Full title: Egocentric and Allocentric Representations in Auditory Cortex

Short title: Egocentric and Allocentric Representations in Auditory Cortex

Authors: Stephen M. Town[1], W. Owen Brimijoin[2] & Jennifer K. Bizley[1]

Affiliations:

(1) Ear Institute, University College London, London, UK

(2) MRC/CSO Institute of Hearing Research – Scottish Section, Glasgow, UK

Corresponding Author:

Stephen Town

s.town@ucl.ac.uk

0000-0003-1375-7769 (ORCID iD)

Author Contributions

S.M.T., W.O.B. and J.K.B. designed the experiment and wrote the paper; S.M.T. conducted the experiments and analysed the data.

Author Summary

When we hear a sound, we can describe its location relative to ourselves (e.g. “the phone is on my right”) or relative to the world (e.g. “the phone is in the corner”). These descriptions of space are known as egocentric and allocentric respectively and illustrate the representation of sound location in coordinate frames defined by the observer or the world in which the observer can move. In the brain, we know that neurons in static subjects can represent sound locations. However because sound location relative to the subject and the world are always fixed when the subject is static, it is impossible to tell if neurons represent egocentric or allocentric sound location. Here we recorded neurons in auditory cortex of freely moving ferrets and showed that most cells represent egocentric sound location. We also recorded a smaller number of allocentric cells that represented sound location in the world across movement. Additionally we used the movement of subjects to investigate the neural encoding of sound source distance and the modulation of auditory processing by the speed of head movement.

Abstract

A key function of the brain is to provide a stable representation of an object's location in the world. In hearing, sound azimuth and elevation are encoded by neurons throughout the auditory system and auditory cortex is necessary for sound localization. However the coordinate frame in which neurons represent sound space remains undefined: classical spatial receptive fields in head-fixed subjects can be explained either by sensitivity to sound source location relative to the head (egocentric) or relative to the world (allocentric encoding). This coordinate frame ambiguity can be resolved by studying freely moving subjects and here we recorded spatial receptive fields in the auditory cortex of freely moving ferrets. We found that most spatially tuned neurons represented sound source location relative to the head across changes in head position and direction. In addition, we also recorded a small number of neurons in which sound location was represented in a world-centered coordinate frame. We used measurements of spatial tuning across changes in head position and direction to explore the influence of sound source distance and speed of head movement on auditory cortical activity and spatial tuning. Modulation depth of spatial tuning increased with distance for egocentric but not allocentric units, whereas for both populations modulation was stronger at faster movement speeds. Our findings suggest that early auditory cortex primarily represents sound source location relative to ourselves but that a minority of cells can represent sound location in the world independent of our own position.

Introduction

A central role of the brain is to build a model of the world and objects within it that remains stable across changes in sensory input when we move. In hearing, this requires an observer maintains the identification of an auditory object as they move through an environment. Movement is a critical aspect of sensing [1] that contributes to sound localization and other auditory behaviors [2-7], however the neural basis underpinning active hearing and how the brain constructs world-centered sound location remains unknown.

For a moving observer, it is possible to represent sound location either relative to oneself (egocentric representation) or relative to the world through which one moves (allocentric representation). Allocentric representations provide a consistent report of object location across movement of an observer [8], as well as a common reference frame for mapping information across several observers or multiple sensory systems [9, 10]. Despite the computational value and perceptual relevance of allocentric representations to hearing, studies of auditory processing have only recently considered the coordinate frames in which sound location is represented [11-13]. Both electroencephalography (EEG) and modelling studies hint that sound location might be represented in cortex in both head-centered and head-independent spaces. However EEG has not yet revealed the precise location of these representations and cannot determine how individual neurons in tonotopic auditory cortex define space.

In static subjects, auditory cortical neurons encode sound azimuth and elevation [14-18] and localization of sound sources requires an intact auditory cortex [19-21]. However, in static subjects with a fixed head position, neural tuning to sound location is ambiguous, as the head and world coordinate frames are fixed in alignment and so allocentric and egocentric sound location are always equivalent. While it has been largely assumed cortical neurons represent sound location relative to the head, the spatial coordinate frame in which location is encoded remains to be demonstrated.

Furthermore, though the acoustic cues to sound localisation are explicitly head-centered, information about head direction necessary to form a world-centered representation is present at early levels of the ascending auditory system [22]. Thus it may be possible for neurons in the auditory system to represent space in an allocentric, world-centered coordinate frame that would preserve sound location across changes in head position and direction.

Here we resolve the coordinate frame ambiguity of spatial tuning in auditory cortex by recording from neurons in freely moving subjects. In moving conditions, the head and world coordinate frames are no longer fixed and so we can determine in which coordinate frame a given cell is most sensitive. Our approach reveals head-centered and world-centered units that suggest the coexistence of egocentric and allocentric representations in auditory cortex. We also explore the impact of distance from a sound source and the speed of subject's movement on spatial tuning in auditory cortex.

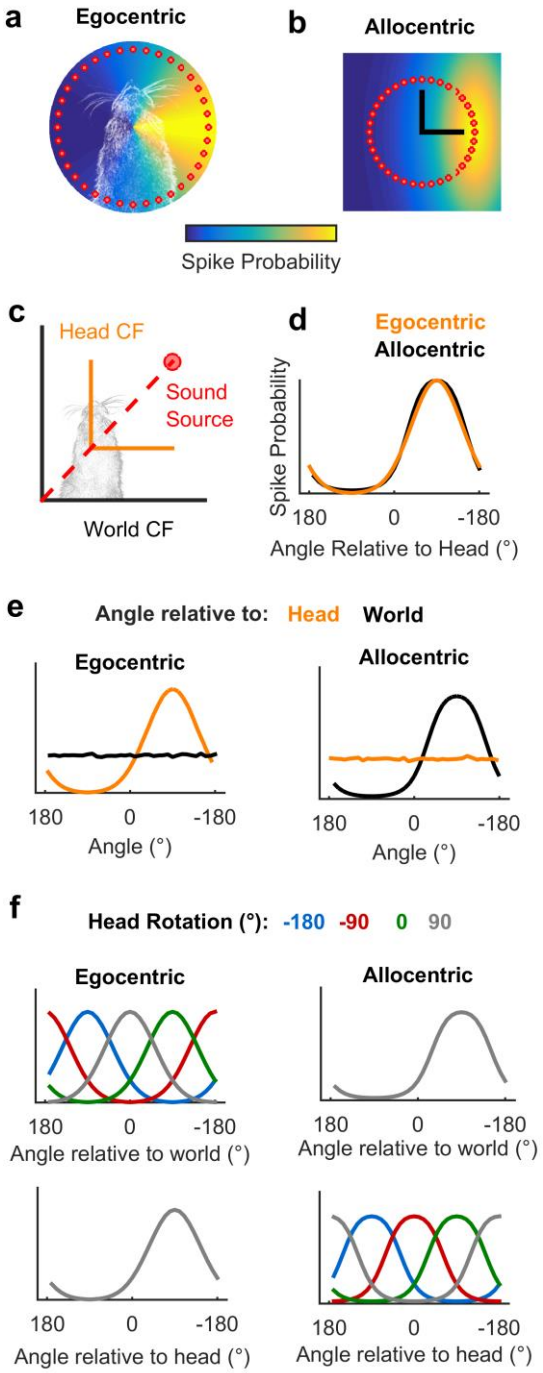
Results

We hypothesised that measuring spatial tuning in moving subjects would allow us to distinguish between egocentric (head-centered) and allocentric (world-centered) representations of sound location (Fig 1). To formalize this theory and develop quantitative predictions about the effects of observer movement on spatial tuning, we first simulated egocentric and allocentric neurons that were tuned to sound locations defined relative to the head (Fig 1a) and world (independent of the subject) respectively (Fig 1b). We simulated allocentric and egocentric units using parameters fitted to produce identical spatial receptive fields when tested in the classical condition in which the head is in a fixed location at the center of a speaker ring (Fig 1c-d), illustrating the coordinate frame ambiguity. However our simulation confirmed that when the observer moved freely with a uniform distribution of head directions, spatial tuning would only be apparent in the coordinate frame relevant for neural output (Fig 1e). Additionally, changes in head direction would

produce systematic shifts in tuning curves in the coordinate frame that was irrelevant for neural output while tuning in the relevant coordinate frame would be invariant across head direction (Fig 1f). We subsequently demonstrated that tuning curves of many shapes and preferred locations can theoretically be explained by spatial receptive fields based within an allocentric coordinate frame (Supplementary Fig. 1). With simulations providing a foundation, we then made recordings in freely moving animals to determine whether the spatial tuning of auditory cortical neurons followed egocentric or allocentric predictions.

Fig 1 Simulated receptive fields show that observer movement resolves coordinate frame ambiguity

a-b, Simulated neurons with receptive fields tuned to sound location relative to the head (a, Egocentric) or in the world (b, Allocentric). Circles show hypothetical sound sources in a classical speaker ring; black lines indicate axes and origin of the simulated world. **c**, Schematic of world and head coordinate frames (CF). **d**, Sound-evoked tuning curves according to allocentric and egocentric hypotheses when head and world coordinate frames were aligned. **e-f**, Predictions of allocentric and egocentric hypotheses showing mean spike probability averaged across uniform distributions of head rotation and position (e) and at specific head directions (f).



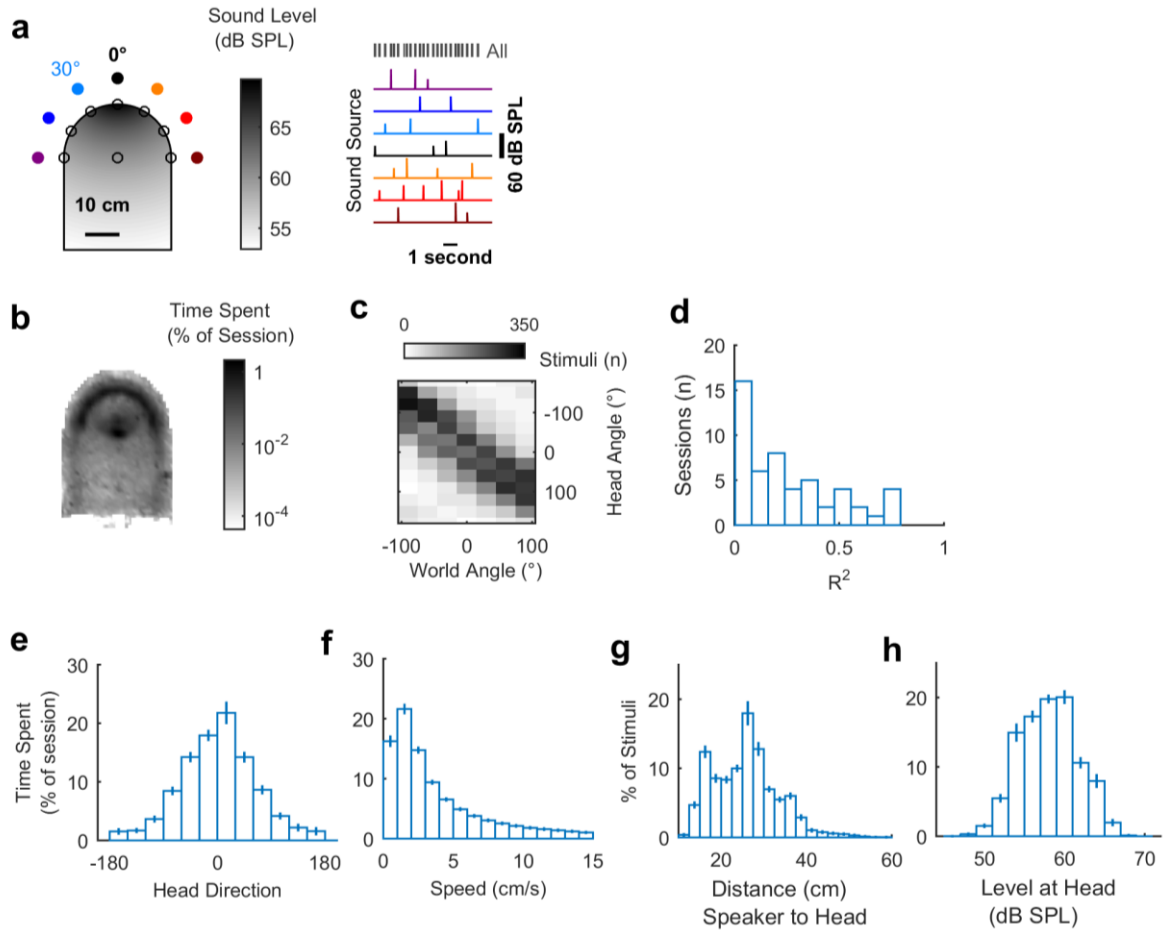
114

115 To measure spatial tuning in moving subjects, we implanted ferrets (n = 5) with multi-
116 channel tungsten electrode arrays allowing the recording of single and multi-unit activity during
117 behavior. During neural recording each ferret was placed in an arena, which the animal explored for
118 water rewards while the surrounding speakers played click sounds (Fig 2a). To measure the animal's
119 head position, direction and speed in the world during exploration (Fig. 2b-f) we tracked LEDs placed

on the midline of the head (Supplementary Video 1). During exploration, click sounds were presented from speakers arranged at 30° intervals between ± 90° relative to the arena center with speaker order and inter-click interval (250 – 500 ms) varied pseudo-randomly. We also roved the level of clicks between 54 and 60 dB SPL such that absolute sound level varied both as a function of sound source level and distance between head and speaker, to reduce cues about sound location provided by absolute sound level (Fig 2g-h). Clicks were used as they provided instantaneous energy and thus ensured minimal movement of the animal during stimulus presentation (Supplementary Fig. 3c-d). The locations (speaker angle) from which the clicks originated were used alone to estimate *allocentric* receptive fields and were used in conjunction with the animal's head direction and position to measure *egocentric* spatial receptive fields.

Fig 2 Experimental design and exploratory behavior in a sound field

a, Arena with speakers (filled circles) and water ports (unfilled circles). Shading indicates the sound field generated by a click from the speaker at 0° calibrated to be 60 dB SPL at the center of the chamber. Stimuli were presented with a pseudorandom interval and order across speakers. **b**, Mean proportion of time in each recording session spent within the arena. **c**, Stimulus angles relative to the head and world for one session that was representative of behavior in all sessions (n = 57). **d,** Correlation coefficients (R^2) between sound angles in head and world coordinate frames across all behavioral sessions. **e-h**, Distributions of head direction, head speed, distance between head and sound source and the sound level at the animal's head during behavior. Bars indicate mean ± s.e.m. across sessions.



140

---

141 We observed that animals moved throughout the arena to collect water (Fig 2b) and used a
142 range of head directions during exploration (Fig 2e). In contrast to our initial simulations, the
143 distribution of the animal's head direction was notably non-uniform, leading to correlations between
144 sound source angle relative to the head and the world (e.g. Fig 2c; mean ± s.e.m. $R^2 = 0.247 \pm$
145 0.0311). This correlation between sound source angles resulted because the animal preferred to
146 orient towards the front of the arena (0˚) and thus sounds that were to the right of the animal were
147 more often on the right of the arena than would result from random behavior. The preference of the
148 animal was likely a consequence of the shape of the arena, and the location of the water spouts
149 within it. Although the correlation between sound source locations relative to the head and within
150 the world was relatively small, we sought to determine how the animal's head direction preference
151 affected our experimental predictions.

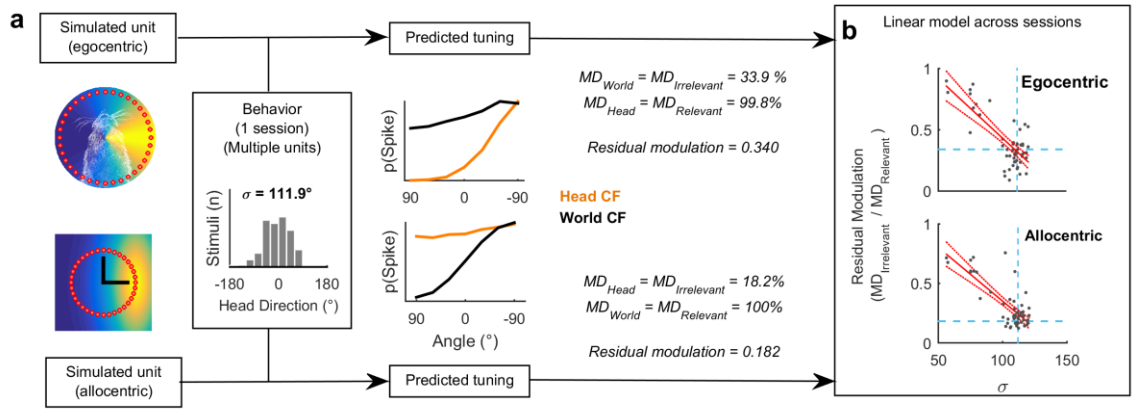
To assess the influence of real animal behavior on our ability to distinguish coordinate frames, we combined our simulated egocentric and allocentric receptive fields (Fig 1) with the animal's head position and direction across each single behavioural testing session (Fig 3a). This allowed us to calculate the spatial tuning for known allocentric and egocentric receptive fields in both head and world coordinate frames. Our simulation predictions (Fig.1) demonstrated that for a uniform distribution of head angles the tuning function of allocentric or egocentric units should be flat when considered in the irrelevant coordinate frame. However, a bias in head location over time would produce spatial modulation in firing rate with location in the irrelevant coordinate frame (Fig 3a). In order to account for this we therefore measured residual modulation as the ratio of modulation depth in each coordinate frame (Fig. 3; $MD_{Irrelevant}$ / $MD_{Relevant}$). Residual modulation thus represents the degree of indirect spatial tuning in one coordinate frame observed as a by-product of the animal's behavior combined with spatial tuning in another coordinate frame.

Across all behavioral sessions, residual modulation was inversely correlated with variation in the animal's head direction (expressed as standard deviation) for both egocentric ($R^2 = 0.562$, $p = 1.07 \times 10^{-10}$) and allocentric simulated units ($R^2 = 0.615$, $p = 3.73 \times 10^{-12}$) (Fig 3b). This indicated that for real animal behavior, we would not expect to see the complete abolition of tuning but rather spatial tuning in both coordinate frames, with the weaker tuning potentially attributable to the animal's bias in head direction. In our neural analysis, we thus used the relationship between behavior and residual modulation to provide a statistical framework in which to assess the significance of spatial tuning of real neurons.

Fig 3 Estimating residual modulation

a, Example workflow for estimating residual modulation in coordinate frames irrelevant for neural output that result from biases in head direction. Residual modulation was defined as: ($MD_{Irrelevant}$ / $MD_{Relevant}$). Estimations performed separately using simulated units for each behavioral session. **b** Residual modulation was inversely correlated with standard deviation of head directions (σ). Red

filled lines indicate regression fit and confidence intervals. Dashed lines indicate data point for the single session in (a).



Egocentric and allocentric tuning in auditory cortex

During exploration we recorded the activity of 186 sound-responsive units (50 single units, 136 multi-units) in auditory cortex (Supplementary Fig. 2). Electrode arrays were targeted to span the low frequency areas where the middle and posterior ectosylvian gyral regions meet and thus units were sampled from primary auditory cortex and two tonotopically organised secondary fields: the posterior pseudosylvian and posterior suprasylvian fields. We analysed the firing rates of units in the 50 milliseconds after the onset of each click; this window was wide enough to capture the neural response while being sufficiently short that the animal's head moved less than 1 cm (median 4 mm, Supplementary Fig. 3) and less than 30° (median 12.6°) – the interval between speakers. The time interval between stimuli always exceeded 250 ms.

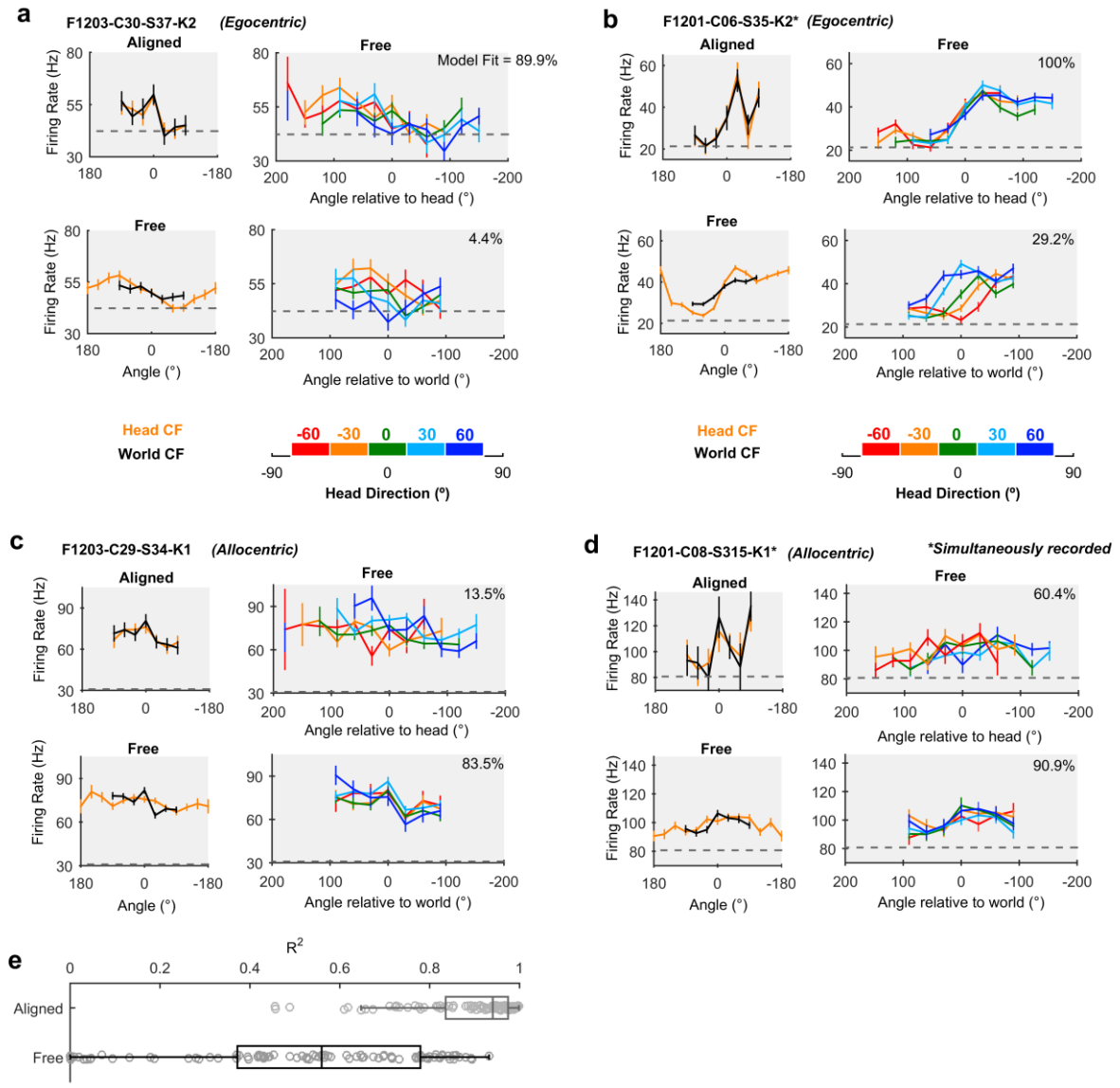
We identified periods of time when the animal was facing forwards (± 15° of the arena midline) at the center of the speaker ring (Supplementary Fig. 4): in this situation we mimic classic neurophysiological investigations of spatial tuning in which head and world coordinate frames are aligned. In the aligned case, we recorded 92 units that were significantly modulated by sound source location (Fig 4a-d: *Top left,* GLM analysis of deviance, $p \leq 0.05$) and for which spatial tuning curves computed in head and world coordinate frames were highly correlated (mean ± s.e.m: R^2 = 0.889 ±

196 0.0131). We then compared the aligned control condition with all data when the head and world
197 coordinate frames were free to vary. Compared to the aligned condition, correlation between
198 egocentric and allocentric tuning curves when coordinate frames were free to vary was significantly
199 reduced (Fig 4e, $R^2 = 0.522 \pm 0.0294$) (paired t-test: free vs. aligned $t_{91} = 8.76$, $p < 0.001$) and
200 differences in spatial tuning in head and world coordinate frames became visible (Fig 4a-d: *Bottom*
201 *left*).

---

202 **Fig 4 Spatial tuning of egocentric and allocentric units**

203 **a-d**, Spatial tuning of four example units that were classified as egocentric (a-b) or allocentric (c-d).
204 In each panel, *top left*: Tuning curves calculated for sound angle in head and world coordinate frames
205 (CF) when both frames were aligned. *Bottom left*: Tuning curve when head and world CFs were free
206 to vary. *Top and bottom right*: Tuning curves plotted at specific head rotations. Model fit refers to
207 percentage of explainable deviance calculated according to Figure 6 across all data in which
208 coordinate frames were free to vary. Data for all tuning curves are shown as mean ± s.e.m. firing
209 rates. Dotted lines show the mean background activity measured in the 50 ms before stimulus
210 presentation. **e,** Correlation coefficients between tuning curves in head and world CFs when aligned
211 or free to vary as the animal foraged around the arena. Boxplots show median and inter-quartile range;
212 symbols show coefficients for individual units.



213

---

214 When animals moved freely through the arena, and head and world coordinate frames were
215 thus dissociated, we observed units consistent with egocentric (Fig 4a-b and Supplementary Fig. 5)
216 and allocentric hypotheses (Fig 4c-d and Supplementary Fig. 6). For units consistent with the
217 egocentric hypothesis, spatial receptive fields were more strongly modulated by sound angle in the
218 head than world coordinate frame. For the unit shown in figure 4a, modulation depth values in the
219 head and world coordinate frames were 28.3% and 10.1% respectively. In figure 4b, modulation
220 depth was 49.0% in the head coordinate frame and 30.3% in the world coordinate frame.
221 Furthermore tuning curves for sounds plotted relative to the head remained consistent across head

rotation but shifted systematically when plotted relative to the world (Fig 4a-b: *Right columns*). Both outcomes are highly consistent with our simulation predictions (Fig 1).

In addition to identifying head-centered spatial tuning across movement, we also found units with spatial tuning that realized the predictions generated by the allocentric hypothesis. These units showed greater modulation depth to sound angle in the world coordinate frame than head coordinate frame (Fig 4c-d and Supplementary Fig. 6): For putative allocentric units, modulation depths for tuning curves were 21.2% and 13.4% in the world and head coordinate frames respectively for the unit shown in Fig. 4c, and 12.7% and 10.1% respectively for the unit shown in Fig. 4d. While these modulation depth values for these units were relatively low (possibly due to the high background activity), their tuning curves were consistent with representations of world-centered sound location. Specifically for allocentric units, spatial tuning in the world coordinate frame was robust to head rotation whereas tuning curves expressed relative to the head were systematically shifted when mapped according to head direction (Fig 4c-d: *Right column*).

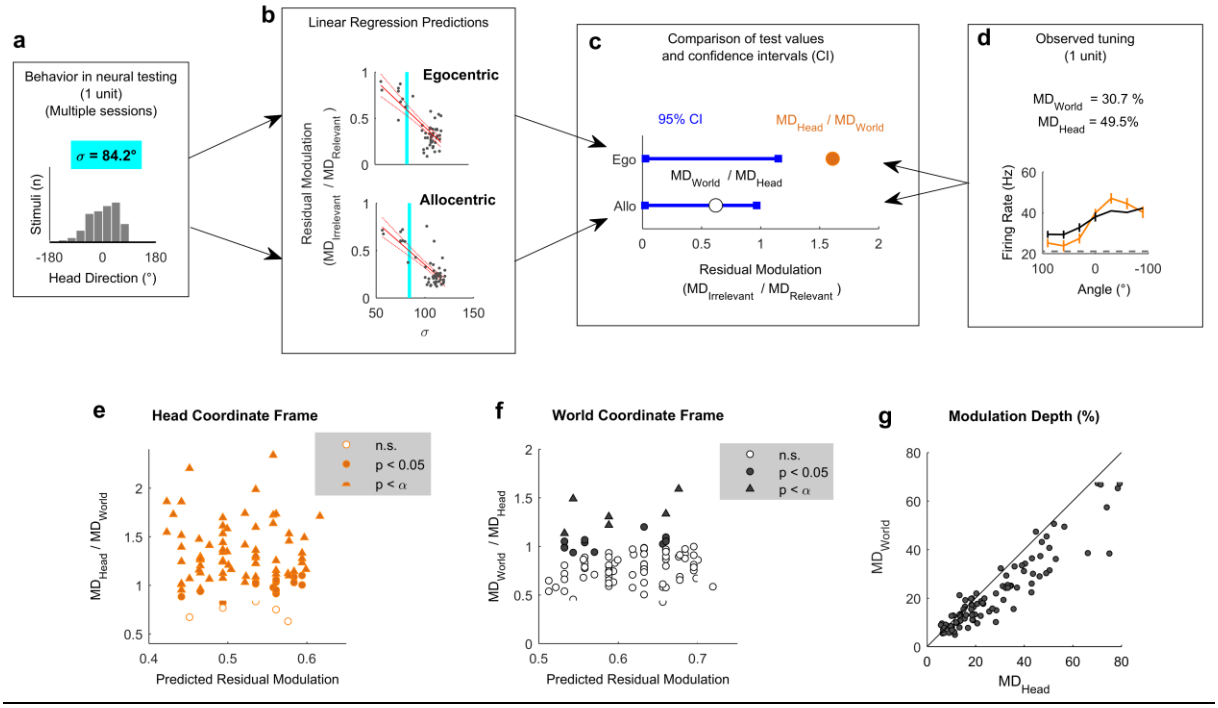
Modulation depth across coordinate frames

To quantify the observations we made above and systematically compare spatial tuning in world and head coordinate frames, we calculated modulation depth for both tuning curves for each unit. We next asked if modulation depth observed in either head or world coordinate frames was greater than the residual modulation predicted by our earlier simulations (Fig.3b). A linear regression model developed using simulated receptive fields was used to predict the magnitude of residual tuning for each coordinate frame from the animal's behavior during the recording of each unit (Fig 5a-d). To describe the animal's behavior across the relevant testing sessions for each neural recording, we calculated the standard deviation of head directions (Fig. 5a). A smaller standard deviation indicates a less uniform range of head-directions and when combined with our regression model (Fig 5b) would predict higher residual modulation in both coordinate frames. Thus for a given standard deviation, we could use linear regression to obtain a predicted confidence interval for the

residual modulation in head and world coordinate frames arising from allocentric or egocentric tuning respectively (Fig. 5c). The observed modulation values were calculated for each unit as the ratio of modulation depth in one coordinate frame divided by the other coordinate frame (Fig. 5d) and significance was attributed when test values exceeded the confidence interval of residual modulation.

Fig 5 Modulation depth across coordinate frames

a-d Workflow illustrating the use of animal behavior (a: summarized using the standard deviation of head directions during neural testing, σ) and linear regression models (b – see also Fig 3) to generate confidence intervals for residual modulation (c) that were compared to observed modulation depth values (d) normalized relative to the alternative coordinate frame. Blue vertical lines in b show the σ value in a. **e-f,** Normalized modulation depth observed for each spatially tuned unit compared against the mean residual modulation predicted from behavior in head (e) or world (f) coordinate frames. Bonferroni-corrected statistical criterion (p = 5.43 x 10^{-4}) is denoted by α. **g**, Modulation depth for all spatially modulated units (n = 92) compared in world (MD_{World}) and head coordinate frames (MD_{Head}) during exploration.



Across all spatially tuned units, modulation in the head coordinate frame was significantly greater than predicted for 87 / 92 units (94.6%)($p < 0.05$, Fig 5e); modulation in the world coordinate frame was significant for 19 / 92 units (20.7%, Fig 5f). For 14 / 92 units (15.2%), modulation depth was significantly greater than expected in both coordinate frames. When Bonferroni corrected for multiple ($n = 92$, $\alpha = 5.43 \times 10^{-4}$) comparisons these numbers dropped to 69 / 92 units (75%) for modulation in the head coordinate frame, none of which were additionally modulated in a world coordinate frame, and 6 / 92 units (6.5%) for modulation in the world coordinate frame – none of which showed significant head-centered modulation. With this more conservative statistical threshold, modulation depths were not significantly greater than expected in either coordinate frame for the remaining 17 / 92 units (18.5%). Together these observations suggest that response types occupy a continuum from purely egocentric to purely allocentric. The existence of units with significant modulation in both coordinate frames with a less conservative statistical cut-off, or no significant modulation with corrected threshold may indicate mixed spatial sensitivity comparable with other reports in auditory cortex [23].

A key prediction from our simulations with both uniform head-directions (Fig.1) and actual head-directions (Fig 3a) was that modulation depth would be greater in the coordinate frame that was relevant for neural activity than the irrelevant coordinate frame (i.e. Head > World for egocentric; World > Head for allocentric). Having demonstrated that modulation within both co-ordinate frames was greater than expected based on the non-uniform sampling of head direction, we compared the modulation depth across coordinate frames for all spatially tuned units (Fig 5g). For 76 / 92 units (82.6%), we observed greater modulation depth in the head than world coordinate frame indicating a predominance of egocentric tuning and a minority of units (16 / 92, 17.4%) in which allocentric tuning was strongest.

General Linear Modelling to define egocentric and allocentric populations

Our analysis of modulation depth indicated the presence of both egocentric and allocentric representations in auditory cortex but also highlighted that the analysis of modulation depth alone was sometimes unable to resolve the coordinate frame in which units encoded sound location. To calculate modulation depth requires we discretize sound location into distinct angular bins, average neural responses across trials and thus ignore single trial variation in firing rates. General linear models (GLMs) potentially offer a more sensitive method as they permit the analysis of single trial data and allow us to treat sound angle as a continuous variable. We considered two models which either characterized neural activity as a function of sound source angle relative to the head (GLM_{HEAD}) or in the world (GLM_{WORLD}). For all units for which at least one GLM provided a statistically significant fit (relative to a constant model, analysis of deviance; $p < 0.05$, 91 / 92 units), we compared model performance using the Akaike information criterion (AIC)[24] for model selection. In accordance with the modulation depth analysis, the majority of units were better modelled by sound angle relative to the head than world (72/91 units; 79.1%; four animals) consistent with egocentric tuning. However, we also observed a smaller number of units (19/91 units; 20.9%; three animals) whose responses were better modelled by sound angle in the world and thus showed stronger representation of allocentric sound location.

To visualize GLM performance and explore egocentric and allocentric tuning further, we plotted a normalized metric of the deviance value usually used to assess model fit. Here we defined *model fit* as the proportion of explainable deviance (Fig 6a) where a test model (e.g. GLM_{WORLD}) is considered in the context of GLMs that have no variable predictors of neural activity (a constant model) or use sound angle in both coordinate frames as predictors (a full model). This normalization step is critical in comparing model fit across units as deviance values alone are unbounded. In contrast, model fit is limited from 0 (indicating the sound angle provides little information about the neuron's response) to 100% (indicating the sound angle in one coordinate frame accounts for the

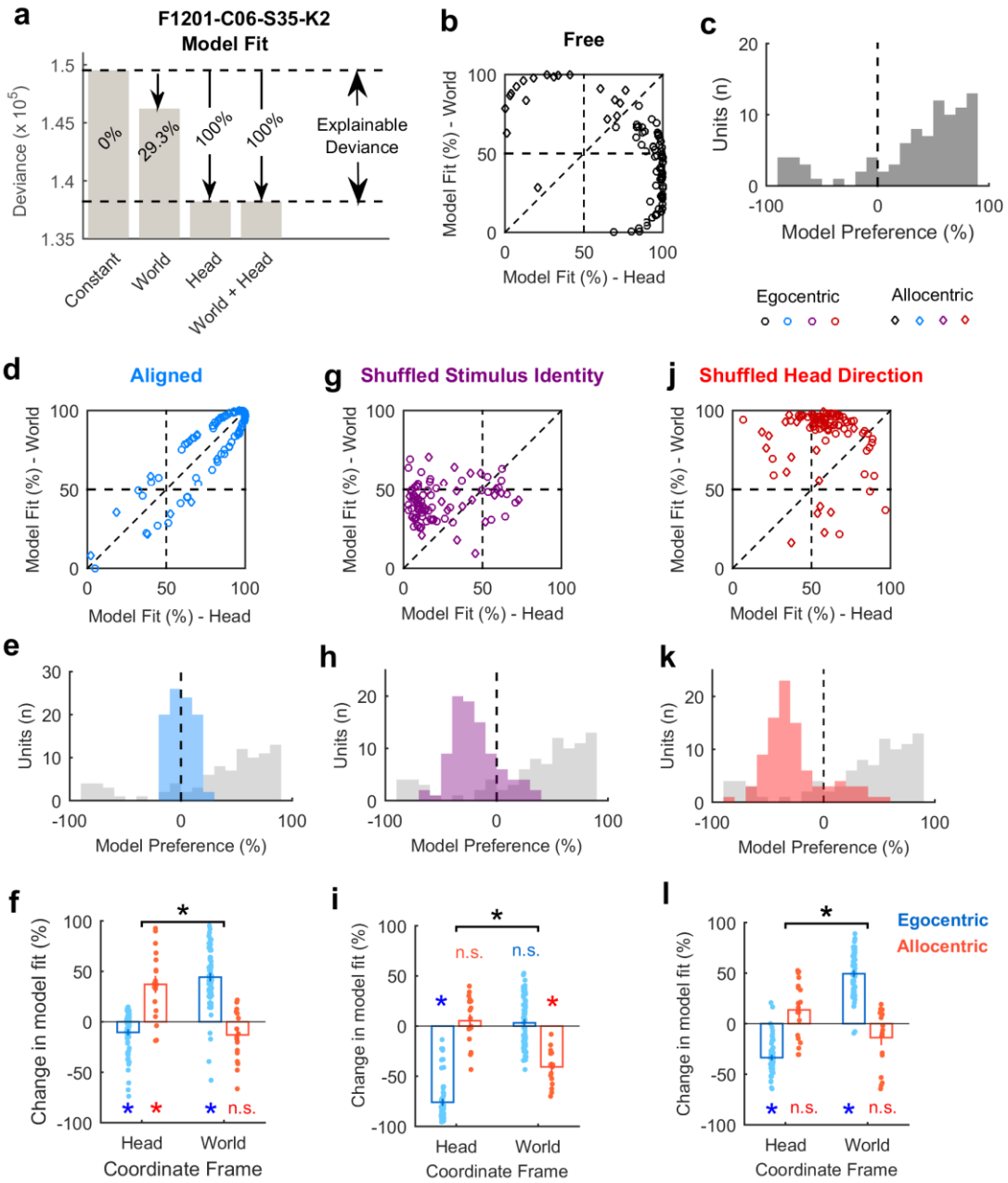
neuron's response as well as sound angles in both frames). While the units we recorded formed a continuum in this space, for the purpose of further analysis we defined egocentric and allocentric units according to the coordinate frame (head / world) that provided the best model fit as determined by the AIC above.

Using a GLM based analysis, we predicted that egocentric units would have a high percentage of the full model fit by sound angles relative to the head, and a low model fit for sound angles relative to the world, and that allocentric units would show the reciprocal relationship. To test these predictions we generated a *model preference* score; the model fit for sound angles relative to the head minus the model fit for sound angles in the world. Accordingly, negative values of model preference should identify allocentric units while positive values should indicate egocentric units. Neurons in which both sound angles relative to the world and head provide high model fit values may represent sounds in intermediary or mixed coordinate frames and would have model preference scores close to zero, as would neurons in which we were unable to disambiguate coordinate frame preference due to non-uniform head angle distributions.

Fig 6 General linear modelling of spatial sensitivity

a, Calculation of model fit for sound angle relative to the head or world. Raw deviance values are normalized as a proportion of explainable deviance; the change in deviance between a constant and a full model. **b**, Model fit comparisons for all units when the animal was free to move through the arena. **c**, Model preference that indicates the distribution of units across the diagonal line of equality in (b). **d-e,** Model fit and model preference for data when the head and world coordinate frames were aligned. Grey background in (e) shows the distribution of model preference in the freely varying condition for reference. **f,** Change in *model fit* between freely moving and aligned states for egocentric and allocentric units in head and world coordinate frames. **g-h,** Model fit and model preference for freely moving data when speaker identity was shuffled. Data points in (g) show for each unit the median model fit averaged across 1000 shuffles. **i**, Change in model fit between unshuffled and

shuffled data. **j-k,** Model fit and model preference for freely moving data when the animal's head direction was shuffled. Data points in (j) show for each unit median model fit averaged across 1000 shuffles. **l**, Change in model fit between unshuffled and shuffled data. Asterisks with horizontal bars in (f), (i) and (l) indicate significant interactions (two-way anova on change in model fit with shuffle) between coordinate frame (head / world) and unit type (egocentric / allocentric) ($p < 0.001$). Asterisks / n.s. for each bar represent significant / non-significant effects of shuffle on model fit in specific coordinate frames and for specific unity types (red/blue; t-test, $p < 0.05$).



In the space defined by model fit for sound angles relative to the head and world (Fig 6b), units clustered in opposite areas supporting the existence of both egocentric and allocentric representations. This clustering was also evident in the model preference scores, which showed a bimodal distribution (Fig 6c). Repeating this analysis on data in which the head and world coordinate frames were aligned (due to the animals position at the center of the speaker ring) demonstrated that model fit values for head and world coordinate frames became more similar and model preference scores were centered around zero (Fig 6d-e). When we compared the change in model fit with alignment (two-way anova), this was reflected as a significant interaction between coordinate frame (head or world) and unit type (egocentric or allocentric, determined by the coordinate frame that provided best model fit using the AIC)($F_{1,178}$ = 130.1, p = 5.71 x 10^{-23}). Post-hoc comparison confirmed that model fit in the head coordinate frame decreased significantly for egocentric units (Bonferroni corrected for multiple comparisons, p = 5.04 x 10^{-5}) and increased significantly for allocentric units (p = 4.36 x 10^{-4}). In contrast in the world coordinate frame, alignment led to a significant increase in model fit for egocentric units (p = 2.22 x 10^{-20}) and a non-significant decrease in model fit for allocentric units (p = 0.155).

We performed two additional control analyses on the freely moving dataset: firstly, we randomly shuffled the speaker identity while maintaining the same information about the animal's head direction. Randomising the speaker identity should affect the ability to model both egocentric and allocentric neurons and we would therefore predict that model fits for spatially relevant coordinate frames would be worse and model preference scores would tend to zero. (I.e. shuffling would shift model preference scores in the negative direction for egocentric units and the positive direction for allocentric units). As predicted, shuffling speaker identity eliminated clustering of egocentric and allocentric units in the space defined by model fit (Fig 6g) and lead to opposing effects on model preference (Fig. 6h-i): Model fit scores for egocentric and allocentric units were both affected by shuffling speaker identity but in opposite directions (unit x coordinate frame interaction, $F_{1, 178}$ = 227.4, p = 1.22 x 10^{-33}): For egocentric units, the model fit for sound angle

relative to the head declined significantly with shuffling (p = 7.44 x 10^{-41}) while fit for sound angle in the world did not change significantly (p = 0.271). For allocentric units, model fit for sound angle in the world declined significantly (p = 1.29 x 10^{-8}) but was not significantly different in the head coordinate frame (p = 0.35). When shuffling stimulus angle (averaging across 1000 shuffles), we identified 12 / 19 (63.2%) allocentric and 63 / 72 (87.5%) egocentric units with model preference scores beyond the 97.5% percentile limits of the shuffled distribution.

Secondly, we shuffled information about the animal's head direction while maintaining information about speaker identity. This should cause model fit values to decline for sound angle relative to the head for egocentric units and should therefore result in egocentric units shifting their model preference scores towards zero. For allocentric units the model fit for sound location in the world should be maintained, and we would not predict a systematic change in model preference. This was indeed the case (Fig 6j-l; interaction between coordinate frame and unit type: $F_{1, 178}$ = 216.7, p = 1.35 x 10^{-32}): For egocentric units shuffling head direction significantly reduced model fit in the head coordinate frame (p = 1.40 x 10^{-27}) and increased model fit in the world coordinate frame (p = 2.20 x 10^{-31}). For allocentric units, shuffling head direction did not significantly affect model fit in either head (p = 0.211) or world coordinate frames (p = 0.178).

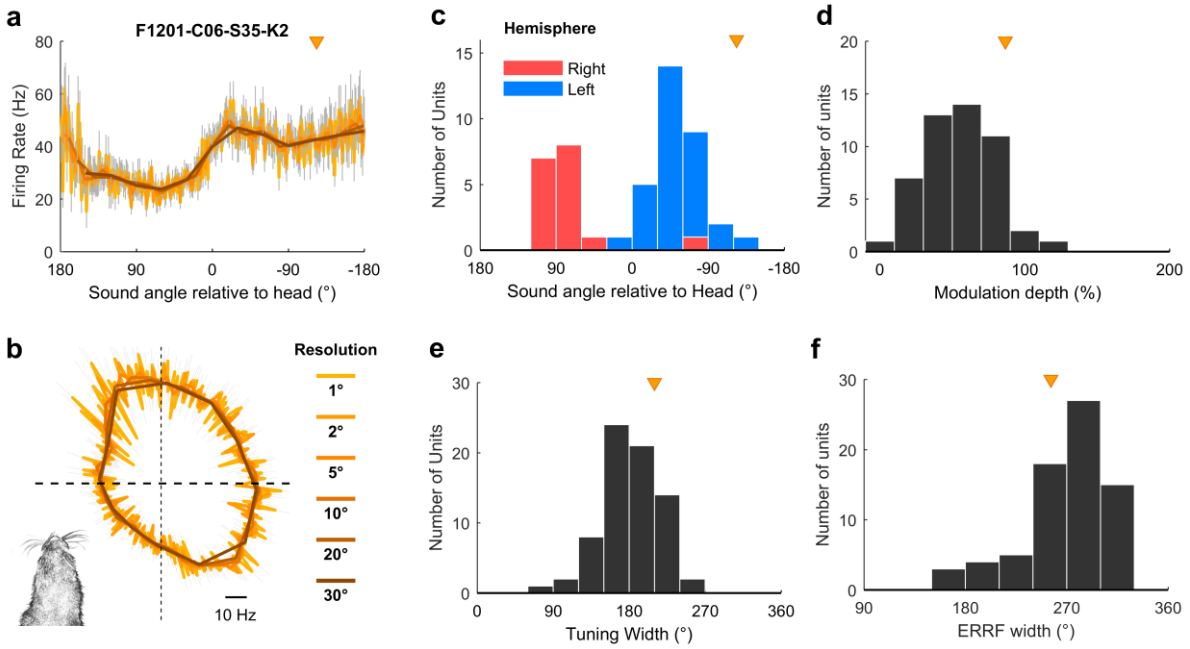
Egocentric and allocentric units – population characteristics

For egocentric units that encoded sound location in the head coordinate frame, it was possible to characterize the full extent of tuning curves in 360° around the head (Fig 5a-b) despite our speaker array only spanning 180°. This was possible because the animal's head direction varied continuously across 360° and so removed the constraints on measurement of spatial tuning imposed by the range of speaker angles used. Indeed, we were able to extend our approach further to characterize super-resolution tuning functions with an angular resolution more precise than the interval between speakers (Fig 7a-b and Supplementary Fig 7). Together these findings show it is possible to use changes in head direction to recover the spatial tuning of individual units with

greater detail than would be possible if the subject's head position and direction remained constant relative to the sound sources. Egocentric units shared spatial receptive field properties typical of previous studies [15, 16, 18, 25]: units predominantly responded most strongly to contralateral space (Fig 7c) with broad tuning width (Fig 7e-f) that typifies auditory cortical neurons. We also found similar, if slightly weaker spatial modulation when calculating modulation depth according to Ref. [18] (Fig 7d).

Fig 7 Egocentric unit characteristics

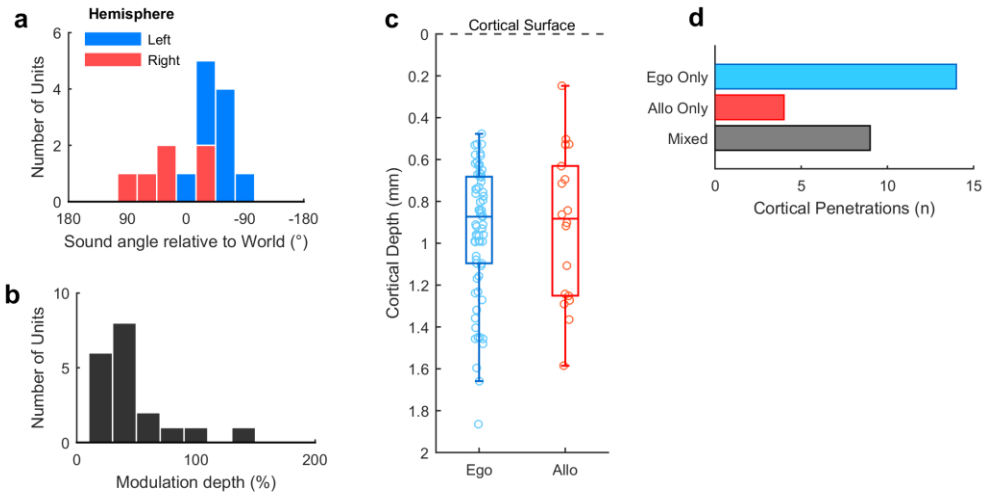
a-b, Spatial tuning of example egocentric unit at multiple angular resolutions. Data shown as mean ± s.e.m. firing rate plotted in Cartesian (a) or polar (b) coordinates. Triangle indicates preferred location of unit. Inset (b) illustrates the corresponding head direction onto which spatial tuning can be super-imposed. **c**, Preferred location of all egocentric units (n = 72) in left and right auditory cortex. **d**, Modulation depth calculated according to [18] across 360° for units in both hemispheres. **e-f**, Tuning width (e) and equivalent rectangular receptive field width (ERRF)(f) for all units. Triangle indicates the preferred location, modulation depth, tuning width and ERRF of the example unit in (a).



For allocentric units, we observed a similar contralateral tuning bias in preferred location (Fig 8a) to egocentric units and that allocentric units had relatively low modulation depths (Fig 8b). These tuning features may not be surprising as an allocentric receptive field could presumably fall anywhere within or beyond the arena, and might therefore be poorly sampled by circular speaker arrangements. If the tuning curves measured here were in fact sampling a more complex receptive field that related to a world-centered coordinate frame (Fig 1b) we would predict that the receptive fields would be correspondingly noisier. Allocentric and egocentric units were recorded at similar cortical depths (Fig 8c) and on the same cortical penetrations as egocentric units were recorded on 9 of 13 electrodes (69.2%) on which we also identified allocentric units (Fig 8d).

Fig 8 Allocentric unit characteristics

a, Preferred location of all allocentric units (n = 19) in left and right auditory cortex. **b**, Modulation depth calculated across 180° for units in both hemispheres. **c,** Comparison of cortical depth at which egocentric and allocentric units were recorded. Ferret auditory cortex varies in thickness between 1.5 and 2 mm and electrode depths were confirmed histologically (Supplementary Fig. 2). **d**, Number of cortical penetrations on which we recorded only egocentric units, only allocentric units or a combination of both (mixed). All 92 spatially tuned units were recorded on 27 unique electrodes, with recorded units being distributed throughout cortex as the electrodes were descended.

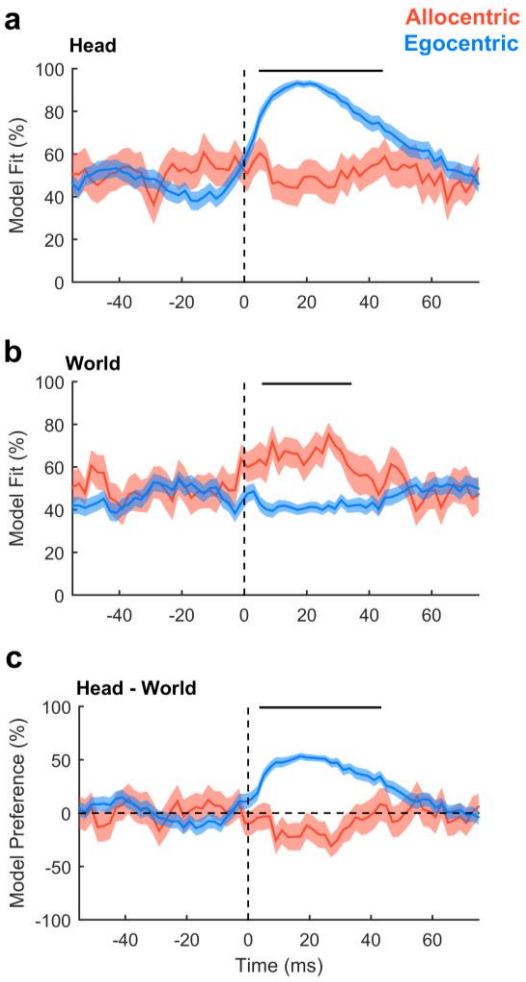


Timing of spatial information

Our findings suggested the existence of egocentric and allocentric tuning in auditory cortex. As these descriptions were functionally defined, we hypothesized that differences between allocentric and egocentric tuning should only arise after stimulus presentation. To test this, we analyzed the time course of unit activity in a moving 20 ms window and compared model fit and model preference of egocentric and allocentric units (defined based on the AIC analysis above) using cluster-based statistics to assess significance [26]. Model fit for sound angles relative to the head was greater for egocentric than allocentric units only between 5 and 44 ms after stimulus onset (Fig 9a, $p = 0.001996$). Model fit for sound angles in the world was greater for allocentric than egocentric units only between 6 and 34 ms after stimulus (Fig 9b, $p = 0.001996$). Model preference diverged only in the window between 4 and 43 ms after stimulus onset (Fig 9c, $p = 0.001996$). We observed no differences between egocentric and allocentric units before stimulus onset or when coordinate frames were aligned (Supplementary Figure 8). Thus the differences between egocentric and allocentric units reflected a stimulus-evoked effect that was only observed when head and world coordinate frames were free to vary.

Fig 9 Egocentric and allocentric tuning over time

a, Model fit for predicting neural activity from sound angles relative to the head. **b**, Model fit for predicting neural activity from sound angles in the world. **c**, Model preference. Data shown as mean ± s.e.m. for egocentric and allocentric unit populations. Black lines indicate periods of statistical significance (cluster based unpaired t-test, $p < 0.05$).



448

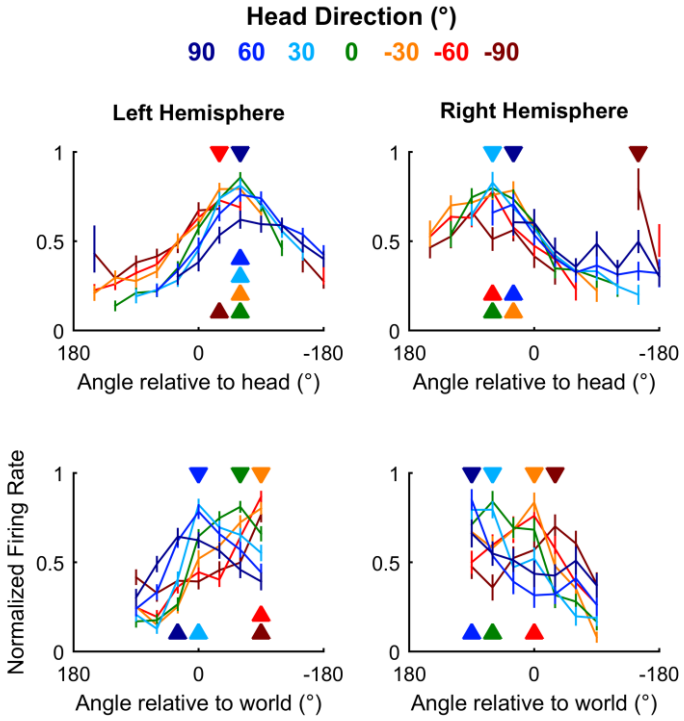
449 **Population representations of space**

450 We next asked how auditory cortical neurons behaved as a population when spatial tuning
451 was compared across head directions. In contrast to individual units, population activity more closely
452 reflects the large-scale signals observed in human studies using electroencephalography (EEG) to
453 distinguish coordinate frame representations [11, 12]. To form populations, we took the unweighted
454 mean of normalized firing rates from all units recorded from left (n = 64) or right (n = 28)
455 hemispheres and compared tuning curves measured at different head directions. As would be
456 expected from the predominance of egocentric units, we found that tuning curves for both left and
457 right auditory cortical populations were consistent within the head but not world coordinate frame
458 (Fig 10). Thus the allocentric units we find here are sufficiently rare as to be masked in overall

population readouts of spatial tuning, potentially accounting for conflicting findings of coordinate frame representations from EEG recordings.

Fig 10 Auditory cortical tuning

Population tuning curves plotted across head direction for mean (± s.e.m.) normalized response of all units in left and right auditory cortex; filled triangles indicate sound angle of maximum response at each head direction.



Distance modulation of cortical neurons and spatial tuning

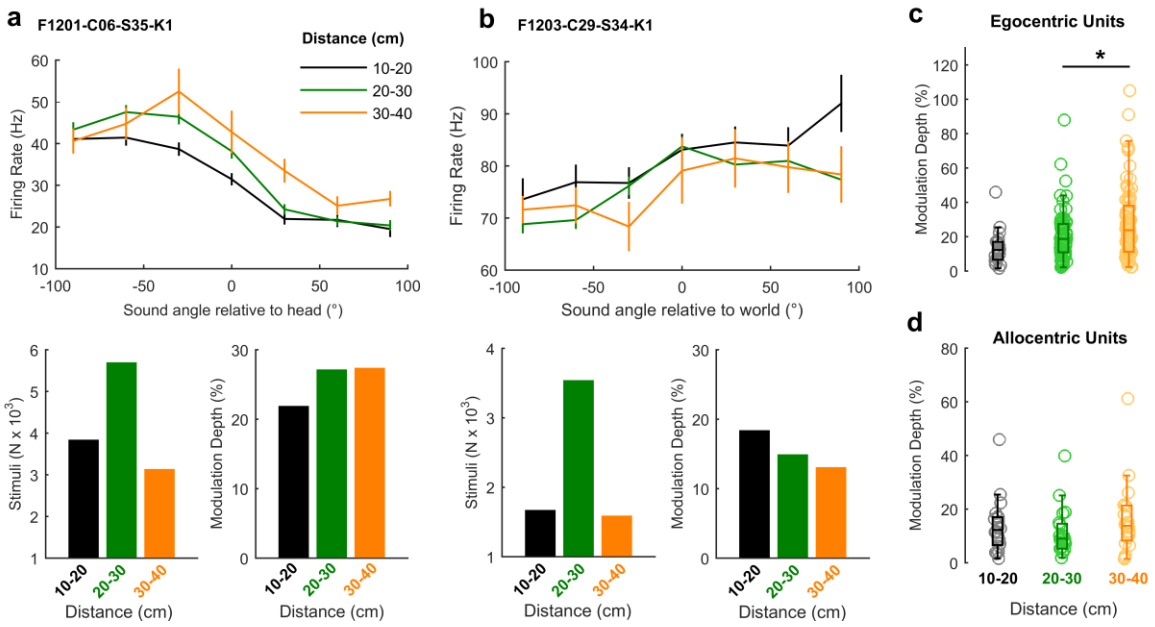
Studying auditory processing in moving subjects allowed us to resolve coordinate frame ambiguity so that we could determine the spaces in which neurons represent sound location. However recording in freely moving subjects also made it possible to go beyond angular measurements of the source location and address how neurons represented the distance of sound sources. Though often overlooked, distance is a critical component of egocentric models of neural tuning as the acoustic cues indicating sound location such as inter-aural level differences (ILDs)

change as sound sources approach the head. For distant sound sources (typically > 1 m) ILDs are small relative to distance-related attenuation of sound; however ILDs become much larger as sounds approach the head and source-receiver distance decreases [27, 28]. Neurons must therefore accommodate for distance dependent cues to accurately represent sound location across changes in head position. In our study, the distance between head and sound source ranged from 10 cm (the minimum distance imposed by the arena walls) to 40 centimetres, with only 3.37% of stimuli (mean across 57 test sessions) presented at greater distances (Fig. 2g) and thus all stimuli were likely to produce large ILDs [27, 28].

Spatial tuning was observed at all distances studied in both egocentric (Fig 11a) and allocentric units (Fig 11b) however modulation depth increased with distance for egocentric units (ANOVA, $F_{2,216} = 3.45$, $p = 0.0334$). Pairwise post-hoc comparisons showed that modulation depth was largest for sounds at greatest distances (Fig 11c) though the only significant difference was found for sounds presented 20 to 30 cm and 30 to 40 cm away ($t_{72} = -3.54$, $p = 0.0279$). In contrast, modulation depth did not change significantly with distance for allocentric units (Fig 11d, $F_{2, 57} = 0.962$, $p > 0.1$).

Fig 11 Spatial tuning across distance

a-b, Tuning curves of an egocentric (**a**) and allocentric (**b**) unit obtained with sound sources at varying distances from the animal's head. Bar plots show the number of stimuli and modulation depth for each tuning curve. **c-d** Distributions of modulation depth measured across distance for egocentric and allocentric units. Asterisk indicates significant pair-wise comparison (Tukey-Kramer corrected, $p < 0.05$).



494

---

495 **Speed modulation of cortical neurons and spatial tuning**

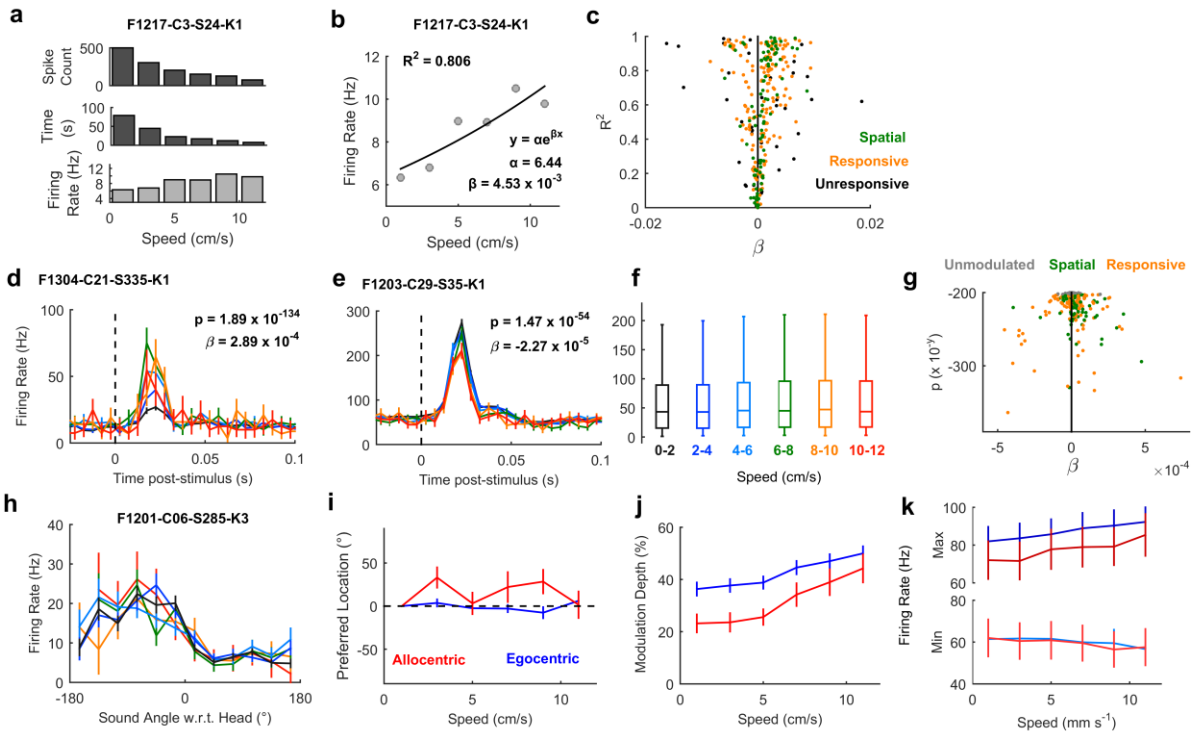
496 Changes in head position and direction also allowed us to investigate how speed of head
497 movement (Fig. 2f) affected neural activity. Movement is known to affect auditory processing in
498 rodents [29-31] but its effects on spatial representations of sound location and also on auditory
499 cortical processing in other phyla such as carnivores remain unknown. Here we presented click
500 sounds for which dynamic acoustic cues would be negligible and thus we could isolate the effects of
501 head movement on neural activity.

502 To address auditory cortical processing first, we asked how many of our recorded units
503 (regardless of auditory responsiveness or spatial modulation) showed baseline activity that varied
504 with speed. For each unit, we took all periods of exploration (excluding the 50 ms after each click
505 onset) and calculated the speed of the animal at the time of each action potential. We then
506 discretized the distribution of spike-triggered speeds to obtain spike counts as a function of speed
507 and normalized spike counts by the duration over which each speed range was measured. This
508 process yielded a speed-rate function for baseline activity (Fig 12a). We then fitted an exponential
509 regression curve to each function (Fig 12b) and plotted the correlation (R^2) and regression

coefficients (β) of each curve to map the magnitude and direction of association between speed and baseline activity (Fig 12c). Across the recorded population, we saw both positive and negative correlations representing units for which firing rate increased or decreased respectively with speed. However a significantly larger proportion of the population increased firing rate with speed across all units (t-test vs. 0; $t_{308} = 3.77$, $p = 1.97 \times 10^{-4}$). This was also true if we considered only sound responsive units ($t_{267} = 5.15$, $p = 5.17 \times 10^{-7}$) or only spatially tuned units ($t_{91} = 4.12$, $p = 8.41 \times 10^{-5}$).

Fig 12 Speed related auditory cortical activity and sensory processing

a Example calculation of speed-related modulation of baseline firing of one unit using reverse correlation. **b**, Example speed-firing rate function summarized using regression (β) and correlation (R^2) coefficients for the same unit as (a). **c**, Population distribution of regression and correlation coefficients showing the predominance of units with increasing speed-rate functions (β >0). **d-e**, Peri-stimulus time histogram of sound evoked responses for units that were modulated by speed: In (d) activity was enhanced as speed increased from 1 to 7 cm s^{-1} and decreased thereafter. In (e) firing decreased with increasing speeds, although speed-related modulation was smaller relative to sound evoked activity than (d). **f,** Box plots showing distributions (median and inter-quartile range) of evoked firing rates in response to clicks across speed for all sound responsive units. **g**, Population distribution of regression coefficients (β) and model fit (analysis of deviance p values) for all sound-responsive units. Units for which speed was not a significant predictor of neural activity ($p > 0.001$) denoted in grey. **h**, Spatial tuning curve for one unit for clicks presented at different head movement speeds. **i-k**, Change in preferred location (i), modulation depth (j) and min / max firing rates (k) of egocentric and allocentric units as a function of speed. Data for d-e and h-k shown as mean ± s.e.m.



531

---

532 We also observed speed-related modulation of sound-evoked responses in individual units
533 (Fig. 12d-e). For each individual unit, we characterized the relationship between head speed and
534 single trial firing rates (averaged over the 50 ms post-stimulus onset) using a general linear model
535 that measured both the strength (analysis of deviance *p* value) and direction (model coefficient, β)
536 of association. Positive β values indicated an increase in firing rate with increasing speed whereas
537 negative β values indicated a fall in firing rate with increasing speed. Thus the direction of the
538 relationship between firing rate and speed was summarized by the model coefficient, allowing us to
539 map the effects of movement speed across the population (Fig. 12f). For 199/268 sound-responsive
540 units (74.3%), speed was a significant predictor of firing rate (analysis of deviance vs. a constant
541 model, $p < 0.001$) however the mean coefficient value for movement sensitive units did not differ
542 significantly from zero ($t_{199} = 0.643$, $p = 0.521$). This suggests the sound-responsive population was
543 evenly split between units that increased or decreased firing with speed. We noted that a
544 significantly greater proportion of spatially modulated units (74/92, 80.4%) had sound evoked
545 responses that were sensitive to speed than units that were either not spatially modulated or for
546 which we had insufficient sample sizes to test spatial modulation (125/176, 71.0%)(Chi-squared test,

χ^2 = 3.96, p < 0.05). For those 74 spatially modulated and speed sensitive units, coefficients were mostly larger than zero (mean ± s.e.m. = 2.73 x 10^{-5} ± 1.53 x 10^{-5}) however this effect was not statistically significant (t_{73} = 1.80, p = 0.076). For the remaining speed sensitive units, the mean coefficients was closer to zero (mean ± s.e.m. = -5.02 x 10^{-6} ± 1.49 x 10^{-5}).

Lastly, we asked if speed affected spatial tuning. Spatial tuning could be observed at all speeds of movement in spatially tuned units (Fig. 12h) and the preferred locations of units did not vary systematically with speed (Fig 12i). For neither egocentric nor allocentric units was there a significant effect of speed in an ANOVA on preferred locations (egocentric: $F_{5, 432}$ = 0.53, p = 0.753; allocentric: $F_{5, 108}$ = 1.53, p = 0.188). However we did observe significant changes in modulation depth (Fig 12j) both for egocentric ($F_{5, 432}$ = 4.91, p = 0.0002) and allocentric units ($F_{5, 108}$ = 5.09, p = 0.0003), indicating that spatial modulation was greater when the head was moving fastest. Change in modulation depth resulted from both a gradual suppression in minimum and enhancement in maximum firing rates with speed (Fig 12k). However none of these changes in minimum or maximum firing rate were significant in comparisons across speed (ANOVA with speed as main factor, p > 0.5) indicating that it was only through the aggregative change in responses to both preferred and non-preferred locations that modulation depth increased with speed.

Discussion

Here we have shown that by measuring spatial tuning curves in freely moving animals it is possible to demonstrate the coordinate frames in which neurons define sound location. For the majority of auditory cortical neurons, we found egocentric tuning that confirm the broadly held but untested assumption that within the central auditory pathway sound location is represented in head-centered coordinates. We also identified a small number of units with allocentric tuning, whose responses were spatially locked to sound location in the world, suggesting that multiple coordinate frames are represented at the auditory cortical level. These units were consistently

identified across different analyses, observed in several subjects and could be dissociated from egocentric receptive fields recorded in the same behavioral sessions. Finally we explored the dependence of neural activity and spatial tuning on sound source distance and the speed of head movements, demonstrating that both factors can modulate firing rates and spatial tuning in auditory cortex.

Our results show that an animal's movement can be successfully tracked to measure head-centered egocentric tuning during behavior. While we used speakers placed at 30° intervals across a range of 180°, we were nonetheless able to characterize spatial tuning of egocentric units around the full circumference of the head (i.e. Fig. 7b). This illustrates the practical benefit of using moving subjects to characterize head-centered spatial tuning as the animal's head rotation generates the additional variation in sound angle relative necessary to fully map azimuthal tuning with a reduced number of sound sources. Furthermore, as the animal's head direction was continuous, the stimulus angle was also continuous and thus it was possible to measure spatial tuning at finer resolutions than that of the speaker ring from which stimuli were presented. In contrast to egocentric tuning, our ability to map allocentric receptive fields was limited by the speaker arrangement that only sparsely sampled world coordinates (Fig 2a). This may in part explain the low number of allocentric units in our population and denser sampling of the world may reveal unseen allocentric tuning – for example in the 94 units we recorded that were not spatially modulated by sound locations in the world that we tested. While a full 360° speaker ring may offer a minor improvement in sampling density, the radial organization of the ring remains a suboptimal design for sampling rectangular or irregular environments. To fully explore the shape of allocentric receptive fields will require denser, uniform speaker grids or lattices in environments through which animals can move between sound sources.

A notable property of egocentric units was the relationship between modulation depth of spatial tuning and distance, which was absent in allocentric units. This distance sensitivity may

596 largely be driven by changes in inter-aural level differences, although other auditory and non-
597 auditory cues can affect distance perception [32-36]. However modulation depth in our study was
598 lowest for proximal sounds when ILDs would be larger, and when other cues such as inter-aural time
599 differences remain relatively constant [27, 28]. Localization of nearby sound sources (< 1 m) is
600 possible for ferrets and humans [37, 38], though within the range of distances we considered here,
601 angular error of azimuthal localization in humans increases slightly as sounds approach the head
602 [37]. Thus our findings are consistent with human psychophysical performance but suggest larger
603 localization cues may not always produce better sound localization by neurons in auditory cortex. In
604 future it will be critical to validate our findings for sound sources at greater distances where sound
605 localization has been more widely studied.

606 In addition to recording many egocentric units, we also recorded a small number of
607 allocentric units, supporting the idea that auditory cortex represents sound location in multiple
608 coordinate frames [23] and parses an auditory scene into distinct objects [39, 40]. We believe this is
609 the first study to look for world-centered encoding of sound locations at the cellular level and thus
610 the units we recorded, while small in number, reflect a novel spatial representation in the auditory
611 system.

612 A key question is where allocentric units reside in cortex: Egocentric and allocentric units
613 were located on the same electrodes and cortical depths, in the primary and posterior regions of
614 auditory cortex. However, the low density of electrodes in our arrays, and their placement over a
615 low frequency border, prevented us from mapping the precise tonotopic boundaries necessary to
616 attribute units to specific cortical subfields [41]. Future use of denser sampling arrays may enable
617 cortical mapping in behaving animals and thus precise localization of allocentric units. We targeted
618 the low-frequency reversal between primary and posterior auditory cortex as these areas are likely
619 to be sensitive to inter-aural timing cues, and the animals involved in this work were also trained to
620 discriminate non-spatial features of low frequency sounds in another study [42]. Posterior regions

621 may correspond to part of the ‘what’ pathway in auditory processing whereas the anterior
622 ectosylvian gyrus may correspond to the ‘where’ pathway in which spatial tuning is more extensive
623 [43, 44]. It is thus likely that the coordinate frames represented in our population (where only 49%
624 of units were spatially sensitive) may be more ubiquitous in anterior regions of ferret auditory
625 cortex. Indeed given sensorimotor and cross-modal coordinate frame transformations are a key
626 feature of activity in parietal cortex [10], it is likely that allocentric representations exist beyond
627 auditory cortex.

628 While the observation of allocentric receptive fields in tonotopic auditory cortex is novel, the
629 existence of allocentric representations has been predicted by behavioural studies in humans [6, 13].
630 Furthermore, coordinate transformations occur elsewhere in the auditory system [23, 45] and
631 behavioral movements can influence auditory subcortical and cortical processing [29, 30]. Perhaps
632 most importantly, vestibular signals are integrated into auditory processing already at the level of
633 the cochlea nucleus [46] allowing the distinction between self and source motion [22]. Auditory-
634 vestibular integration, together with visual, proprioceptive and motor corollary discharge systems,
635 provides a mechanism through which changes in head direction can partially offset changes in
636 acoustic input during movement to create allocentric representations. At present it is unclear
637 whether allocentric representations require self-generated movement and it will be interesting to
638 test if world-centered tuning is present if head direction is systematically varied in stationary
639 animals.

640 It is also unclear how head position is also integrated into auditory processing. Positional
641 information within the medial temporal lobe (or its carnivore equivalent) is a leading candidate given
642 the connections between entorhinal and auditory cortex [47, 48]; however the functional
643 interactions between these areas and their potential contributions to allocentric processing remain
644 to be addressed. Another critical question relates to the visual (or other sensory) cues that animals
645 may use to orient in the world and define allocentric representations. Given that head and place

cells remap in different settings [49, 50] and that auditory cortex receives tactile and visual information through multiple pathways [48, 51], it will be interesting to determine if changes in visual / tactile environment affects tuning of allocentric receptive fields, and if so, what environmental features anchor auditory processing.

Variation in the animal's head position with movement also allowed us to study the effects of head movement speed on auditory processing and spatial tuning. In contrast to other studies in freely behaving animals that reported movement-related suppression of activity [29-31] we found that neurons tended to fire more strongly when the head moved faster (Fig 12). One reason for this difference may lie in the behavior measured: Other investigators have covered a diverse range of actions including locomotion in which both the head and body move and self-generated sounds are more likely. We only considered the head speed of an animal and did not track the body position that would distinguish head movements from locomotion (which was relatively limited given the size of the animal and the arena). It is thus likely that much of the variation in speed we observe is a product of head movement during foraging without locomotion and thus with relatively little self-generated sound. The behavior of our subjects may therefore present different requirements for auditory-motor integration that result in distinct neural effects.

We also observed that spatial modulation was also greater when the animal was moving faster, which may be consistent with the sharpening of tuning curves during behavioral engagement [15]. While sharpening of engagement-related spatial tuning was linked to a reduction in spiking responses at untuned locations, we observed non-significant decreases in peak and minimum firing rates (which together increased modulation depth) suggesting that the mechanisms underlying speed-related modulation of spatial tuning may be subtly different. At the acoustic level, faster movements provide larger dynamic cues [2, 3] that improve sound localization abilities in humans [3, 52, 53] and may explain the increase in modulation depth of units at greater speeds observed here.

In summary, we recorded spatial tuning curves in auditory cortex of freely moving animals to resolve coordinate frame ambiguity. We demonstrated many egocentric units representing location relative to the head and a small number of allocentric units representing sound location relative to the world. We also studied the role of distance and speed in auditory cortical processing. Together our findings illustrate that auditory cortical processing of sound space may extend to multiple coordinate frames and spatial dimensions such as azimuth and distance, as well as non-auditory variables such as speed of movement.

Methods

Ethics statement

All experimental procedures were approved by a local ethical review committee (AWERB; University College London and The Royal Veterinary College, University of London) and performed under license from the UK Home Office (Project License 70/7267) and in accordance with the Animals (Scientific Procedures) Act 1986.

Simulated spatial receptive fields

Egocentric

Egocentric tuning described the relationship between spike probability (*P*) and sound source angle relative to the midline of the subject's head (θ_{HS}) and was simulated in Matlab (MathWorks) using a Gaussian function:

$$P(\theta) = ae^{-\frac{(\theta_{HS}-b)^2}{2c^2}} \quad (1)$$

In the example shown in figure 1a, parameters (a = 1.044, b = 0° and c = 75.7°) were determined by manual fitting to find values for which egocentric and allocentric tuning matched qualitatively. The theta domain was between ±180° binned at 1° intervals and distance of sound sources was not included in the simulation.

Allocentric

Allocentric tuning describes the relationship between neural output (reported here as spike probability; *P*) and sound source position within the world measured in Cartesian (x, y) coordinates. Spatial tuning was simulated as the dot product of spike probability vectors returned from functions defined separately for positions on the x- and y-axes:

699 $$P(x, y) = f(x) \cdot f(y) \quad (2)$$

700 In figure 1b, logistic probability functions were used for both dimensions:

701 $$f(x) = \frac{e^{-\frac{(x-\mu)}{s}}}{s\left(1 + e^{-\frac{(x-\mu)}{s}}\right)^2} \quad (3)$$

702

704 $$f(y) = \frac{e^{-\frac{(y-\mu)}{s}}}{s\left(1 + e^{-\frac{(y-\mu)}{s}}\right)^2} \quad (4)$$

703

705 With μ = 1000 mm and s = 400 mm for the x-axis, and μ = 0 mm and s = 1000 mm for the y-axis. For
706 both axes, spike probability vectors were generated for domains between ± 1500 mm binned at 2
707 mm intervals.

708 *Stimulus presentation, head pose and movement*

709 The position and orientation of the subject's head within the world was described as a
710 coordinate frame transform composed of a translation vector between origins (indicating the head
711 position) and a rotation matrix between axes (indicating the head direction). Stimuli were presented
712 on each time step of the simulation from each speaker in a ring at 10° intervals, 1000 mm from the
713 origin of the world coordinate frame. As the simulation was deterministic, each stimulus was
714 presented to static simulations only once to calculate the response. When simulating motion, a
715 'pirouette' trajectory was constructed in which the subject's head translated on a circular trajectory
716 (radius = 50 mm; angular speed = 30° per time-step) while simultaneously rotating (angular speed =
717 10° per time-step) for 7200 stimulus presentations (Supplementary Video 2). For each stimulus
718 presentation, the stimulus angle was calculated relative to both the midline of the head and the
719 vertical axis of the arena (Supplementary Fig 9). Simulation responses were quantized in 1° bins.

720 **Animals**

721 Subjects were five pigmented female ferrets (1-5 years old) trained in a variety of
722 psychophysical tasks that did not involve the stimuli presented or the experimental chamber used in
723 the current study. Each ferret was chronically implanted with Warp-16 microdrives (Neuralynx, MT)
724 housing sixteen independently moveable tungsten microelectrodes (WPI Inc., FL) positioned over
725 middle and posterior fields of left or right auditory cortex. Details of the surgical procedures for
726 implantation and confirmation of electrode position are described elsewhere[54]. The weight and
727 water consumption of all animals was measured throughout the experiment. Regular otoscopic
728 examinations were made to ensure the cleanliness and health of ferrets' ears.

729 Subjects were water-restricted prior to testing, during which they explored the experimental
730 arena to find freely available sources of water. On each day of testing, subjects received a minimum
731 of 60ml/kg of water either during testing or supplemented as a wet mash made from water and
732 ground high-protein pellets. Subjects were tested in morning and afternoon sessions on each day for
733 up to three days in a week (i.e. a maximum of six consecutive testing sessions); on the remaining
734 weekdays subjects obtained water in performance of other psychophysical tasks. Test sessions
735 lasted between 10 and 50 minutes and were ended when the animal lost interest in exploring the
736 arena.

737

738 **Experimental Design and Stimuli**

739 In each test session, a ferret was placed within a D-shaped arena (Fig. 2a, rectangular
740 section: 35 x 30 cm [width x length]; semi-circular section: 17.5 cm radius; 50 cm tall) with seven
741 speakers positioned at 30° intervals, 26 cm away from a central pillar from which the animal could
742 find water. The periphery of the circular half of the arena was also fitted with spouts from which
743 water could be obtained. Animals were encouraged either to explore the arena by delivery of water

744 at all spouts, or to hold their head at the center spout by restricted water delivery at this location.
745 The arena and speakers were housed within a sound-attenuating chamber lined with 45 mm sound-
746 absorbing foam.

747 During exploration (n = 57 test sessions), clicks were presented from each speaker with
748 random inter-stimulus intervals (250 – 500 ms). The instantaneous energy of clicks minimized
749 dynamic cues, simplifying neural analysis and comparisons with other work on spatial encoding.
750 Clicks were presented at 60 dB SPL when measured from the center of the arena using a measuring
751 amplifier (Bruel & Kjaer 2636). However because sound level varied across the arena, we roved
752 sound levels over a ±6 dB range to reduce changes in level arising from differences in position of the
753 head within the sound field. The frequency response of each speaker (Visaton SC 5.9) was measured
754 using golay codes [55] and compensated for to produce a flat spectral output between 20 Hz and 20
755 kHz. Stimulus location and water delivery were independent and subjects were not required to
756 attend to stimuli in order to find water rewards. To avoid characterizing neural responses to the
757 sound of solenoid control signals, stimulus presentation and water reward were delivered in
758 separate alternating time windows; water was delivered in a short period of 1 to 2 seconds when
759 each solenoid was rapidly opened (100 ms duration) with a 10 second interval between delivery
760 windows in which click stimuli were presented. Sessions typically lasted approximately 15 - 20
761 minutes (median = 16.5 minutes; range = 6.15 – 48.0 minutes) in which several thousand stimuli
762 could be presented (median = 1984; range = 304 – 3937).

763 **Head Tracking**

764 During exploration of the experimental arena, the animal's head position and orientation
765 were tracked using two LEDs (red and green) placed along the midline of the head and recorded
766 using an RV2 video acquisition system (TDT) sampling at 30 frames per second and synchronized
767 with the electrophysiology recording hardware. For each video frame, the red and green LED
768 positions were identified in Matlab from a weighted image in which the channel color of the target

LED was positively weighted and all other channels negatively weighted. Each LED position was then taken as the center of the largest cluster of pixels containing the maximum weighted value. To maximize the frame rate of the camera, we recorded with a low exposure time (10-20 ms). Lower exposure also improved LED identification by reducing signal intensity in the background of each frame.

In cases where an LED went out of view of the camera (usually due to the roll or pitch of the head, or the recording cables obscuring the LED), the maximum weighted value identified as the LED would be a random point within the arena resulting either from a weak reflection or image noise. To remove such data, we set a minimum intensity threshold based on the distribution of maximum values in weighted images across all frames. In cases where the LED intensity failed to match the specified threshold, the LED position was noted as missing. To compensate for missing data, we estimated LED positions across runs of up to a maximum of ten frames (333 ms) using spline interpolation. Longer runs of missing data were discarded.

We then mapped each LED position in the image (M) into the behavioral arena to give the new position N using the transformation:

$$N = T + RM \quad (5)$$

Where T is the translation between the origin of the image coordinate frame (i.e. pixel [0,0]) and the origin of the arena coordinate frame (the center of the arena). And, R is the three-dimensional rotation matrix describing the rotation between the arena and image coordinate frames. T was obtained by manually identifying the pixel closest to center of the arena (i.e. the equidistant point between all speakers) in a calibration image captured at the start of each test session. R was estimated from singular value decomposition using the position and distance between known points in the arena (also identified manually from each calibration image). Here we estimated a 3D rotation matrix to take into account the position of the camera relative to the arena (i.e. above the arena

rather than below). All coordinate frames were represented using the right-hand rule (i.e. positive values for counter-clockwise rotation about the z-axis) to ensure consistency with the *atan2* function.

The animal's head position ($\overrightarrow{AH}$) was then calculated as the mid-point between the LEDs within the arena and was used to define the origin of the head-centered coordinate frame (Supplementary Fig 9). The animal's head direction (θ_{HA}) was calculated from the two argument arctangent function (atan2) of the vector between LEDs that defined the midline (*Y*-axis) of the head-centered coordinate frame ($\widehat{f_H}$). The *Z* axis was undefined by the tracking system as we only measured two points (red and green LEDs) with a single camera; this leads to ambiguity about the pitch and roll of the head. To compensate for this deficiency we assumed that when LEDs were visible, the *XY* plane of the head always matched the plane of the arena floor and that the Z-axis of the head was orthogonal to this plane and oriented towards the camera. Such assumptions are justified by the properties of the tracking system – as the head rolls or pitches away from the assumed conditions, it becomes impossible to identify both LEDs within the image due to the limited angular range of the each diode. Therefore tracking was impossible (in which case data was discarded) in the same conditions in which our assumptions became untenable.

By using the frame times recorded on the device, it was possible to create a time series of head position and direction within the arena that could be compared to the spiking pattern of neurons. We used the inter-frame interval and change in position of the head origin, smoothed with a nine-point Hann window to calculate the speed of head movement.

Neural Recording

Neural activity in auditory cortex was recorded continuously throughout exploration. On each electrode, voltage traces were recorded using TDT System III hardware (RX8 and RZ2) and OpenEx software (Tucker-Davis Technologies, Alachua, FL) with a sample rate of 25 kHz. For

817 extraction of action potentials, data were band-pass filtered between 300 and 5000 Hz and motion
818 artefacts were removed using a decorrelation procedure applied to all voltage traces recorded from
819 the same microdrive in a given session [56]. For each channel within the array, we identified
820 candidate events as those with amplitudes between -2.5 and -6 times the RMS value of the voltage
821 trace and defined waveforms of events using a 32-sample window centered on threshold crossings.
822 Waveforms were then interpolated (128 points) and candidate events combined across sessions
823 within a test run for spike sorting. Waveforms were sorted using MClust (A.D. Redish, University of
824 Minnesota, http://redishlab.neuroscience.umn.edu/MClust/) so that candidate events were
825 assigned to either single-unit, multi-unit clusters or residual hash clusters. Single units were defined
826 as those with less than 1% of inter-spike intervals shorter than 1 millisecond. In total 331 units were
827 recorded, including 116 single units (35.1%).

828 **Tracking unit identity across recording sessions**

829 Through the experiment, electrodes were descended progressively deeper into cortex at
830 intervals of 50 – 100 μm to ensure sampling of different neural populations. At most recording sites,
831 we tested animals on multiple sessions (1-6 sessions) across several (1-3) consecutive days.
832 Conducting test sessions over multiple days makes possible the recording of different units at a
833 single recording site over time (i.e. through electrode drift, gliosis etc.). To constrain our analysis to
834 units with a consistent identity we tracked the waveform of recorded units across sessions within a
835 test run. Our rationale was that a unit should have a constant waveform shape across test sessions
836 and any differences in waveform shape should be small relative to differences in the waveforms of
837 units measured on other electrodes or at other depths by the same electrode. Thus for one test
838 session at a given recording site, we calculated the Euclidean distance matrix between the mean
839 waveform recorded on that session (W_{Test}) and the mean waveform recorded on each additional
840 session at the same recording site (D_{Test}). We also calculated the distances between W_{Test} and the
841 mean waveform recorded for every session at different recording sites ($D_{Control}$). $D_{Control}$ provided null

distributions for waveform distances between pairs of neurons known to have separate identities (due to the spatial separation between electrodes at recording sites [>50 μm in depth, >500 μm laterally]). For a given waveform, we then calculated the statistical probability of observing distances between test waveform and waveforms *at the same recording site* given the distribution of distances between test waveforms and waveforms *at other recording sites*. For waveforms exceeding statistical significance (t-test; $p < 0.05$, Bonferroni corrected for the number of sessions conducted at the recording site), we concluded that the same neuron or neuronal population was recorded.

For runs of test sessions, we took the longest continuous run for which waveform distances were significantly smaller than expected by chance. The majority of units tested more than once could be tracked over all sessions tested (72.4%: 126 / 174 units) although the number of neurons tracked fell off with time.

Data Analysis

During exploration we characterized sound evoked responses from auditory cortical units. Each click stimulus and the concomitant neural response could be related to controlled variables determined by the experimental design and measured variables observed from head tracking. Controlled variables were the position of the sound source relative to the arena and sound source level in dB SPL whereas measured variables were the position and direction of the head relative to the arena, as well as head speed. Controlled and measured variables were combined to determine several experimental parameters: Stimulus position relative to the head was calculated as the vector $\overrightarrow{HS}$:

$$\overrightarrow{HS} = \overrightarrow{AS} - \overrightarrow{AH} \quad (6)$$

Where $\overrightarrow{AH}$ is the vector from arena origin to head origin and $\overrightarrow{AS}$ is the vector from arena origin to the sound source. Stimulus angle relative to the head (θ_{HS}) was calculated by subtracting head

direction in the arena (θ_{HA}) from the stimulus angle relative to the origin of the head coordinate frame:

$$\theta_{HS} = \text{atan2}\left(\overrightarrow{HS}_y, \overrightarrow{HS}_x\right) - \theta_{HA} \quad (7)$$

The distance between head and stimulus was calculated as the magnitude of $\overrightarrow{HS}$ and expressed as a ratio relative to the distance between arena and stimulus to calculate sound level at the head:

$$p_H = p_A \cdot \frac{|\overrightarrow{AS}|}{|\overrightarrow{HS}|} \quad (8)$$

Where P_H and P_A are sound pressures at the head and center of the arena expressed in pascals and sound level is expressed in dB SPL:

$$L = 20 \cdot log_{10}\left(\frac{\tilde{p}}{2 \times 10^{-5}}\right) \quad (9)$$

Sound level was calibrated to 60 dB SPL (0.02 Pa) at the center of the arena.

For each variable we calculated the value at the time of stimulus presentation (i.e. with a lag of 0 ms) and contrasted these values with the spiking responses of neurons. To study encoding of stimulus features (both measured and control variables) by neurons, single trial responses of individual units were summarized as the mean firing rate 0 – 50 ms after stimulus onset. This window was sufficiently long to characterize the response of units (Supplementary Fig. 5a and 6a) but also short enough that changes in head direction and position during the analysis window were small (Supplementary Fig 3). Sound-responsive units (268/336) were first identified as those with evoked firing rates that differed significantly from background activity measured in the 50 ms before

887 stimulus presentation (GLM analysis of deviance using Poisson distributions and log link function; $p \leq$
888 0.05).

889 Spatially tuned units were then identified using sound-evoked responses collected with the
890 animal at the center of the arena with the head and world coordinate frames in approximate
891 alignment. For a stimulus presentation to be included in this analysis, the animal's head origin was
892 required to be within 5 cm of a point 2.5 cm behind the arena center (Supplementary Fig. 4a). The
893 2.5 cm offset was applied to provide an approximate account for the distance between the animal's
894 snout and head center. Head direction was also required to be within ± 15° of the midline of the
895 arena (i.e. the line of symmetry of the arena, so that the animal was facing forward). Sound-evoked
896 responses under these constraints were then fitted with a GLM (Poisson distribution; log link
897 function) with sound source angle relative to the head binned in 30° intervals as predictor. To ensure
898 adequate data for statistical testing, units were only assessed if responses were recorded for ≥ 5
899 stimulus presentations in each angular bin (186/268 units). Units for which sound source angle
900 significantly reduced model deviance (χ^2distribution, $p \leq 0.05$) were classed as spatially tuned
901 (92/186 units). While this approach may not identify all spatially informative neurons (some of which
902 may signal sound location by spike timing rather than rate [18, 57] or that may be tuned only to
903 sounds behind the head that were not sampled by speakers in the aligned condition) it identified a
904 sub-population of spatially sensitive units on which further analysis could be performed.

905 To calculate spatial tuning curves, analysis was expanded to include all head positions and
906 directions recorded. To calculate world-based tuning curves, mean firing rate across trials (0 - 50 ms)
907 was plotted for each sound source angle relative to the arena. For head-based tuning curves, sound
908 source angle relative to the head was binned at 30° intervals and mean spike rate plotted as a
909 function of the bin center angle. To study super-resolution tuning of egocentric units (Fig 5a-b), the
910 bin width used to calculate curves was reduced to 20°, 10°, 5°, 2°, or 1°. To compare spatial tuning of
911 egocentric units with other studies, we also calculated preferred location, modulation depth and

tuning width and equivalent rectangular receptive field width for spatial tuning curves calculated relative to the head across 360° according to the methods used for awake cats [15, 18]. For allocentric units, we calculated preferred location and modulation depth for across sound location in the world.

Modulation depth analysis

For each unit we calculated the depth of spatial modulation for tuning curves in each coordinate frame. Unless otherwise stated, modulation depth (MD) was calculated as:

$$MD = \frac{\max(x) - \min(x)}{\max(x)} \times 100 \quad (10)$$

Where x is the vector of firing rates in response to sounds located in each 30° bin between ±90° either of the world or head coordinate frame.

Modulation depth could also be calculated for simulated neurons using the same equation but with x being a vector of spike probabilities. This approach allowed us to calculate modulation depth for simulated allocentric and egocentric units when presented with sounds during observed animal movement (Fig 3). In simulations, modulation depth could be calculated in head and world coordinate frames that were either relevant or irrelevant for neural activity depending on whether the simulation was allocentric or egocentric. We termed modulation depth in the irrelevant coordinate frame *residual modulation* when expressed as a ratio of modulation in the represented coordinate frame:

$$Residual\ Modulation = \frac{MD_{Irrelevant}}{MD_{Relevant}} \quad (11)$$

For allocentric simulations, the world was relevant and the head was irrelevant; whereas for egocentric simulations, the head was relevant and the world was irrelevant.

For each test session in which we observed animal behavior, we compared the relationship between the residual modulation calculated during simulations of both allocentric and egocentric units, with the standard deviation of the head directions (σ, Fig 3). We fitted a linear regression model to this relationship that was subsequently used to test if observed modulation depth values of real units were significantly greater than the residual modulation expected from the animal's behavior. The linear regression model was fitted using the *fitlm* function in Matlab (R2015a). For each observed unit we measured the standard deviation of head angles during neural testing (σ) and, together with the regression models, predicted the 95% or 99.95% (post-Bonferroni correction for 92 units) confidence interval of residual modulation values in the head and in the world coordinate frame. Prediction was performed in Matlab using the *predict* function with the most conservative options selected (simultaneous confidence bounds and prediction for new observations rather than fitted mean values) to give the widest confidence intervals and thus minimize the probability of false positives. If the observed modulation depth of a unit in a particular coordinate frame exceeded the upper bound of the confidence interval for that frame, we identified it as significantly modulated.

General Linear Models (GLMs)

To compare the relationship between single trial firing rates and sound source angles in the head and world coordinate frames, we fitted the average firing rate on each trial (0-50 ms) with a generalized linear regression model (Matlab, *fitglm* function: Poisson distribution, log link function). For both sound source angles relative to the head and relative to the world, we measured the deviance of models fitted separately with each parameter (D_{Test}). The Akaike information criterion [24] was used to compare test models and distinguish allocentric and egocentric units as those for which sound source angle relative to the world or head respectively provided the best model. For all

but one unit that was excluded from further analysis, either sound source angle relative to the world or head improved model fit compared to a constant model (analysis of deviance; Bonferroni correction for two comparisons; $p < 0.05$).

To visualize GLM performance (Fig 6), we calculated *model fit* for each unit and coordinate frame as:

$$Model\ Fit = \frac{D_{Const} - D_{Test}}{D_{Const} - D_{Full}} \quad (12)$$

Where D_{const} was the deviance resulting from a constant model, and D_{Full} was the deviance resulting from a full linear model that included both sound source angle relative to the head and relative to the world. We compared the model fit for data obtained when the head and world coordinate frames were free to vary, and when we restricted data to cases when the head and world coordinate frames were aligned (see above). We also repeated our analysis but with speaker identity or head direction information randomly shuffled between stimulus presentations prior to calculation of spatial tuning curves. Shuffling was repeated 1000 times and, across shuffles, we calculated the median model fit in head and world coordinate frames. Here we used the median rather than mean of the permuted values across shuffles were not always normally distribution. To test the effect of shuffle on model fit of all units, we performed a two-way anova on change in model fit with shuffle, with coordinate frame (head / world) and unit class (egocentric / allocentric). Post-hoc tests were conducted on change in model fit vs zero (t-test) with Bonferroni correction for multiple comparisons.

For each analysis in which we calculated model fit, we also calculated *model preference* as:

$$Model\ Preference = Model\ Fit_{Head} - Model\ Fit_{World} \quad (13)$$

Model preference could thus vary between -100% (better fit for neural data based on sound angle in the world) and +100% (better fit for neural data based on sound angle relative to the head).

For time-based comparison of model performance, we reduced the time over which firing rates were considered (from 50 to 20 ms) and repeated the analysis with a window offset by -60 ms to 90 ms after stimulus presentation that moved with a 2 ms interval. Model fit and preference values for allocentric and egocentric units were compared across time using a non-parametric cluster-based statistical test [26] implemented in Matlab through the FieldTrip toolbox [58].

Acknowledgements

We would like to thank Prof Jan Schnupp, Prof David McAlpine, Dr Daniel Bendor and Dr Peter Keating for discussions of this work.

References

1. Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P. Dynamics of Active Sensing and perceptual selection. Curr Opin Neurobiol. 2010;20(2):172-6. doi: 10.1016/j.conb.2010.02.010. PubMed PMID: 20307966; PubMed Central PMCID: PMCPMC2963579.

2. Wallach H. The role of head movements and vestibular and visual cues in sound localization. Journal of Experimental Psychology. 1940;27(4):339-68.

3. Perrett S, Noble W. The contribution of head motion cues to localization of low-pass noise. Percept Psychophys. 1997;59(7):1018-26. PubMed PMID: 9360475.

4. Wightman FL, Kistler DJ. Resolution of front-back ambiguity in spatial hearing by listener and source movement. J Acoust Soc Am. 1999;105(5):2841-53. PubMed PMID: 10335634.

5. Brimijoin WO, Boyd AW, Akeroyd MA. The contribution of head movement to the externalization and internalization of sounds. PLoS One. 2013;8(12):e83068. doi: 10.1371/journal.pone.0083068. PubMed PMID: 24312677; PubMed Central PMCID: PMCPMC3846779.

6. Brimijoin WO, Akeroyd MA. The moving minimum audible angle is smaller during self motion than during source motion. Front Neurosci. 2014;8:273. doi: 10.3389/fnins.2014.00273. PubMed PMID: 25228856; PubMed Central PMCID: PMCPMC4151253.

7. Brimijoin WO, Akeroyd MA. The role of head movements and signal spectrum in an auditory front/back illusion. Iperception. 2012;3(3):179-82. doi: 10.1068/i7173sas. PubMed PMID: 23145279; PubMed Central PMCID: PMCPMC3485843.

8. Davison AJ, Murray DW. Simultaneous localization and map-building using active vision. Ieee T Pattern Anal. 2002;24(7):865-80. doi: Doi 10.1109/Tpami.2002.1017615. PubMed PMID: WOS:000176446100001.

9. Castellanos JA, Neira J, Tardos JD. Multisensor fusion for simultaneous localization and map building. Ieee T Robotic Autom. 2001;17(6):908-14. doi: Doi 10.1109/70.976024. PubMed PMID: WOS:000173337600014.

10. Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. Nat Rev Neurosci. 2002;3(7):553-62. doi: 10.1038/nrn873. PubMed PMID: 12094211.

11. Altmann CF, Wilczek E, Kaiser J. Processing of auditory location changes after horizontal head rotation. J Neurosci. 2009;29(41):13074-8. doi: 10.1523/JNEUROSCI.1708-09.2009. PubMed PMID: 19828820.

12. Schechtman E, Shrem T, Deouell LY. Spatial localization of auditory stimuli in human auditory cortex is based on both head-independent and head-centered coordinate systems. J Neurosci. 2012;32(39):13501-9. doi: 10.1523/JNEUROSCI.1315-12.2012. PubMed PMID: 23015439.

13. Van Grootel TJ, Van Wanrooij MM, Van Opstal AJ. Influence of static eye and head position on tone-evoked gaze shifts. J Neurosci. 2011;31(48):17496-504. doi: 10.1523/JNEUROSCI.5030-10.2011. PubMed PMID: 22131411.

14. Schnupp JW, Mrsic-Flogel TD, King AJ. Linear processing of spatial cues in primary auditory cortex. Nature. 2001;414(6860):200-4. doi: 10.1038/35102568. PubMed PMID: 11700557.

15. Lee CC, Middlebrooks JC. Auditory cortex spatial sensitivity sharpens during task performance. Nat Neurosci. 2011;14(1):108-14. doi: 10.1038/nn.2713. PubMed PMID: 21151120; PubMed Central PMCID: PMCPMC3076022.

16. Stecker GC, Harrington IA, Middlebrooks JC. Location coding by opponent neural populations in the auditory cortex. PLoS Biol. 2005;3(3):e78. doi: 10.1371/journal.pbio.0030078. PubMed PMID: 15736980; PubMed Central PMCID: PMCPMC1044834.

17. Recanzone GH, Guard DC, Phan ML, Su TK. Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. J Neurophysiol. 2000;83(5):2723-39. PubMed PMID: 10805672.

18. Mickey BJ, Middlebrooks JC. Representation of auditory space by cortical neurons in awake cats. J Neurosci. 2003;23(25):8649-63. PubMed PMID: 14507964.

19. Malhotra S, Hall AJ, Lomber SG. Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. J Neurophysiol. 2004;92(3):1625-43. doi: 10.1152/jn.01205.2003. PubMed PMID: 15331649.

20. Smith AL, Parsons CH, Lanyon RG, Bizley JK, Akerman CJ, Baker GE, et al. An investigation of the role of auditory cortex in sound localization using muscimol-releasing Elvax. Eur J Neurosci. 2004;19(11):3059-72. doi: 10.1111/j.0953-816X.2004.03379.x. PubMed PMID: 15182314.

21. Nodal FR, Kacelnik O, Bajo VM, Bizley JK, Moore DR, King AJ. Lesions of the auditory cortex impair azimuthal sound localization and its recalibration in ferrets. J Neurophysiol. 2010;103(3):1209-25. doi: 10.1152/jn.00991.2009. PubMed PMID: 20032231; PubMed Central PMCID: PMCPMC2887622.

22. Wigderson E, Nelken I, Yarom Y. Early multisensory integration of self and source motion in the auditory system. Proc Natl Acad Sci U S A. 2016;113(29):8308-13. doi: 10.1073/pnas.1522615113. PubMed PMID: 27357667; PubMed Central PMCID: PMCPMC4961155.

23. Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM. Eye position affects activity in primary auditory cortex of primates. Curr Biol. 2003;13(7):554-62. PubMed PMID: 12676085.

24. Akaike H. New Look at Statistical-Model Identification. Ieee T Automat Contr. 1974;Ac19(6):716-23. doi: Doi 10.1109/Tac.1974.1100705. PubMed PMID: WOS:A1974U921700011.

25. Miller LM, Recanzone GH. Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. Proc Natl Acad Sci U S A. 2009;106(14):5931-5. doi: 10.1073/pnas.0901023106. PubMed PMID: 19321750; PubMed Central PMCID: PMCPMC2667094.

26. Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods. 2007;164(1):177-90. doi: 10.1016/j.jneumeth.2007.03.024. PubMed PMID: 17517438.

27. Brungart DS, Rabinowitz WM. Auditory localization of nearby sources. Head-related transfer functions. J Acoust Soc Am. 1999;106(3 Pt 1):1465-79. PubMed PMID: 10489704.

28. Brungart DS, Rabinowitz WM, Durlach NI. Auditory localization of a nearby point source. J Acoust Soc Am. 1996;100:2593.

29. Schneider DM, Nelson A, Mooney R. A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature. 2014;513(7517):189-94. doi: 10.1038/nature13724. PubMed PMID: 25162524; PubMed Central PMCID: PMCPMC4248668.

30. Williamson RS, Hancock KE, Shinn-Cunningham BG, Polley DB. Locomotion and Task Demands Differentially Modulate Thalamic Audiovisual Processing during Active Search. Curr Biol. 2015;25(14):1885-91. doi: 10.1016/j.cub.2015.05.045. PubMed PMID: 26119749; PubMed Central PMCID: PMCPMC4511122.

31. Zhou M, Liang F, Xiong XR, Li L, Li H, Xiao Z, et al. Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. Nat Neurosci. 2014;17(6):841-50. doi: 10.1038/nn.3701. PubMed PMID: 24747575; PubMed Central PMCID: PMCPMC4108079.

32. Mershon DH, King EL. Intensity and reverberation as factors in the auditory perception of egocentric distance. Perception & Psychophysics. 1975;18(6):409-15. doi: doi:10.3758/BF03204113.

33. Traer J, McDermott JH. Statistics of natural reverberation enable perceptual separation of sound and space. Proc Natl Acad Sci U S A. 2016;113(48):E7856-E65. doi: 10.1073/pnas.1612524113. PubMed PMID: 27834730.

1083 34. Kopco N, Shinn-Cunningham BG. Effect of stimulus spectrum on distance perception for
1084 nearby sources. J Acoust Soc Am. 2011;130(3):1530-41. doi: 10.1121/1.3613705. PubMed PMID:
1085 21895092; PubMed Central PMCID: PMCPMC3188969.
1086 35. Zahorik P, Brungart DS, Bronkhorst AW. Auditory distance perception in humans: A summary
1087 of past and present research. Acta Acustica. 2005;91:409-20.
1088 36. Kolarik AJ, Moore BC, Zahorik P, Cirstea S, Pardhan S. Auditory distance perception in
1089 humans: a review of cues, development, neuronal bases, and effects of sensory loss. Atten Percept
1090 Psychophys. 2016;78(2):373-95. doi: 10.3758/s13414-015-1015-1. PubMed PMID: 26590050;
1091 PubMed Central PMCID: PMCPMC4744263.
1092 37. Brungart DS, Durlach NI, Rabinowitz WM. Auditory localization of nearby sources. II.
1093 Localization of a broadband source. J Acoust Soc Am. 1999;106(4 Pt 1):1956-68. PubMed PMID:
1094 10530020.
1095 38. Wood KC, Town SM, Atilgan H, Jones GP, Bizley JK. Acute Inactivation of Primary Auditory
1096 Cortex Causes a Sound Localisation Deficit in Ferrets. PLoS One. 2017;12(1):e0170264. doi:
1097 10.1371/journal.pone.0170264. PubMed PMID: 28099489.
1098 39. Bizley JK, Cohen YE. The what, where and how of auditory-object perception. Nat Rev
1099 Neurosci. 2013;14(10):693-707. doi: 10.1038/nrn3565. PubMed PMID: 24052177; PubMed Central
1100 PMCID: PMCPMC4082027.
1101 40. Bregman AS. Auditory scene analysis. Cambridge, MA: MIT Press; 1990.
1102 41. Bizley JK, Nodal FR, Nelken I, King AJ. Functional organization of ferret auditory cortex. Cereb
1103 Cortex. 2005;15(10):1637-53. doi: 10.1093/cercor/bhi042. PubMed PMID: 15703254.
1104 42. Town SM, Wood KC, Bizley JK. Neural correlates of perceptual constancy in Auditory Cortex.
1105 BioRxiv. 2017. doi: https://doi.org/10.1101/102889
1106 43. Bizley JK, King AJ. Visual-auditory spatial processing in auditory cortical neurons. Brain Res.
1107 2008;1242:24-36. doi: 10.1016/j.brainres.2008.02.087. PubMed PMID: 18407249; PubMed Central
1108 PMCID: PMCPMC4340571.

44. Bizley JK, Bajo VM, Nodal FR, King AJ. Cortico-Cortical Connectivity Within Ferret Auditory Cortex. J Comp Neurol. 2015;523(15):2187-210. doi: 10.1002/cne.23784. PubMed PMID: 25845831.

45. Groh JM, Trause AS, Underhill AM, Clark KR, Inati S. Eye position influences auditory responses in primate inferior colliculus. Neuron. 2001;29(2):509-18. PubMed PMID: 11239439.

46. Shore SE, Roberts LE, Langguth B. Maladaptive plasticity in tinnitus - triggers, mechanisms and treatment. Nat Rev Neurol. 2016;12(3):150-60. doi: 10.1038/nrneurol.2016.12. PubMed PMID: 26868680.

47. Insausti R, Herrero MT, Witter MP. Entorhinal cortex of the rat: cytoarchitectonic subdivisions and the origin and distribution of cortical efferents. Hippocampus. 1997;7(2):146-83. doi: 10.1002/(SICI)1098-1063(1997)7:2<146::AID-HIPO4>3.0.CO;2-L. PubMed PMID: 9136047.

48. Budinger E, Heil P, Scheich H. Functional organization of auditory cortex in the Mongolian gerbil (Meriones unguiculatus). III. Anatomical subdivisions and corticocortical connections. Eur J Neurosci. 2000;12(7):2425-51. PubMed PMID: 10947821.

49. Muller RU, Kubie JL. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J Neurosci. 1987;7(7):1951-68. PubMed PMID: 3612226.

50. Fyhn M, Hafting T, Treves A, Moser MB, Moser EI. Hippocampal remapping and grid realignment in entorhinal cortex. Nature. 2007;446(7132):190-4. doi: 10.1038/nature05601. PubMed PMID: 17322902.

51. Bizley JK, Nodal FR, Bajo VM, Nelken I, King AJ. Physiological and anatomical evidence for multisensory interactions in auditory cortex. Cereb Cortex. 2007;17(9):2172-89. doi: 10.1093/cercor/bhl128. PubMed PMID: 17135481.

52. Kato MU, H., Kashino M, Hirahara T. The effect of head motion on the accuracy of sound localization. Acoustical Science and Technology. 2003;24(5):315-7.

53. Thurlow WR, Runge PS. Effect of induced head movements on localization of direction of sounds. J Acoust Soc Am. 1967;42(2):480-8. PubMed PMID: 6075941.

54. Bizley JK, Walker KM, Nodal FR, King AJ, Schnupp JW. Auditory cortex represents both pitch judgments and the corresponding acoustic cues. Curr Biol. 2013;23(7):620-5. doi: 10.1016/j.cub.2013.03.003. PubMed PMID: 23523247; PubMed Central PMCID: PMCPMC3696731.

55. Zhou B, Green DM, Middlebrooks JC. Characterization of external ear impulse responses using Golay codes. J Acoust Soc Am. 1992;92(2 Pt 1):1169-71. PubMed PMID: 1506522.

56. Musial PG, Baker SN, Gerstein GL, King EA, Keating JG. Signal-to-noise ratio improvement in multiple electrode recording. J Neurosci Methods. 2002;115(1):29-43. PubMed PMID: 11897361.

57. Middlebrooks JC, Clock AE, Xu L, Green DM. A panoramic code for sound location by cortical neurons. Science. 1994;264(5160):842-4. PubMed PMID: 8171339.

58. Oostenveld R, Fries P, Maris E, Schoffelen JM. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci. 2011;2011:156869. doi: 10.1155/2011/156869. PubMed PMID: 21253357; PubMed Central PMCID: PMCPMC3021840.