calculated for each  
796 single unit the trial-averaged firing rate change with respect to baseline over the 1s  
797 optogenetic stimulation for all locations, yielding two-dimensional spatial response maps.  
798 Using the centers of each location, a 2D map was created for each unit and fitted with a two-  
799 dimensional Gaussian model that is composed of an offset term (constant over space) and a  
800 Gaussian spatial modulation term. Significance of the modulation was assessed using a  
801 bootstrap analysis (significance threshold, 0.05). Cells that did not show significant spatial  
802 modulation were described by the offset term only.

803

804 **Auditory cortex lesions and immunohistochemistry.** Mice after learning a PTvsPT  
805 discrimination task were anesthetized (~1.5% isoflurane delivered with SomnoSuite, Kent  
806 Scientific) and placed on a thermal blanket. Craniotomies were performed as described above  
807 and focused thermo-coagulation lesions were done bilaterally. The area was then covered  
808 using Kwik-Cast silicon (World precision instruments) and closed with dental cement (Ortho-  
809 Jet, Lang). After a period of recovery on a heating pad with accessible food pellets, mice were  
810 taken back to their home cage and a nonsteroidal anti-inflammatory agent (Metacam®,  
811 Boehringer Ingelheim) was injected intramuscularly. To test if mice could still perform the  
812 discrimination task, they were placed on the behavioral setup the next day after surgery. If a  
813 mouse presented signs of pain, liquid meloxicam was given in drinking water or via  
814 subcutaneous injection. Discrimination was tested in a normal session and performance  
815 calculated as previously described. After the experiment, mice were transcardially perfused  
816 with saline followed by 4% paraformaldehyde (PFA) in 0.1 M phosphate buffer (PB) and then  
817 brains post-fixed overnight at 4°C. After washing with phosphate-buffered saline (PBS),  
818 brains were cut in 80 µm coronal slides and immuno-histochemical reactions were performed  
819 on free-floating brain slices as follows. Slides were blocked for 2 h using PBS + 10% goat

820 serum and 1% Triton-X 100 at room temperature. After washing in PBS (10min×3), the  
821 sections were incubated 2h at room temperature with a dilution 1:100 of mouse anti-NeuN  
822 conjugated with Alexa Fluor® 488 (MAB377X, Merck). Slides were washed and mounted for  
823 imaging using a Nikon eclipse 90i microscope (Intensilight, Nikon).

824

825 **Intrinsic imaging and alignment of tonotopic maps across animals.** Intrinsic imaging was  
826 performed to localize AC in mice under light isoflurane anesthesia (~1% delivered with  
827 SomnoSuite, Kent Scientific) on a thermal blanket. Images were acquired using a 50 mm  
828 objective (1.2 NA, NIKKOR, Nikon) with a CCD camera (GC651MP, Smartek Vision)  
829 equipped with a 50 mm objective (Fujinon, HF50HA-1B, Fujifilm) through a cranial window  
830 implanted 1-2 weeks before the experiment (4-pixel binning, field of view between 3.7 x 2.8  
831 mm or 164 × 124 binned pixels, 5.58- $\mu$ m pixel size, 20 fps). Signals were obtained under 780  
832 nm LED illumination (M780D2, Thorlabs). Images of the vasculature over the same field of  
833 view were taken under 480 nm LED illumination (NSPG310B, Conrad). Two second  
834 sequences of short pure tones at 80 dB SPL were repeated every 30 seconds with a maximum  
835 of 10 trials per sound. Acquisition was triggered and synchronized using a custom made GUI  
836 in Matlab. For each sound, we computed baseline and response images, 3 seconds before and  
837 3 seconds after sound onset, respectively. The change in light reflectance  $\Delta R/R_0$  was  
838 calculated over repetitions for each sound frequency (4, 8, 16, 32 kHz, white noise). Response  
839 images were smoothed applying a 2D Gaussian filter ( $\sigma = 3$  pixels). The different subdomains  
840 of AC corresponding to the tonotopic areas appeared as regions with reduced light  
841 reflectance. To align intrinsic imaging responses from different animals, the 4 kHz response  
842 was used as a functional landmark. The spatial locations of maximal amplitude responses in  
843 the 4 kHz response map for the A1, A2 and AAF (three points) was extracted for each mouse  
844 and a Euclidean transformation matrix was calculated by minimizing the sum of squared

845 deviations (RMSD) for the distance between the three landmarks across mice. This procedure  
846 yielded a matrix of rotation and translation for each mouse that was applied to compute  
847 intrinsic imaging responses averaged across a population of mice.

848

849 **Two-photon calcium imaging.** Two-photon imaging during behavior was performed using a  
850 two-photon microscope equipped with an 8 kHz resonant scanner (Femtonics, Budapest,  
851 Hungary) coupled to a pulsed Ti:Sapphire laser system (MaiTai DS, Spectra Physics, Santa  
852 Clara, CA). The laser was tuned at 900 nm during recordings and light was collected through  
853 a 20x (XLUMPLFLN-W) or 10x (XLPLN10XSVMP) Olympus objective. Images were  
854 acquired at 31.5 Hz. For each behavioral trial, imaging duration was 7s with a pause of 5s in  
855 between trials. Sound delivery was randomized during the trial to prevent mice from starting  
856 licking with the onset of the sound (45 dB SPL) emitted by the microscope scanners at the  
857 beginning of each trial. All sounds were delivered at 192 kHz with a National Instruments  
858 card (NI-PCI-6221) driven by homemade software (Elphy, G. Sadoc, UNIC, France), through  
859 an amplifier (SA1 Stereo power amp, Tucker-Davis Technologies) and high frequency  
860 loudspeakers (SA1 and MF1-S, Tucker-Davis Technologies, Alachua, FL). During the active  
861 context mice performed the discrimination task, a regular session consisted in ~250 trials,  
862 followed immediately after by a passive context session with ~150 trials (lick tube was  
863 withdrawn). Data analysis was performed using Matlab and Python scripts. Motion artifacts  
864 were first corrected frame by frame, using a rigid body registration algorithm. Regions of  
865 interest (ROIs) corresponding to the neurons were selected using Autocell a semi-automated  
866 hierarchical clustering algorithm based on pixel covariance over time (Roland et al., 2017).  
867 Neuropil contamination was subtracted (Kerlin et al., 2010) by applying the following  
868 equation:  $F_{\text{corrected}}(t) = F_{\text{measured}}(t) - 0.7 F_{\text{neuropil}}(t)$ , where  $F_{\text{neuropil}}(t)$  is estimated from the  
869 immediate surroundings (Gaussian smoothing kernel, excluding the ROIs (Deneux et al.,

870 2016),  $\sigma = 170\mu\text{m}$ ). Then the change in fluorescence ( $\Delta F/F_0$ ) was calculated as  $(F_{\text{corrected}}(t) -$   
871  $F_0) / F_0$ , where  $F_0$  is estimated as the 3<sup>rd</sup> percentile of the low-pass filtered fluorescence over  
872  $\sim 40$  s time windows period. To estimate the time-course of the firing rate, the calcium signal  
873 was temporally deconvolved using the following formula:  $r(t) = f'(t) + f(t) / \tau$  in which  $f'$  is  
874 the first time derivative of  $f$  and  $\tau$  the decay constant set to 2 seconds for GCaMP6s. In total  
875 7605 neurons were recorded across 11 active and 11 passive sessions in 3 mice. We kept for  
876 analysis only 1008 neurons significantly responding to the S+ or S- stimuli with respect to  
877 baseline activity (Wilcoxon rank-sum test,  $p = 0.01$  and Bonferroni correction for multiple  
878 testing). The deconvolved signals were smoothed with a Gaussian kernel ( $\sigma = 33\text{ms}$ ). To  
879 estimate the discriminability of two sounds based on cortical population responses, linear  
880 Support Vector Machine classifiers were trained independently on each time point to  
881 discriminate population activity vectors obtained from half of the presentations of each sound  
882 and context (training set), and were tested on activity vectors obtained on the remaining  
883 presentations of the same sounds and context or on all presentations of sounds in the non-  
884 trained context (test sets). To estimate behavioral categorization of sounds based on  
885 population activity in AC, we trained linear SVM classifier to discriminate AC responses to  
886 single presentations of the trained target sounds, and tested the classifiers with AC responses  
887 in single presentations of the non-trained sounds. Classification results were averaged across  
888 presentations to generate the categorization probability.

889

## 890 **QUANTIFICATION AND STATISTICAL ANALYSIS**

891 All quantification and statistical analysis were performed with custom Matlab or Python  
892 scripts. Statistical assessment was based on non-parametric tests reported in figure legends  
893 together with the mean and SEM values of the measurements, the number of samples used for  
894 the test and the nature of the sample (number of neurons, recording sessions or mice). A

895 custom bootstrap test (see Methods) was used for identification of optogenetic stimulation  
896 locations producing a behavioural effect. Unless otherwise mentioned, false positive rates  
897 below 0.05 were considered significant and the Benjamini-Hochberg correction for multiple  
898 testing was applied for repeated measurements in the experiments. In all analyses, all subjects  
899 which underwent a particular protocol in the study were included. For small groups, a  
900 minimum of 4 samples in each compared group was used to allow significance detection by  
901 standard non-parametric tests (e.g. Wilcoxon ranksum test).

902

### 903 **DATA AND SOFTWARE AVAILABILITY**

904 All data and analysis code are available from the corresponding author upon reasonable  
905 request.

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925



926 **KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Anti-NeuN alexa fluor 488 conjugated	Merk	MAB377X // RRID:AB_2149209
<b>Bacterial and Virus Strains</b>		
pAAV1.Syn.GCaMP6s.WPRE.SV40	Vector Core – University of Pennsylvania	Cat#: AV-1- PV2824 // RRID:Addgene_100843
<b>Biological Samples</b>		
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
<b>Critical Commercial Assays</b>		
<b>Deposited Data</b>		
<b>Experimental Models: Cell Lines</b>		
<b>Experimental Models: Organisms/Strains</b>		
PV-Cre Mouse B6;129P2-Pvalbtm1(cre)Arbr/J	The Jackson Laboratory	JAX: 008069 // RRID:IMSR_JAX:008069

Ai 27 Mouse: B6.Cg-Gt(ROSA)26Sortm27.1(CAG-COP4*H134R/tdTomato)Hze/J	The Jackson Laboratory	JAX: 012567 // RRID:IMSR_JAX:012567
Emx1-Cre Mouse B6.129S2-Emx1tm1(cre)Krl/J	The Jackson Laboratory	JAX: 005628 // RRID:IMSR_JAX:005628
Oligonucleotides		
Recombinant DNA		
Software and Algorithms		
AutoCell	Roland et al. 2017	<a href="https://github.com/thomasdeneux/autocell">https://github.com/thomasdeneux/autocell</a>
Klusta spike sorting	CortexLab (UCL)	<a href="https://github.com/kwikteam/klusta">https://github.com/kwikteam/klusta</a>
Other		

927

928

929 **References**

- 930 Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J.J., Natan, R.G., and Geffen, M.N.  
931 (2015). Bidirectional Regulation of Innate and Learned Behaviors That Rely on Frequency  
932 Discrimination by Cortical Inhibitory Neurons. *PLoS Biol* 13, e1002308.
- 933 Bandyopadhyay, S., Shamma, S.A., and Kanold, P.O. (2010). Dichotomy of functional  
934 organization in the mouse auditory cortex. *Nat Neurosci* 13, 361-368.
- 935 Bathellier, B., Ushakova, L., and Rumpel, S. (2012). Discrete neocortical dynamics predict  
936 behavioral categorization of sounds. *Neuron* 76, 435-449.
- 937 Choi, G.B., Stettler, D.D., Kallman, B.R., Bhaskar, S.T., Fleischmann, A., and Axel, R.  
938 (2011). Driving opposing behaviors with ensembles of piriform neurons. *Cell* 146, 1004-  
939 1015.
- 940 Cowey, A. (2010). Visual system: how does blindsight arise? *Curr Biol* 20, R702-704.
- 941 Dalmay, T., Abs, E., Poorthuis, R.B., Onasch, S., Signoret, J., Navarro, Y.L., Tovote, P.,  
942 Gjorgjieva, J., and Letzkus, J.J. A critical role for neocortical processing of fear memory. Co-  
943 submitted.
- 944 Deneux, T., Kempf, A., Daret, A., Ponsot, E., and Bathellier, B. (2016). Temporal  
945 asymmetries in auditory coding and perception reflect multi-layered nonlinearities. *Nature*  
946 *Communications* 7, 12682.
- 947 Dhawale, A.K., Hagiwara, A., Bhalla, U.S., Murthy, V.N., and Albeanu, D.F. (2010). Non-  
948 redundant odor coding by sister mitral cells revealed by light addressable glomeruli in the  
949 mouse. *Nat Neurosci* 13, 1404-1412.
- 950 Diamond, I.T., and Neff, W.D. (1957). Ablation of temporal cortex and discrimination of  
951 auditory patterns. *J Neurophysiol* 20, 300-315.
- 952 Francis, N.A., Winkowski, D.E., Sheikhattar, A., Armengol, K., Babadi, B., and Kanold, P.O.  
953 (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. *Neuron* 97,  
954 885-897 e886.
- 955 Fritz, J.B., David, S.V., Radtke-Schuller, S., Yin, P., and Shamma, S.A. (2010). Adaptive,  
956 behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal  
957 cortex. *Nat Neurosci* 13, 1011-1019.
- 958 Gimenez, T.L., Lorenc, M., and Jaramillo, S. (2015). Adaptive categorization of sound  
959 frequency does not require the auditory cortex in rats. *J Neurophysiol* 114, 1137-1145.
- 960 Guo, L., Ponvert, N.D., and Jaramillo, S. (2017). The role of sensory cortex in behavioral  
961 flexibility. *Neuroscience* 345, 3-11.
- 962 Guo, L., Walker, W.I., Ponvert, N.D., Penix, P.L., and Jaramillo, S. (2018). Stable  
963 representation of sounds in the posterior striatum during flexible auditory decisions. *Nat*  
964 *Commun* 9, 1534.
- 965 Guo, W., Hight, A.E., Chen, J.X., Klapoetke, N.C., Hancock, K.E., Shinn-Cunningham, B.G.,  
966 Boyden, E.S., Lee, D.J., and Polley, D.B. (2015). Hearing the light: neural and perceptual  
967 encoding of optogenetic stimulation in the central auditory pathway. *Sci Rep* 5, 10319.
- 968 Harrington, I.A., Heffner, R.S., and Heffner, H.E. (2001). An investigation of sensory deficits  
969 underlying the aphasia-like behavior of macaques with auditory cortex lesions. *Neuroreport*  
970 12, 1217-1221.

971 Harris, K.D., Henze, D.A., Csicsvari, J., Hirase, H., and Buzsaki, G. (2000). Accuracy of  
972 tetrode spike separation as determined by simultaneous intracellular and extracellular  
973 measurements. *J Neurophysiol* 84, 401-414.

974 Hong, Y.K., Lacefield, C.O., Rodgers, C.C., and Bruno, R.M. (2018). Sensation, movement  
975 and learning in the absence of barrel cortex. *Nature* 561, 542-546.

976 Honma, Y., Tsukano, H., Horie, M., Ohshima, S., Tohmi, M., Kubota, Y., Takahashi, K.,  
977 Hishida, R., Takahashi, S., and Shibuki, K. (2013). Auditory cortical areas activated by slow  
978 frequency-modulated sounds in mice. *PLoS One* 8, e68113.

979 Houweling, A.R., and Brecht, M. (2008). Behavioural report of single neuron stimulation in  
980 somatosensory cortex. *Nature* 451, 65-68.

981 Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromadka, T., Mainen, Z., and Svoboda, K.  
982 (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely  
983 moving mice. *Nature* 451, 61-64.

984 Issa, J.B., Haefele, B.D., Agarwal, A., Bergles, D.E., Young, E.D., and Yue, D.T. (2014).  
985 Multiscale optical Ca<sup>2+</sup> imaging of tonal organization in mouse auditory cortex. *Neuron* 83,  
986 944-959.

987 Issa, J.B., Haefele, B.D., Young, E.D., and Yue, D.T. (2017). Multiscale mapping of  
988 frequency sweep rate in mouse auditory cortex. *Hear Res* 344, 207-222.

989 Jaramillo, S., and Zador, A.M. (2011). The auditory cortex mediates the perceptual effects of  
990 acoustic temporal expectation. *Nat Neurosci* 14, 246-251.

991 Kalatsky, V.A., Polley, D.B., Merzenich, M.M., Schreiner, C.E., and Stryker, M.P. (2005).  
992 Fine functional organization of auditory cortex revealed by Fourier optical imaging. *Proc Natl*  
993 *Acad Sci U S A* 102, 13325-13330.

994 Kerlin, A.M., Andermann, M.L., Berezovskii, V.K., and Reid, R.C. (2010). Broadly tuned  
995 response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* 67,  
996 858-871.

997 Kuchibhotla, K., and Bathellier, B. (2018). Neural encoding of sensory and behavioral  
998 complexity in the auditory cortex. *Curr Opin Neurobiol* 52, 65-71.

999 Kuchibhotla, K.V., Gill, J.V., Lindsay, G.W., Papadoyannis, E.S., Field, R.E., Sten, T.A.,  
1000 Miller, K.D., and Froemke, R.C. (2017). Parallel processing by cortical inhibition enables  
1001 context-dependent behavior. *Nat Neurosci* 20, 62-71.

1002 Lashley, K.S. (1950). In search of the engram. Paper presented at: Physiological mechanisms  
1003 in animal behavior (Society's Symposium IV) (Academic Press).

1004 LeDoux, J.E., Farb, C.R., and Romanski, L.M. (1991). Overlapping projections to the  
1005 amygdala and striatum from auditory processing areas of the thalamus and cortex. *Neurosci*  
1006 *Lett* 134, 139-144.

1007 LeDoux, J.E., Sakaguchi, A., and Reis, D.J. (1984). Subcortical efferent projections of the  
1008 medial geniculate nucleus mediate emotional responses conditioned to acoustic stimuli. *J*  
1009 *Neurosci* 4, 683-698.

1010 Letzkus, J.J., Wolff, S.B., Meyer, E.M., Tovote, P., Courtin, J., Herry, C., and Luthi, A.  
1011 (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex.  
1012 *Nature* 480, 331-335.

1013 Manzur, H.E., Alvarez, J., Babul, C., and Maldonado, P.E. (2013). Synchronization across  
1014 sensory cortical areas by electrical microstimulation is sufficient for behavioral  
1015 discrimination. *Cereb Cortex* 23, 2976-2986.

1016 Musall, S., von der Behrens, W., Mayrhofer, J.M., Weber, B., Helmchen, F., and Haiss, F.  
1017 (2014). Tactile frequency discrimination is enhanced by circumventing neocortical adaptation.  
1018 *Nat Neurosci* 17, 1567-1573.

1019 Nelken, I., Bizley, J.K., Nodal, F.R., Ahmed, B., Schnupp, J.W., and King, A.J. (2004).  
1020 Large-scale organization of ferret auditory cortex revealed using continuous acquisition of  
1021 intrinsic optical signals. *J Neurophysiol* 92, 2574-2588.

1022 O'Connor, D.H., Clack, N.G., Huber, D., Komiyama, T., Myers, E.W., and Svoboda, K.  
1023 (2010). Vibrissa-based object localization in head-fixed mice. *J Neurosci* 30, 1947-1967.

1024 O'Connor, D.H., Hires, S.A., Guo, Z.V., Li, N., Yu, J., Sun, Q.Q., Huber, D., and Svoboda, K.  
1025 (2013). Neural coding during active somatosensation revealed using illusory touch. *Nat*  
1026 *Neurosci* 16, 958-965.

1027 Ohl, F.W., Wetzel, W., Wagner, T., Rech, A., and Scheich, H. (1999). Bilateral ablation of  
1028 auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but  
1029 not of pure tones. *Learn Mem* 6, 347-362.

1030 Otchy, T.M., Wolff, S.B., Rhee, J.Y., Pehlevan, C., Kawai, R., Kempf, A., Gobes, S.M., and  
1031 Olveczky, B.P. (2015). Acute off-target effects of neural circuit manipulations. *Nature* 528,  
1032 358-363.

1033 Packer, A.M., Russell, L.E., Dalglish, H.W., and Hausser, M. (2015). Simultaneous all-  
1034 optical manipulation and recording of neural circuit activity with cellular resolution in vivo.  
1035 *Nat Methods* 12, 140-146.

1036 Pai, S., Erlich, J.C., Kopec, C., and Brody, C.D. (2011). Minimal impairment in a rat model of  
1037 duration discrimination following excitotoxic lesions of primary auditory and prefrontal  
1038 cortices. *Front Syst Neurosci* 5, 74.

1039 Peng, Y., Gillis-Smith, S., Jin, H., Trankner, D., Ryba, N.J., and Zuker, C.S. (2015). Sweet  
1040 and bitter taste in the brain of awake behaving animals. *Nature* 527, 512-515.

1041 Poort, J., Khan, A.G., Pachitariu, M., Nemri, A., Orsolich, I., Krupic, J., Bauza, M., Sahani,  
1042 M., Keller, G.B., Mrsic-Flogel, T.D., *et al.* (2015). Learning Enhances Sensory and Multiple  
1043 Non-sensory Representations in Primary Visual Cortex. *Neuron* 86, 1478-1490.

1044 Roland, B., Deneux, T., Franks, K.M., Bathellier, B., and Fleischmann, A. (2017). Odor  
1045 identity coding by distributed ensembles of neurons in the mouse olfactory cortex. *Elife* 6.

1046 Rothschild, G., Nelken, I., and Mizrahi, A. (2010). Functional organization and population  
1047 dynamics in the mouse primary auditory cortex. *Nat Neurosci* 13, 353-360.

1048 Rybalko, N., Suta, D., Nwabueze-Ogbo, F., and Syka, J. (2006). Effect of auditory cortex  
1049 lesions on the discrimination of frequency-modulated tones in rats. *Eur J Neurosci* 23, 1614-  
1050 1622.

1051 Sachidhanandam, S., Sreenivasan, V., Kyriakatos, A., Kremer, Y., and Petersen, C.C. (2013).  
1052 Membrane potential correlates of sensory perception in mouse barrel cortex. *Nat Neurosci* 16,  
1053 1671-1677.

1054 Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation  
1055 influences perceptual judgements of motion direction. *Nature* 346, 174-177.

1056 Sanders, M.D., Warrington, E.K., Marshall, J., and Wieskrantz, L. (1974). "Blindsight":  
1057 Vision in a field defect. *Lancet* *1*, 707-708.

1058 Schmid, M.C., Mrowka, S.W., Turchi, J., Saunders, R.C., Wilke, M., Peters, A.J., Ye, F.Q.,  
1059 and Leopold, D.A. (2010). Blindsight depends on the lateral geniculate nucleus. *Nature* *466*,  
1060 373-377.

1061 Stein, B.E., and Stanford, T.R. (2008). Multisensory integration: current issues from the  
1062 perspective of the single neuron. *Nat Rev Neurosci* *9*, 255-266.

1063 Talwar, S.K., Musial, P.G., and Gerstein, G.L. (2001). Role of mammalian auditory cortex in  
1064 the perception of elementary sound properties. *J Neurophysiol* *85*, 2350-2358.

1065 Tye, K.M., Prakash, R., Kim, S.Y., Fenno, L.E., Grosenick, L., Zarabi, H., Thompson, K.R.,  
1066 Gradinaru, V., Ramakrishnan, C., and Deisseroth, K. (2011). Amygdala circuitry mediating  
1067 reversible and bidirectional control of anxiety. *Nature* *471*, 358-362.

1068 Wetzel, W., Ohl, F.W., and Scheich, H. (2008). Global versus local processing of frequency-  
1069 modulated tones in gerbils: an animal model of lateralized auditory cortex functions. *Proc*  
1070 *Natl Acad Sci U S A* *105*, 6753-6758.

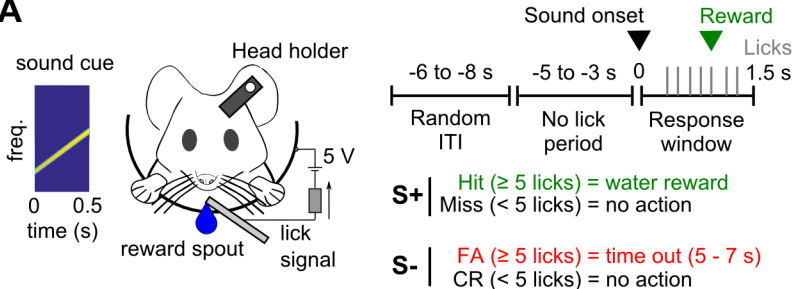
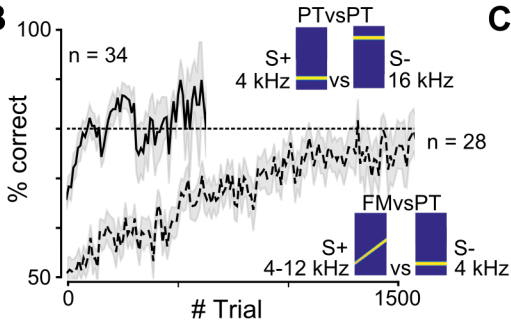
1071 Yang, Y., DeWeese, M.R., Otazu, G.H., and Zador, A.M. (2008). Millisecond-scale  
1072 differences in neural activity in auditory cortex can drive decisions. *Nat Neurosci* *11*, 1262-  
1073 1263.

1074 Zhu, P., Fajardo, O., Shum, J., Zhang Scharer, Y.P., and Friedrich, R.W. (2012). High-  
1075 resolution optical control of spatiotemporal neuronal activity patterns in zebrafish using a  
1076 digital micromirror device. *Nat Protoc* *7*, 1410-1425.

1077 Znamenskiy, P., and Zador, A.M. (2013). Corticostriatal neurons in auditory cortex drive  
1078 decisions during auditory discrimination. *Nature* *497*, 482-485.

1079

1080

**A****B****C**