

# SCIENTIFIC REPORTS



OPEN

## A novel concept for dynamic adjustment of auditory space

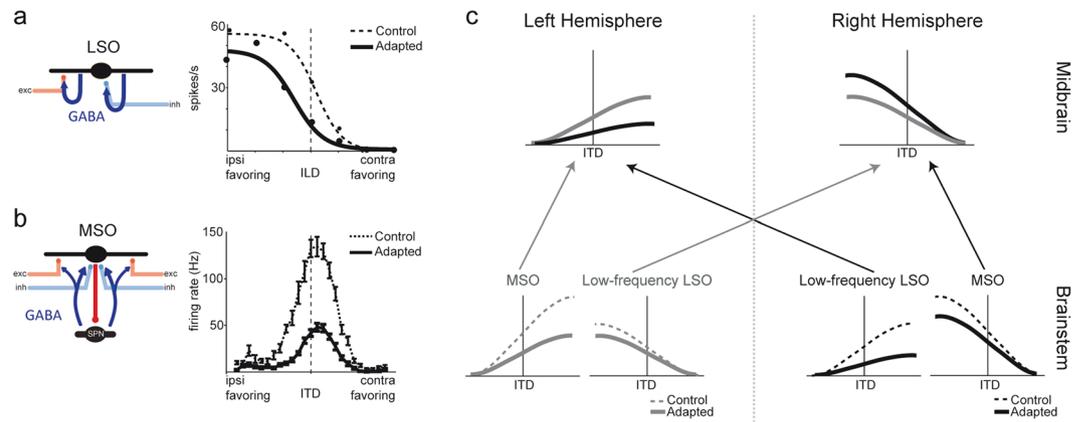
A. Lingner<sup>1</sup>, M. Pecka<sup>1</sup>, C. Leibold<sup>2</sup> & B. Grothe<sup>1</sup>

**Traditionally, the auditory system is thought to serve reliable sound localization. Stimulus-history driven feedback circuits in the early binaural pathway, however, contradict this canonical concept and raise questions about their functional significance. Here we show that stimulus-history dependent changes in absolute space perception are poorly captured by the traditional labeled-line and hemispheric-difference models of auditory space coding. We therefore developed a new decoding model incorporating recent electrophysiological findings in which sound location is initially computed in both brain hemispheres independently and combined to yield a hemispherically balanced code. This model closely captures the observed absolute localization errors caused by stimulus history, and furthermore predicts a selective dilation and compression of perceptual space. These model predictions are confirmed by improvement and degradation of spatial resolution in human listeners. Thus, dynamic perception of auditory space facilitates focal sound source segregation at the expense of absolute sound localization, questioning existing concepts of spatial hearing.**

The localization of sound sources is important for navigation and communication. This ability is neuronally conveyed by an exquisite sensitivity to physical differences of sounds between the two ears (interaural time and intensity differences, ITD and ILD, respectively). For example, human ITD sensitivity permits an angular separation acuity of only a few degrees<sup>1,2</sup> and is closely matched by the ITD sensitivity found on the level of single neurons in the mammalian auditory brainstem<sup>3–5</sup>. However, how the single cell activity of both ITD and ILD sensitive neurons is transformed into a neuronal code that is underlying the perceptual sensitivity has been debated intensively for decades: Several anatomical and physiological studies in Archosaurs (birds and crocodiles) confirmed the theoretical work of Jeffress<sup>6</sup>, proposing that ITDs are encoded via a so-called labeled-line mechanism resulting in a neuronal place code or map of auditory space<sup>7–10</sup>. Each ITD-sensitive neuron within a population exhibits a narrow receptive field tuned to a specific location in azimuth (=best ITD) with neighboring neurons encoding adjacent positions in space, reminiscent of spatial maps in other sensory modalities. However, in the mammalian ITD detector, the medial superior olive (MSO), as well as in downstream areas, the best ITDs of single neurons do not systematically vary with their location in the nucleus. Rather, MSO neurons exhibited relatively broad ITD sensitivity, where the slopes of single ITD functions span the entire occurring range of ITDs<sup>11–13</sup>. Individual neurons within the MSO of each brainstem hemisphere are similarly tuned (within each frequency band) with opposite tuning characteristics across both hemispheres during control conditions (Fig. 1, dashed lines). A similar hemispheric tuning has also been found for neurons in the Lateral Superior Olive (LSO), the mammalian ILD detector<sup>14</sup>. Owing to this fact, it has been concluded that the ITD and ILD of a sound source is encoded by the relative difference between average activities in the two MSO/LSO populations and this coding strategy is therefore often referred to as two-channel hemispheric (or hemispheric difference) model<sup>11,13,15</sup>.

Common to both theories on sound localization is that they are based on data from single sound source localization in isolation. However, there is psychophysical and phylogenetic evidence suggesting that absolute sound localization may not be the primary task of the mammalian auditory system<sup>16–19</sup>. Correspondingly, it was shown that binaural processing in MSO and LSO is subject to short-term adaptation based on feed-back loops in the mammalian auditory brainstem<sup>18,20,21</sup>, through which the neuronal firing rate is modulated by recent stimulus history (Fig. 1a,b). Specifically, in both nuclei, the inhibitory transmitter GABA (gamma-Aminobutyric acid) is released in activity-dependent manner and binds to pre-synaptic GABA-B receptors to mediated gain adaptation on the time scale of seconds. Such adaptation effects on the single cell level should influence the neuronal coding for acoustic space in the ITD and ILD pathway similarly and thus the perception of space<sup>18,20</sup>.

<sup>1</sup>Division of Neurobiology, Department Biology II, Ludwig-Maximilians-Universitaet Muenchen, Großhaderner Str. 2-4, D-82152, Martinsried, Planegg, Germany. <sup>2</sup>Bernstein Center for Computational Neuroscience Munich, Großhaderner Straße 2-4, D-82152, Martinsried, Germany. A. Lingner and M. Pecka contributed equally to this work. Correspondence and requests for materials should be addressed to B.G. (email: [grothe@lmu.de](mailto:grothe@lmu.de))



**Figure 1.** Neuronal coding adapts to stimulus history. (a,b) Firing rates of both LSO and MSO are modulated by recent stimulus history via a GABAergic feedback loop. Modified from Grothe and Pecka (2014) (c) Schematic of the Hemispheric Balanced Model including the effects of hemispherically specific adaptation: sound source azimuth is estimated based on a population vector analysis of MSO and low-frequency LSO for both hemispheres independently. The two estimates are then combined to yield a hemispherically balanced estimate of sound source location at the level of the midbrain.

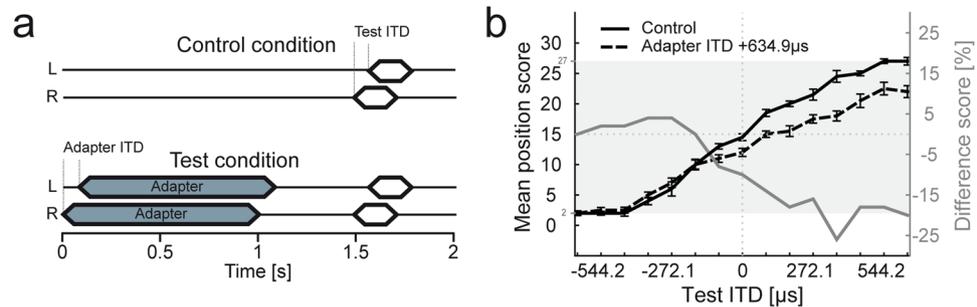
Indeed, several psychophysical studies demonstrated pronounced context-dependent sound localization in human listeners<sup>17,18,22,23</sup>. That is, the perception of a sound location changes depending on the stimulus context, i.e., the stimulus location and frequency preceding the target stimulus. Specifically, locations of sounds presented in close proximity or at the position of a preceding (tens or hundreds of milliseconds) sound (the “adapter”) are perceptually shifted away from the “real” position. These dynamic effects of stimulus history on auditory space perception were mostly in agreement with the hemispheric difference model of sound localization<sup>18</sup>. Nonetheless, findings of perceptual localization shifts for centralized adapter positions gave rise to the suggestions of an additional midline population channel<sup>24,25</sup>, implying that neither the labeled-line nor the hemispheric difference model might apply to human spatial processing. Thus, short-term history dependency of human ITD perception represents a crucial test for the validity of the assumed underlying processing models. However, so far these models were lacking an account of population responses in the brainstem including adaptation due to stimulus history and hence require refinement (compare Fig. 1a,b).

It is also unclear to what extent compromised absolute sound localization due to history-dependent modulation might serve an ecological purpose. Given the existence of dedicated feedback circuits in the auditory brainstem that are mediating this modulation<sup>18</sup>, we had previously suggested that it might serve to increase spatial sensitivity in complex listening situations. Similarly, Sach and colleagues<sup>26</sup> and Getzmann<sup>16</sup> had shown that the presence of a stimulus improves sound source separability for subsequent stimuli at the same spatial position (restricted to midline also in Maier *et al.*<sup>27</sup>). However, changes in absolute localization were not tracked by these studies, rendering a conclusion about the causal role of compromised absolute sound localization for the observed improvement impossible.

Here we studied context-dependent sound localization in human listeners, determined its effect on ITD resolution and used modeling of the data to decipher the underlying spatial code and to predict localization performance in more complex acoustic context. In accordance with previous studies<sup>17,18,22,23</sup>, we found pronounced shifts in the spatial perception of sounds away from the preceding adapter sounds, corresponding to a substantial miss-judgment of absolute sound source positions in the range of tens of degrees of auditory space. For strongly lateralized adapters, these shifts in perception were restricted to the hemisphere of the adapter location. Such hemispheric confinement of localization errors could not be explained by existing models of auditory space coding (i.e. labeled line or hemispheric difference), even when implementing the afore-mentioned history-dependent neuronal adaptation. Motivated by reports that a unilateral lesion of the midbrain still allows for sound localization in the hemisphere contralateral to the lesion<sup>28–30</sup>, and the fact that the low-frequency limb of the LSO is also ITD sensitive<sup>31,32</sup>, we developed a new hierarchical decoding model in which sound source azimuth is estimated based on population vector analysis for both hemispheres independently. The two estimates are then combined to yield a hemispherically balanced estimate of sound source location (Fig. 1c). This model not only explained the observed absolute localization errors, but also predicted a selective dilation and compression of auditory perceptual space relative to the adapter location. In accordance with this hypothesis, human listeners reported a focal increase in spatial resolution near the adapter on the expense of locations further away. Together, our findings indicate the need for a new concept for the coding of auditory space which is based on the assumption that spatial hearing in mammals serves relative separation rather than absolute sound localization.

## Results

**Selective modulation of spatial perception.** We first determined the localization of target tones with 15 different test ITDs without and with a preceding adapter (Fig. 2a). Listeners (N = 18) had to indicate the perceived intracranial position (stimuli delivered through headphones appear to be within the head) of the target tones by choosing a number from a linear scale between 0 (most left) to 30 (most right). In the control condition, without



**Figure 2.** Experimental designs and analyses of perceptual shift. **(a)** Schematic depicting a sample trial sequence to determine the intracranial perception of target tones without and with adapter. Both in the control and test condition target tones (white) were 220 ms in duration and were presented with a randomly chosen test ITD ranging from  $-634.9$  to  $+634.9$   $\mu\text{s}$  in  $90.7$   $\mu\text{s}$  steps. In the test condition the target tone was preceded by the adapter (1 s in duration) with a pause of 500 ms between adapter and target tone. All stimuli were presented with a frequency of 200 Hz **(b)** Position scores (indication of the listeners' intracranial target tone perception, with 0 being most left and 30 being most right) without and with adapter (solid and dashed black line) as a function of test ITD for one exemplary subject (data are represented as mean over ten trials for each test ITD,  $\pm$ S.E.M.). The grey shaded area depicts the individual range (i.e. range between the minimal (2) and maximal (27) mean position score for this subject). To determine the effect of a preceding adapter, the difference score (difference in position scores for each test ITD between control and test condition, normalized to the individual range) was calculated and plotted as a function of the presented test ITDs (grey solid line). Deviations from zero indicate a perceptual shift due to a preceding adapter (negative = shift to the left, positive = shift to the right).

adapter, mean position scores increased linearly with increasing test ITD (from 2 to 27 for an exemplary subject in Fig. 2b, black line), indicating that listeners were reliably able to report the perceived intracranial position of target tone ITDs, despite of small individual hemispheric asymmetries in the behavioral data (compare mean difference scores for test ITDs of  $-453.5$  and  $+453.5$   $\mu\text{s}$  in Fig. 2b).

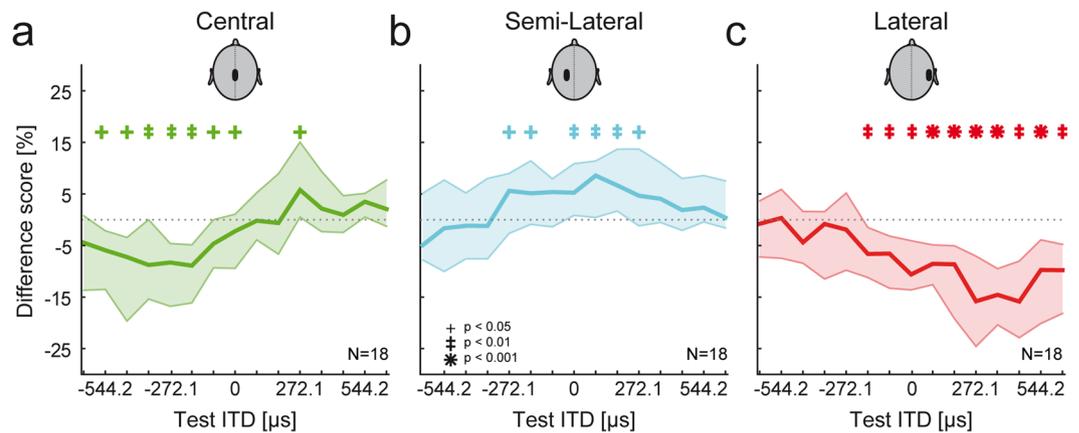
In the presence of a preceding, lateralized adapter (ITD of  $+634.9$   $\mu\text{s}$ , 1 s duration, same frequency as the test tone, see Methods), the mean position scores of listeners typically shifted to lower values (Fig. 2b, dashed black line), most prominently within the hemisphere where the adapter was presented. That is, sounds presented within the same hemisphere as the adapter were perceived closer to midline. Thus, in agreement with previous studies that had used similar paradigms<sup>23,33,34</sup> the perceived intracranial position of target tones was altered by the preceding adapter. This perceptual shift became particularly apparent when the difference score (difference in position scores for each test ITD between control and test condition, normalized to the total range of position scores for each subject) was calculated (grey line, Fig. 2b). Difference scores for test ITDs at and close to the adapter ITD were most negative denoting the strong perceptual shift to the left, i.e. away from the adapter ITD. Difference scores then gradually converged to zero for test ITDs farther away. That is, non-zero difference scores indicate a perceptual shift due to the presence of a preceding adapter.

To test for ITD-specific effects of the adapter on the perceptual shift, we obtained different scores for three distinct adapter ITDs (central: 0  $\mu\text{s}$ ; semi-lateral:  $-181.4$   $\mu\text{s}$ ; lateral:  $634.9$   $\mu\text{s}$ ). Generally, all three adapter ITDs lead to a systematic significant shift in the perception of target tone ITDs (Fig. 3, median difference score over all 18 listeners, individual symbols mark perceptual shifts of different significance levels, two-sided Wilcoxon signed rank test). However, the magnitude, extent and direction of the perceptual shift depended on the test condition. When being preceded by a central adapter, the perceived position of target tone ITDs to both sides of the adapter ITD are shifted away from it, leading to opposite directional shifts for both spatial hemispheres (Fig. 3a). Moving the adapter to a semi-lateral ITD ( $-181.4$   $\mu\text{s}$ ) also leads to a perceptual shift of target tone ITDs within both spatial hemispheres. In contrast to a central adapter, however, the perceptual shift is now restricted to mostly one direction, i.e. a shift to the right as indicated by a positive difference scores (Fig. 3b). Similar to a semi-lateral adapter the presentation of a lateral adapter ( $+634.9$   $\mu\text{s}$ ) leads to a perceptual shift also in only one direction. Beside the effect of adapter ITD on the shift direction, magnitude and extent (number of test ITDs significantly shifted) is also dependent on the adapter ITD being strongest for the lateral adapter, indicating a selectively altered perception of sound location due to the ITD of the adapter.

For all three tested conditions, a preceding adapter moved the perceived stimulus location away from the adapter, indicative of a dynamic process of spatial representation in humans that conflicts with the canonical concept of absolute sound source localization.

**A new model for binaural coding.** How can these changes in perception be explained neuronally? Generally, the observed effects of the adapter based on its ITD are in accordance with the recently described negative feedback loop in the auditory brainstem, where a stronger activation of the MSO by more lateralized adapters lead to more pronounced suppression of activity to trailing stimuli<sup>18</sup>. However, the nature of the MSO read-out ("the spatial code") is still debated<sup>12,35–38</sup>.

Therefore, in order to decipher the underlying coding mechanism we compared our experimental data to the outcome of three different model simulations. Based on anatomical and electrophysiological data of the superior



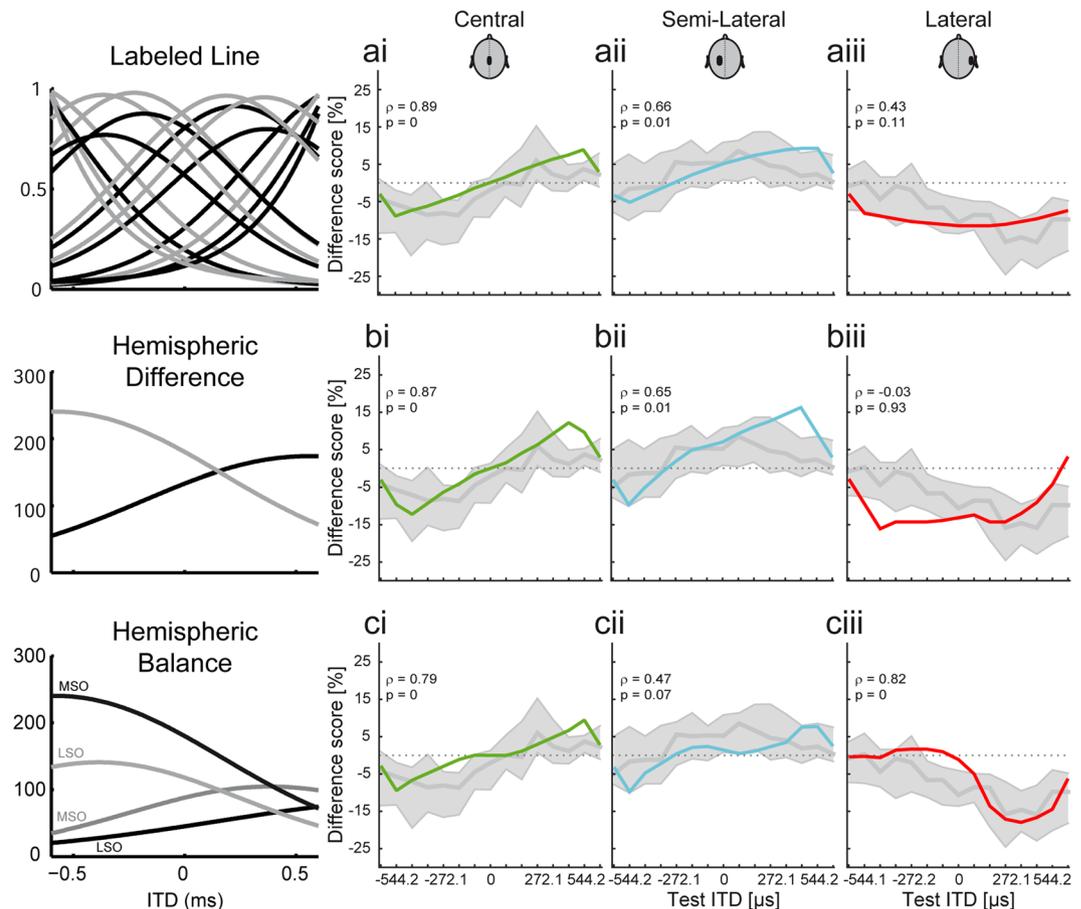
**Figure 3.** Adapter ITD selectively modulates ITD perception. Median difference scores over 18 subjects (solid line) and interquartile ranges (shaded area) for the three different adapter ITDs as a function of the test ITD are shown. Significant perceptual shifts are indicated by symbols (see legend in panel (b)), p-values derived from two-sided Wilcoxon signed rank test. (a) When a central adapter is presented, the perception of target tones with test ITDs to the left and right of the adapter are affected in different ways: the perception of target tones presented left of the adapter are shifted to the left (negative difference scores), whereas target tones presented right of the adapter are shifted to the right (positive difference scores). (b) The presentation of a semi-lateral adapter (ITD of  $-181.4 \mu\text{s}$ ) leads to a shift to the right (positive difference scores) for target tones with test ITDs ranging from  $-272.1$  to  $+272.1 \mu\text{s}$  (except for  $-90.7 \mu\text{s}$ ). (c) The presentation of a preceding adapter with an ITD of  $+634.9 \mu\text{s}$  shifts the perception of target tones with test ITDs ranging from  $-181.4 \mu\text{s}$  to  $+634.9 \mu\text{s}$  to the left (negative difference scores).

olivary complex, we modeled the population statistics of the MSO and the low-frequency limb of the LSO (ILSO), which is known to also be ITD sensitive<sup>31,32</sup> (see Methods). The three most important free model parameters (mean best delay, tuning width, and weight function non-linearity) were optimized according to physiological plausibility and overall fitting performance across all models and conditions (see Methods). The simulated neuronal activities including adaptation via negative feedback are the same for all three models. The models only differ in the applied decoding strategy. Two of them are based on classical principles of space encoding, namely labeled-line and hemispheric difference (Fig. 4a,b). For the labeled line model, each simulated neuron votes for its preferred ITD and the estimate is obtained from the respective population vector decoder. For the hemispheric difference model the activities are summed over all neurons in a hemisphere and the resulting population vector decoder is applied to only two dimensions. The third, referred to as “hemispheric balance”, model is based on a novel concept, which performs a population vector analysis in both brain hemispheres independently, leading to two location estimates that are then combined to yield a hemispherically balanced estimate of sound source location. The idea of independent hemispheric estimates is motivated by reports that a unilateral lesion of the midbrain still allows for sound localization in the hemisphere contralateral to the lesion, whereas there are stronger localization deficits in the ipsilateral hemifield<sup>28–30</sup>. Moreover, lesion of the commissure of Probst seems to specifically affect sound localization around midline<sup>39</sup>, where both hemispheres would contribute to a similar extent (according to eq. (1) in Methods). Hemispherically independent processing is further suggested by the stronger ipsilateral connections of the ascending auditory pathway after the level of the olivary complex and particularly the anatomy of the feedback loop in the MSO<sup>18</sup>, which does not cross the hemispheres and thus the resulting adaptive effects only have a unilateral computational purpose.

For all three tested models, the adapter had a significant effect on the lateralization of the target tones (estimated ITD of a test stimulus is based on the population activities in the adapted state derived from eq. (5)). The ability to predict perceptual shifts caused by the presentation of a preceding central adapter was good for all three decoding models ( $\rho = 0.89$  [labeled line],  $\rho = 0.87$  [hemispheric difference],  $\rho = 0.79$  [hemispheric balance]). Similarly, for the semi-lateral adapter, all three models only fit moderately well ( $\rho = 0.66$  [labeled line],  $\rho = 0.65$  [hemispheric difference],  $\rho = 0.47$  [hemispheric balance] Pearson rank correlation, see also Fig. 4a<sub>i</sub>-c<sub>i</sub>). Crucially, however, the non-uniform perceptual shift as observed after the presentation of a preceding lateral adapter, i.e. a shift that is restricted to target tone ITDs within the adapted hemisphere, was only significantly correlated with the predictions from the hemispheric balance model ( $\rho = 0.82$ ,  $p < 0.01$  Fig. 4c<sub>iii</sub>). The two other models erroneously predicted strong effects on target tone ITDs within the non-adapted hemisphere ( $\rho = 0.43$ ,  $p = 0.11$  [labeled line],  $\rho = -0.03$ ,  $p = 0.93$  [hemispheric difference], Fig. 4a<sub>iii</sub>-b<sub>iii</sub>).

The virtual lack of an effect of the lateral adapter on the contralateral perceptual space is the major feature distinguishing the hemispheric-balance model. The mechanical underpinnings of this difference is the combination of the two hemispheric estimates by the non-linear weight function  $g$  in equation (1), which effectively suppresses the estimate of the less active hemisphere.

Thus, all three models are able to predict the perceptual changes for a central and semi-lateral adapter reasonably well. However, only the hemispheric balance model was able to correctly reproduce the hemispherical constraint of changes in the listeners' perception found for the lateral adapter ITD. To further examine the validity

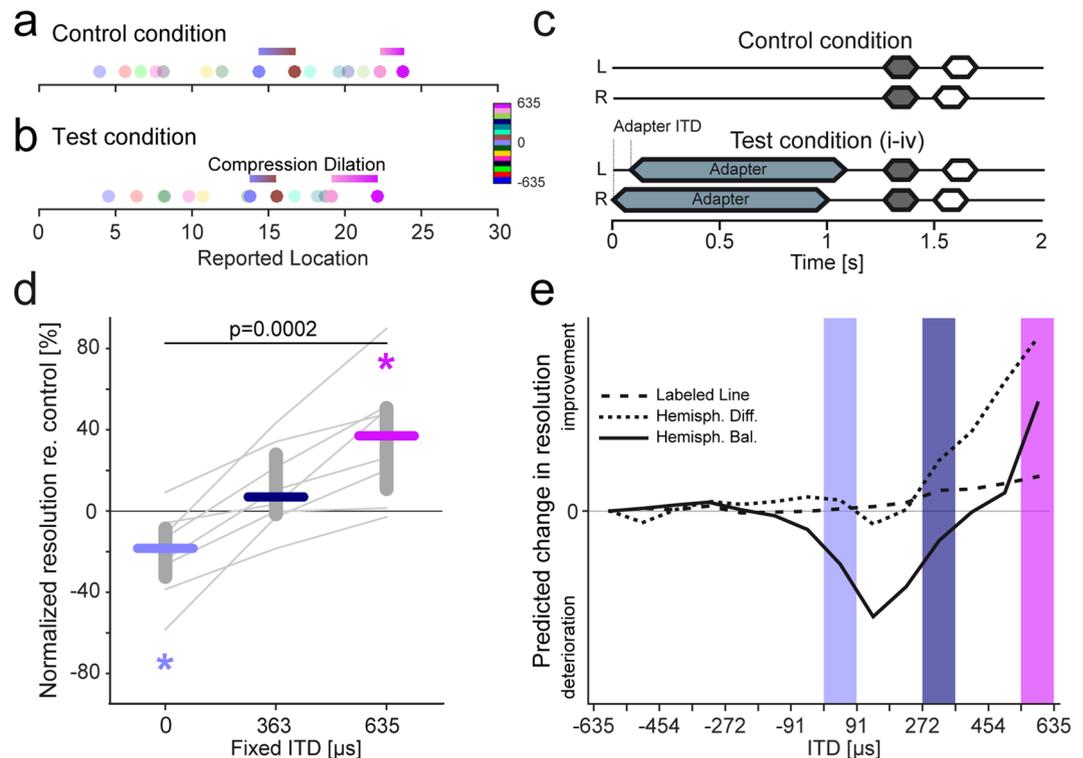


**Figure 4.** Comparison of neuronal computer modelling and experimental data. (a–c) Schematics of the three different model concepts (labeled line, hemispheric difference, and hemispheric balance model, respectively). Within each column the corresponding model predictions (solid, coloured lines) and the experimental data (grey) for one of the three different adapter conditions are shown (Pearson's rank correlation coefficients,  $\rho$ , and p-values are given for comparison). (a<sub>i</sub>–c<sub>i</sub>) All three models are able to predict the opposing shifts of target tone perception within the left and right spatial hemisphere for a preceding central adapter. (a<sub>ii</sub>–c<sub>ii</sub>) Perceptual shifts resulting from a preceding semi-lateral adapter are also reasonable good predicted by all three model concepts. (a<sub>iii</sub>–c<sub>iii</sub>) The most striking effect of the three models predictions is found for the lateral adapter: Only the hemispheric balance model restricts the predicted shifts to target tone ITDs within the adapter hemisphere, whereas the two other models erroneously predict a shift over the whole range of target tone ITDs.

of the models in more detail, we next set out to test their predictive power for human auditory spatial resolution. Specifically, we took advantage of these model-specific differences for the lateralized adapter condition, which result in distinct predictions how a listeners' ability to resolve nearby ITDs should be influenced:

To illustrate the dependency of shifted spatial perception on a listeners' azimuthal resolution, Fig. 5a depicts the distribution of reported locations for the 15 test ITDs of the localization task in the control condition for one representative subject. The resolution of nearby test ITDs can be approximated by the difference in the corresponding estimated locations (illustrated by the horizontal bars for two example ITD pairs). This difference corresponds to the relative distance in perceived spatial location. As we have seen (Fig. 3c), the presence of a lateralized adapter resulted in altered location estimates for a subset of test ITDs, namely those in the same hemisphere as the adapter. This asymmetry of the shifts in the listeners' perception consequently also altered the relative spacing in perceived spatial locations (Fig. 5b) and thus might affect their ability to resolve nearby locations. Specifically, for ITDs near the adapter, the spacing in perceived locations increased, suggesting the possibility of a relatively improved spatial resolution. In contrast, the spacing between nearby location estimates around midline decreased markedly (Fig. 5b), suggesting that the listeners' resolution might be impaired at these ITDs. We directly tested this hypothesis by evaluating the spatial resolution in human listeners ( $N = 8$ ) at three different ITDs.

**The impact of adaptation on sound source separability.** To this end, subjects had to report the more lateralized out of two successive target tones in an adaptive tracking paradigm to determine just noticeable differences in ITD (ITD-JNDs) with and without a preceding lateral adapter (+634.9  $\mu$ s ITD, 1 s duration, same frequency as the test tone, see Methods, Fig. 5c). On average, normalized JNDs (see Methods) during the presence of the lateral adapter were worst compared to control conditions at ITDs around midline and significantly



**Figure 5.** Selective change in resolution in human listeners is captured by the hemispheric balance model only. (a) Distribution of reported location values for the 15 test ITDs during the control condition for a representative subject. The resolvability of nearby test ITDs can be approximated by the difference in the corresponding location values (indicated by horizontal bars). (b) Location values for the test condition for the same subject show pronounced shifts for test ITDs in the hemisphere ipsilateral to the adapter ITD, yet only marginal shifts around midline. These non-uniform shifts result in a compression, i.e. decreased differences between location values around midline (compare horizontal bar lengths near  $0\mu\text{s}$  ITD in (a) and (b)). Simultaneously, the perception of auditory space close to the adapter ITD is dilated, indicated by increased differences between location values for nearby test ITDs (compare horizontal bar lengths near  $634.9\mu\text{s}$  ITD in (a) and (b)). (c) Schematic depicting a trial sequence to investigate the listeners' ITD resolution without and with adapter (ITD of  $+634.9\mu\text{s}$ ). One of the two consecutive target intervals was always presented with a fixed ITD (grey); the other interval was presented with a random ITD (white). The interval containing the fixed ITD was chosen randomly. (d) Normalized resolution (re. control) across 8 subjects (colored bars, data are represented as median; grey lines display individual subjects) is plotted as a function of ITD positions. For each subject, ITD resolution at each position was normalized to the resolution at  $-635\mu\text{s}$  for the control and test condition independently. The change relative to the control condition was calculated subsequently. Positive values indicate improved resolution, whereas negative values indicate deteriorated resolution due to the preceding adapter. Overall, the listeners' resolution increased significantly for locations closer to the adapter ( $p = 0.00019$ , Friedman test). Asterisks represents significantly improved or deteriorated normalized resolution for ITD positions close to the adapter position and at  $0\mu\text{s}$ , respectively ( $p < 0.05$ , two-sided Wilcoxon signed rank test). (e) The combined compression (deterioration) and dilation (improvement) in spatial resolution at  $0\mu\text{s}$  and close to the adapter location, respectively, was only captured correctly by the hemispheric balance model. Both other models did not predict the deterioration around midline, due to the missing hemispheric specificity of their fits to the different scores for the lateral adapter (see Fig. 4).

( $p = 0.0002$ , Friedman test) increased with best ITD-JNDs measured close to the position of the adapter. This gradual improvement of spatial resolution is the combined result of a significant improvement of ITD-JNDs close to the adapter and, importantly, a significant ( $p = 0.02$  two-sided Wilcoxon signed rank test) deterioration of ITD-JNDs around midline, as we had hypothesized based on the single subject data (Fig. 5a,b).

To validate the suitability of the various models (and thus the underlying processing strategy), we next used the estimations from the three models to predict changes in the listeners spatial resolution and compared the predicted change in resolution for all three model version to the human psychophysical data at the lateralized adapter condition at the same three test-ITDs (Fig. 5d,e). While all three models correctly predicted an improvement in ITD-JNDs at ITDs close to the adapter, only the hemispheric balance model correctly predicted a deterioration in ITD JNDs around zero  $\mu\text{s}$ .

Thus, only the hemispheric balance model predicted the combined improvement and deterioration in ITD resolution as a function of the relative distance tot the adapter (Fig. 5d). Importantly, these location-specific

effects are not predicted by the two other models, due to the lack of hemispheric specificity of their different scores for lateral adapters (cf. Fig. 4a<sub>iii</sub>, b<sub>iii</sub>).

Together, these data corroborate the validity of the hemispheric balance model to describe the neuronal processing strategy of mammalian coding of auditory space.

## Discussion

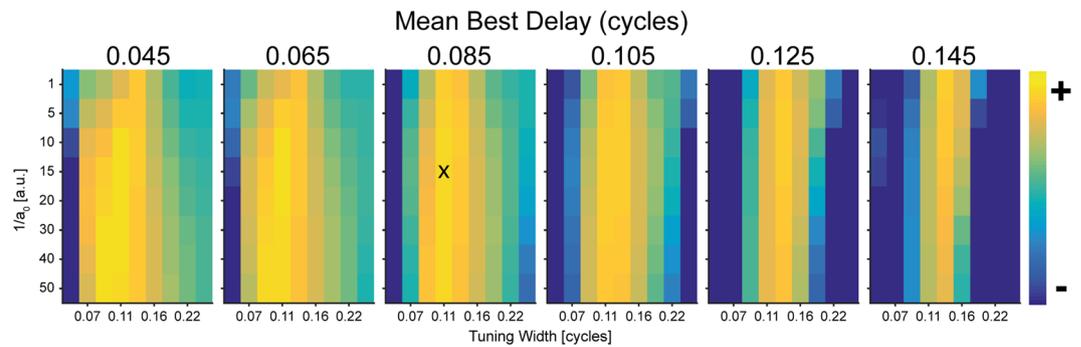
The present study investigated the interrelation between adaptive perception of absolute sound location, sound segregation and the underlying neuronal representation. We found a substantial shift in the subjective perception of sound location away from the preceding adapter, leading to significant errors in the absolute lateralization judgment. This altered perception of sound location resulted in an improved ability to spatially segregate sound sources close to the ITD of a preceding adapter, and in deterioration for ITDs further away. These findings indicate selective compression and dilation of perceptual auditory space coherent with our novel model of spatial coding in mammals based on a population vector analysis performed in both brain hemispheres independently. Together, the results indicate that spatial processing in mammals evolved for relative segregation rather than absolute sound localization.

Similar to other studies examining context-dependent spatial sensitivity (e.g.<sup>18,22,33,40</sup>), we found that test ITDs in close proximity to a preceding adapter are perceptually shifted away from this ITD. The affected hemisphere and also magnitude of perceptual shift was highly dependent on the adapter ITD. Specifically, a highly lateralized adapter (ITD  $\pm$  634.9  $\mu$ s) results in a perceptual shift of all target tones within the ipsilateral hemisphere away from the adapter. Only small or no shifts can be observed for sounds in the hemisphere contralateral to the adapter. For less lateral adapters (ITD  $\pm$  181.4  $\mu$ s) the magnitude of perceptual shift was smaller, however, within both the ipsi- and contralateral hemisphere. Thus, the extent of the perceptual shift scaled with the laterality of the adapter. This is also in agreement both with previous findings<sup>22,33</sup> and with the suggested underlying mechanism in the auditory brainstem, where a stronger activation of the MSO population by more lateralized adapters lead to more pronounced adaptive effects<sup>18</sup>. It is unlikely that the observed resolution improvement/deterioration are effects of altered spatial attention, because we controlled for this through normalization (see Methods) and it has been shown that these effects and the underlying shifts in spatial perception are highly specific for a match between adapter and test tones<sup>18,33</sup>. For the purpose of this study, the time interval between offset of the adapter and onset of the target tone was kept at 500 ms, as this interval has been used before to describe spatial adaptation effects<sup>18</sup>. This rather long-lasting history-dependence of the adapter on test ITDs corresponds well with the likely underlying physiological mechanism, namely a comparatively slow GABA-B mediated modulation of synaptic properties in the MSO and LSO<sup>18,20</sup>. However, at very short time intervals between adapter and target tone (<100 ms), the direction of perceived location shifts can be reversed, possibly due to the additional influence of mechanisms relating to the precedence effect<sup>41,42</sup>.

The present study tested ITD processing of low frequency sounds, representing the most relevant cue and frequency range for human spatial hearing. Stimulus history dependent spatial perception is, however, not only restricted to the low-frequency domain. Perceptual shifts away from the adapter ITD were also found for high-frequency stimuli based on envelope ITDs<sup>43</sup>, as well as based on ILDs<sup>17</sup>. This is consistent with the finding of GABA-B mediated adaptation in the LSO, which is functionally equivalent to the GABA-B mediated adaptation in the MSO. Hence, binaural adaptation of the computation of ILDs and ITDs at the earliest level of binaural processing seem to be a general feature in the mammalian auditory system. The high congruence we found between human perception and models based on physiological parameters determined in rodents<sup>4,5,12,18</sup> provides further corroboration for a common processing origin of mammalian spatial hearing<sup>19</sup>.

The description of the neuronal code for sound localization, and for ITDs in particular, was subject of several studies (see<sup>13</sup> for review). The two most prominent models for sound localization (labeled line and hemispheric difference model), however, are not able to fully account for the stimulus driven adaptation found in the present study (Fig. 4a<sub>i-iii</sub>, b<sub>i-iii</sub>). Both models fail to capture the effect of a highly lateralized adapter, because they predict a perceptual shift that affects both hemispheres in contrast to the experimental data showing only a limited effect on the contralateral hemisphere. There have already been studies that refined existing models of auditory space coding in order to predict human psychophysical adaptation data<sup>22,24,25,44</sup>. Kashino and Nishida<sup>22</sup> implemented a feedback gain control mechanisms in the context of a labeled line ITD processor to explain adaptational effects of sound localization. Other studies<sup>24,25,44</sup> extended the hemispheric difference model by a third midline channel to account for observed perceptual effects of a central adapter. However, – to our understanding – both the additional feedback gain control mechanism and also an additional channel alone cannot account for the aforementioned hemispheric specificity, as it would not be affected by lateralized adapters. In contrast, our newly developed hemispheric balance model was designed to capture this hemispheric specificity and also explains the effects of central adapters. Importantly, this model also made specific predictions about selective changes for sound source segregation due to compressive and dilative effects in perceived auditory space. While it has been suggested before that spatially congruent preceding stimuli might improve spatial resolution (<sup>16,26</sup>but see also<sup>27</sup>), no detrimental effects of an adapter on separability were known so far. Crucially, human perception closely followed both model predictions, showing both improvement and deterioration of ITD resolution at specific locations and thus corroborating the hemispheric balance concept as a potential basis for auditory space coding in mammals.

It follows that in contrast to Archosaurs, as shown for birds and crocodilians<sup>7,10,45,46</sup>, that exhibit a hard-wired labeled line coding of auditory space, mammalian auditory spatial processing is dynamic and does not provide an absolute and stable representation of the environment, but emphasizes focal separability of particular sound sources to the detriment of the overall localization ability. These differences might be based on the independent evolution of the middle-ear in birds and mammals and their different starting point in the evolution of spatial hearing about 200 million years ago (for review<sup>19</sup>): While the avian system is inherited from large predatory animals, the ancestors of tympanic ears in mammals are tiny nocturnal prey animals. Potentially, and in contrast



**Figure 6.** Summed log-likelihoods used for model parameter optimization. The summed log-likelihood (see Methods) for various combinations of tuning width and non-linear weight function ( $1/a_0$ ) at six different mean best delays are plotted. Lighter colors denote better overall fitting performance. A broad range of parameter combinations resulted in similarly good model performance (yellow) with insignificant changes in summed log-likelihood values. Settings used throughout the study (black x) were chosen based on highest log-likelihoods given physiologically plausible mean best delays ( $>0.065$  cycles).

to avian predator evolution<sup>47,48</sup>, the dominant evolutionary driving force of the early mammalian localization system was not the faithful mapping of the surrounding, but the sensitive detection of minute changes in spatially dynamic environments to avoid predatory pressure. Dedicated neuronal circuits allowing for such improvement in spatial resolution can be found already on the primary detector level of the brainstem<sup>18</sup>. We therefore propose that the mammalian auditory system evolved for source segregation rather than absolute localization.

## Material and Methods

**Psychophysics.** *Subjects.* 18 listeners (11 males and 7 females, mean age 24.4 years) participated in the first experiment, eight listeners (five males and three females) participated in the second experiment. All subjects showed normal hearing abilities between 250 Hz and 8000 Hz. All experimental protocols were approved by the ethical board of the Ludwig-Maximilians-Universität medical center and carried out in accordance with the ethical principles of the world medical association for research involving human subjects (Declaration of Helsinki). All subjects gave their written informed consent.

*Stimuli and experimental design.* All stimuli were generated in MatLab (The Mathworks, Inc, Natick, Massachusetts, US) at a sampling rate of 44.1 kHz and were digital-to-analog converted (Audio 2 Dj, Native Instruments GmbH, Berlin, Germany) before being presented over circumaural headphones (HDA 200, Sennheiser Electronic GmbH & Co. KG., Wedemark, Germany) to the listeners. In the first experiment one control and three test conditions were presented. In the control condition (Fig. 2a, top) a single 220-ms target tone (including linear rise-fall times of 10 ms) was presented. In the three test conditions (Fig. 2a, bottom) a 1-s adapter ( $20 \times 50$ -ms tone pips) preceded the target tone, separated by 500 ms silence. The adapter was presented with one of the following ITDs: (i)  $-181.4 \mu\text{s}$ , (ii)  $634.9 \mu\text{s}$ , or (iii)  $0 \mu\text{s}$ . Target tones were presented at 15 different test ITDs ranging from  $-634.9 \mu\text{s}$  to  $634.9 \mu\text{s}$  in steps of  $90.7 \mu\text{s}$ .

In the second experiment, the just noticeable difference in ITDs (ITD-jnds) was determined at three lateral ( $-634.9$ ,  $+362.8$ , and  $+634.9 \mu\text{s}$ ) and one midline ITD, each for a control and a test condition. In the control condition (Fig. 5c, top); two consecutive 160 ms target tones (pause: 140 ms) were presented to the listener. One target tone was presented with a fixed ITD (corresponding to one of the four test ITDs), the other one with an alternating ITD. Alternating ITDs were smaller than the fixed ITD for  $+362.8$ ,  $+634.9 \mu\text{s}$  and  $0 \mu\text{s}$  and bigger for the  $-634.9 \mu\text{s}$ . In the test condition (Fig. 5c, bottom) a 1-s adapter ( $20 \times 50$ -ms tone pips) with an ITD of  $+634.9 \mu\text{s}$  preceded the target tone pair, separated by 200 ms silence. Both for the control and test condition (i.e. with adapter) ITD-jnds determined at  $-634.9 \mu\text{s}$  served as a reference. All stimuli were presented at 200 Hz and 80 dB SPL.

*Procedure.* In the first experiment, subjects specified the perceived intracranial position of a target tone by pressing one of 31 linearly arranged buttons (0 being far left and 30 being far right) displayed on a computer screen. One trial consisted of one of the four experimental conditions. Listeners performed three sessions on the three different days. One session always consisted of the control condition followed by one of the three randomly chosen test conditions (150 trials with 10 trials per test ITD). For each condition, test ITDs were presented randomly. To exclude any effect of training listeners did not receive feedback about the presented test ITD and their decision. Moreover, for each listener and session the test condition was chosen randomly. Within one session, the test ITDs were presented randomly.

In the second experiment, listeners specified the perceived direction of a target tone pair (lateralized by ITDs) to determine ITD-jnds. The mean ITD-jnd for each experimental condition was determined by three consecutive runs, using a one-up three-down rule with feedback, as implemented by the MatLab AFC package<sup>49</sup>.

*Analysis.* For each subject and each test ITD, scores values of experiment one were averaged over all trials for the control condition and the test condition within one session (=mean position score). The individual range

(mean position score for +634.9  $\mu\text{s}$  minus mean position score for -634.9  $\mu\text{s}$ ) was calculated subsequently. To determine the effect of the adapter, a difference score was calculated by subtracting the mean position score of the control condition from the mean position score of the test condition (expressed as the percentage of the subjects' individual range). Mean position scores, individual ranges and difference scores were calculated for each session (i.e. experimental day) separately. The median difference score for each test condition over all 18 listeners was calculated subsequently.

ITD-jnds for each subject, condition, and tested ITD in experiment two were calculated as a median over nine runs. To determine the effect of a preceding adapter the so-called normalized resolution relative to the control condition was calculated. Therefore, ITD-jnds were normalized to the ITD-jnd determined at -634.9  $\mu\text{s}$ . This was done separately for the control and test condition. Normalized ITD-jnds for the control condition were then subtracted from the test condition.

**Simulation.** The fundamental idea of the model is that both brain hemispheres (L: Left, R: Right) independently generate ITD estimates  $\tau_{L/R}$ , based on population vector decoding of the activities in the binaural nuclei (MSO, ILSO). The two estimates are then combined to yield a hemispherically balanced ITD estimate

$$\alpha = \arg[g(a_R) \exp(i\tau_R) + g(a_L) \exp(i\tau_L)] \quad (1)$$

in which  $a_h$  are the total spike counts of the SOC afferents to midbrain hemisphere h (Fig. 1c) and  $g(a) = \exp(a/a_0)$  a monotonically increasing function (Numerical simulations are performed for  $a_0 = 0.07$ ).

*Intra-Hemispheric Estimates.* In analogy to the hemispheric-balance estimate  $\alpha$ , also the intra-hemispheric estimates  $\phi_h$  are assumed to arise from a population vector decoder. An essential parameter for such a decoder is the number of populations (channels) that contribute to it. While, in principle, each neuron in the respective nuclei could act as an individual channel (following the labeled-line code proposed by Jeffress<sup>6</sup>), other studies argued that fewer channels (one or two per hemisphere) might actually be sufficient<sup>11,12,50-53</sup>. Since the behavioural and functional differences between a low number of channels per hemisphere and a full-blown multi-channel code in which all neurons vote for their individual ITD<sup>54,55</sup> are still unclear, we rather follow an argument that is based on anatomy. Particularly, we center our considerations on the feedback systems of the MSO and low-frequency LSO (ILSO) that act on their presynaptic GABA-B receptors and mediate adaptation<sup>18,20</sup>. At the MSO, GABA is released into the extracellular medium by varicosities arising from the GABA-ergic SPN neurons<sup>18</sup>. This GABA release is not cell specific but rather affects a whole sample of the MSO neurons probably a large fraction of them within a best frequency band. For the ILSO the GABA feedback is more cell-specific since there GABA is released from the LSO dendrites<sup>20</sup>. However, also there GABA is released extra-synaptically and will also bind to GABA-B receptors on synapses of neighboring neurons. Both feedback mechanisms thus treat a subset of neurons in MSO and ILSO as computational entity and, thus, we assume that there are only two channels per hemisphere, one associated to MSO and the other one to ILSO. Note, that MSO and LSO have opposite hemispherical preferences, yet their projections to the IC lead to similar hemispherical preferences there because MSO projects ipsilaterally and LSO contralaterally.

If we denote the respective neuronal activities by  $\mu_h$  (MSO) and  $\lambda_h$  (ILSO) the hemispheric estimate can thus be computed as

$$\phi_h = \arg[\mu_h \exp(\pm i\phi_\mu) + \lambda_h \exp(\pm i\phi_\lambda)] \quad (2)$$

where the argument of the exponential function is positive for  $h = R$  and negative for  $h = L$ . The parameters  $\phi_\lambda > \phi_\mu > 0$  denote the azimuthal angles the population of MSO and ILSO neurons vote for. We made a straight-forward symmetric choice  $\phi_\mu = \pi/4$   $\phi_\lambda = \pi/2$ . The estimates  $\phi_h$  in equation (2) do not represent the azimuthal angle or ITD directly, they are conceived as a one-dimensional isomorphism of ITD that may be learned by the readout structures. More specifically, for each ITD (in steps of 5 $\mu\text{s}$ ) in a control situation (without adapter) we compute  $\phi_h$  corresponding to this ITD and subsequently use these ( $\phi_h$ , ITD) pairs as a lookup table to estimate the ITD  $\tau_h$  from the corresponding  $\phi_h$  value in the adapted case. Consequently, the specific choice of  $\phi_\mu$  and  $\phi_\lambda$  does not affect the results (as long as the two vectors  $\exp(i\phi_\mu)$  and  $\exp(i\phi_\lambda)$  are linearly independent). The activities variables  $\mu_h$  and  $\lambda_h$  might not necessarily be the spike counts but some nonlinear function of it. In this paper we introduce a threshold and saturation by choosing  $\mu_h = \tanh(k_h - 0.05)$  for  $k_h > 0.05$  and  $\mu_h = 0$  otherwise, where  $k_h$  denotes the total a scaled spike count of the MSO neurons in one frequency band and the  $\lambda_h$  is computed analogously for the spike count of the ILSO population.

It is important to note that the present version of the model does not include level dependence. It therefore assumes that the outcome is level-invariant. The latter can be achieved for the hemispheric estimates  $\phi_h$  if one assumes that the rates  $\mu_h$  and  $\lambda_h$  in both hemispheres change by the same factor upon a change in level. Moreover, the invariance property also requires that for the non-linear weighting function  $g$ ,  $a_0$  is increased by the same factor as the firing rates  $a$ .

*Two further decoding models.* the hemispheric difference model and the labeled line model are also used in this study for comparison. They both rely on variants of equation (2). For the hemispheric difference model the rates  $\mu_h$  and  $\lambda_h$  are replaced by the populations rates  $\mu_R$  and  $\mu_L$  of the two hemispheres (MSO and ILSO summed together) and the respective directions  $\phi_\lambda$  and  $\phi_\mu$  are taken as  $\pm 45^\circ$  resulting in  $\alpha = \arctan[(\mu_R - \mu_L)/(\mu_R + \mu_L)]$ , which around midline matches the normalized difference of the firing rates two hemispheres. The angle  $\alpha$  is then transformed to an ITD according to the same learning procedure as in the hemispheric balance model.

The labeled line model extends the sum in eq. (2) to as many summands as there are neurons (2000 per hemisphere; 1000 MSO and 1000 ILSO), with respective directions  $\phi_i$  that corresponds to the individual neurons best delay.

**Modelling Firing Rates.** The model for the firing rates of MSO and ILSO neurons underlying the population code takes into account the best frequency of the neuron as well as the frequency-dependence of best ITDs<sup>4,5</sup>

$$\text{best ITD}(f) = \text{CD} + \text{CP}/f, \quad (3)$$

in which CD and CP denote the characteristic delay and characteristic phase, respectively. The populations of MSO and ILSO neurons differ in the distributions of CPs and CDs. For MSO neurons we assume the CPs to be uniformly distributed between  $-1/4$  and  $+1/4$  of a cycle and the best phases (best ITDs at best frequency times best frequency) are normally distributed with mean 0.085 and standard deviation 0.05 cycles<sup>5</sup>. The CDs follow from eq. (3). For the ILSO population we take CPs uniformly distributed between  $+1/4$  and  $+3/4$  of a cycle and the best phases normally distributed with mean 0.775 cycles and standard deviation 0.05 cycles (cf.<sup>32</sup>). Note that 0.775 would correspond to a best phase of 0.225 on the level of the contralateral midbrain. The parameter choice implements the negative correlation between CP and CD observed in the DNLL of gerbils<sup>52,56</sup> see Supplementary Fig. 1a.

The rate  $r(f, \text{ITD})$  as a function of stimulus frequency  $f$  and ITD is thus modelled as

$$r(f, \text{ITD}) = A(f) \exp\{[\cos[2\pi f(\text{ITD} - \text{bestITD}(f))] - 1]/(2\pi\sigma)^2\} \quad (4)$$

where  $A(f) = (f/\text{BF})^{\text{sign}(\text{BF}-f)\kappa}$  introduces the best frequency BF of the neuron as the frequency for which the rate amplitude  $A$  is maximal and is applied to all simulated neurons.

The two remaining free parameters  $\sigma$  and  $\kappa$  determine the width of the ITD tuning curve and the spectral width of the receptive field, respectively. All numerical results in this paper are obtained for  $\sigma = 0.111$  and  $\kappa = 1/\ln(1.3)$  unless mentioned otherwise; see Supplementary Fig. 1b.

**Modelling Adaptation.** Motivated by the anatomy of the feedback loops in the SOC (see Methods), we assume adaptation to act on the level of populations, i.e., the feedback signal  $d$  is the same for every neuron in the same frequency band of the respective MSO or ILSO. Since GABA-B acts on the presynaptic side, we assume the adaptation to multiplicatively scale the neuronal activity and, thus, for every neuron  $i$  we obtain as the firing rate as

$$R_i(f, \text{ITD}) = r_i(f, \text{ITD})d_i(\mathbf{R}) \quad (5)$$

where  $r$  denotes the non-adapted rate from eq. (4) and  $\mathbf{R}$  is the vector of all firing rate in the ipsilateral MSO or ILSO, respectively. Owing to its recursive nature, we solve eq. (5) implicitly using Newton's method. The adaptation  $d_i$  of the  $i$ -th frequency-band (Supplementary Fig. 1c<sub>1</sub>) is thereby modelled as

$$d_i(\mathbf{R}) = 0.5 \tanh[-\beta((\mathbf{WR})_i - \theta)] + 0.5, \quad (6)$$

with  $\beta = 1/2$ , and a feedback matrix  $W$  that incorporates frequency-specificity of the feedback adaptation by

$$W_{ij} = (f_i/f_j)^{\text{sign}(f_i-f_j)/(4 \ln 1.3)}. \quad (7)$$

The best frequencies of the neurons  $i$  and  $j$  are thereby denoted by  $f_i$  and  $f_j$ , respectively. The threshold  $\theta$  equals 5 for the MSO population and 4 for the ILSO population to account for the lower average firing rate in the ILSO populations (Supplementary Fig. 1c<sub>2</sub>).

**Analysis of Model Performance.** To evaluate the goodness of the model fits to the behavioral data we calculated the Pearson's correlation coefficient ( $\rho$  with corresponding p-values for each model and each adapter condition, respectively (see Fig. Figure 4). Positive values of  $\rho$  denote a positive linear correlation, 0 denote no correlation and negative values give a negative correlation between data and model fit. We further determined the log-likelihood values of how well the model predicts the data. To this end we used the mean positional estimate  $m(\text{ITD})$  and variance  $v(\text{ITD})$  from the psychoacoustic data and defined likelihoods for the model predictions  $\alpha$  as

$$L(\alpha) = \prod_{\text{ITD}} \exp[-(\alpha - m(\text{ITD}))^2/(2v(\text{ITD}))](2\pi v(\text{ITD}))^{-1/2} \quad (8)$$

(logarithmic probability density function)  $\text{Log}_{10}$  likelihoods were computed for at each of the 15 test-ITDs separately and subsequently summed these values for each model and each adapter condition and subsequently summed to obtain one likelihood value for each model and parameter choice (See Supplementary Table 1). Here, higher log likelihood values denote better model fits.

We determined the log-likelihood value (using the logarithmic probability density function provided by Matlab, The Mathworks) at each of the 15 test-ITDs separately. In short, this procedure determines the likelihood of an observed value (in our case the model prediction at a given ITD) to stem from a distribution with defined mean and standard deviation (the observed behavioral data). Here, higher values denote better model fits. Log-likelihood were first summed across all test-ITD. subsequently summed these values for each model and adapter condition (See Supplementary Table 1). Here, higher values denote better model fits. To determine an unbiased optimal combination of mean best delay (6 different settings), tuning width (9 different settings) and the non-linear weighting function (8 different settings), these log-likelihood values were summed over all adapter condition and model version for each parameter combination separately. This analysis revealed a broad

range of parameter combinations that resulted in similarly good model performance with insignificant changes in summed log-likelihood values (Fig. 6). Settings used throughout the study were chosen based on highest log-likelihoods given physiologically plausible mean best delays.

## References

- Zwislocki, J. FR. Just noticeable differences in dichotic phase. *The Journal of the Acoustical Society of America* **28**, 860–864 (1956).
- Mills, A. W. On the Minimum Audible Angle. *The Journal of the Acoustical Society of America* **30**, 237–246 (1958).
- Goldberg, J. M. & Brown, P. B. Functional organization of the dog superior olivary complex: an anatomical and electrophysiological study. *J Neurophysiol* **31**, 639–656 (1968).
- Yin, T. C. & Chan, J. C. Interaural time sensitivity in medial superior olive of cat. *J Neurophysiol* **64**, 465–488 (1990).
- Pecka, M., Brand, A., Behrend, O. & Grothe, B. Interaural time difference processing in the mammalian medial superior olive: the role of glycinergic inhibition. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **28**, 6914–6925 (2008).
- Jeffress, L. A. A place theory of sound localization. *J Comp Physiol Psychol* **41**, 35–39 (1948).
- Knudsen, E. I. & Konishi, M. A neural map of auditory space in the owl. *Science* **200**, 795–797 (1978).
- Carr, C. E. & Konishi, M. A circuit for detection of interaural time differences in the brain stem of the barn owl. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **10**, 3227–3246 (1990).
- Hyson, R. L., Overholt, E. M. & Lippe, W. R. Cochlear microphonic measurements of interaural time differences in the chick. *Hearing research* **81**, 109–118 (1994).
- Koppl, C. & Carr, C. E. Maps of interaural time difference in the chicken's brainstem nucleus laminaris. *Biol Cybern* **98**, 541–559 (2008).
- McAlpine, D., Jiang, D. & Palmer, A. R. A neural code for low-frequency sound localization in mammals. *Nature neuroscience* **4**, 396–401 (2001).
- Brand, A., Behrend, O., Marquardt, T., McAlpine, D. & Grothe, B. Precise inhibition is essential for microsecond interaural time difference coding. *Nature* **417**, 543–547 (2002).
- Grothe, B., Pecka, M. & McAlpine, D. Mechanisms of sound localization in mammals. *Physiol Rev* **90**, 983–1012 (2010).
- Park, T. J., Klug, A., Holinstat, M. & Grothe, B. Interaural level difference processing in the lateral superior olive and the inferior colliculus. *J Neurophysiol* **92**, 289–301 (2004).
- Grothe, B. New roles for synaptic inhibition in sound localization. *Nat Rev Neurosci* **4**, 540–550 (2003).
- Getzmann, S. Spatial discrimination of sound sources in the horizontal plane following an adapter sound. *Hearing research* **191**, 14–20 (2004).
- Dahmen, J. C., Keating, P., Nodal, F. R., Schulz, A. L. & King, A. J. Adaptation to stimulus statistics in the perception and neural representation of auditory space. *Neuron* **66**, 937–948 (2010).
- Stange, A. *et al.* Adaptation in sound localization: from GABA(B) receptor-mediated synaptic modulation to perception. *Nature neuroscience* **16**, 1840–1847 (2013).
- Grothe, B. & Pecka, M. The natural history of sound localization in mammals—a story of neuronal inhibition. *Front Neural Circuits* **8**, 116 (2014).
- Magnusson, A. K., Park, T. J., Pecka, M., Grothe, B. & Koch, U. Retrograde GABA signaling adjusts sound localization by balancing excitation and inhibition in the brainstem. *Neuron* **59**, 125–137 (2008).
- Grothe, B. & Koch, U. Dynamics of binaural processing in the mammalian sound localization pathway—the role of GABA(B) receptors. *Hearing research* **279**, 43–50 (2011).
- Kashino, M. & Nishida, S. Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *The Journal of the Acoustical Society of America* **103**, 3597–3604 (1998).
- Phillips, D. P. & Hall, S. E. Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level. *Hearing research* **202**, 188–199 (2005).
- Dingle, R. N., Hall, S. E. & Phillips, D. P. A midline azimuthal channel in human spatial hearing. *Hearing research* **268**, 67–74 (2010).
- Dingle, R. N., Hall, S. E. & Phillips, D. P. The three-channel model of sound localization mechanisms: interaural time differences. *The Journal of the Acoustical Society of America* **133**, 417–424 (2013).
- Sach, A. J., Hill, N. I. & Bailey, P. J. Auditory spatial attention using interaural time differences. *Journal of experimental psychology Human perception and performance* **26**, 717–729 (2000).
- Maier, J. K., McAlpine, D., Klump, G. M. & Pressnitzer, D. Context effects in the discriminability of spatial cues. *Journal of the Association for Research in Otolaryngology: JARO* **11**, 319–328 (2010).
- Kelly, J. B., Li, L. & van Adel, B. Sound localization after kainic acid lesions of the dorsal nucleus of the lateral lemniscus in the albino rat. *Behav Neurosci* **110**, 1445–1455 (1996).
- Furst, M. *et al.* Sound lateralization and interaural discrimination. Effects of brainstem infarcts and multiple sclerosis lesions. *Hearing research* **143**, 29–42 (2000).
- Litovsky, R. Y., Fligor, B. J. & Tramo, M. J. Functional role of the human inferior colliculus in binaural hearing. *Hearing research* **165**, 177–188 (2002).
- Moore, M. J. & Caspary, D. M. Strychnine blocks binaural inhibition in lateral superior olivary neurons. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **3**, 237–242 (1983).
- Tollin, D. J. & Yin, T. C. Interaural phase and level difference sensitivity in low-frequency neurons in the lateral superior olive. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **25**, 10648–10657 (2005).
- Vigneaault-MacLean, B. K., Hall, S. E. & Phillips, D. P. The effects of lateralized adaptors on lateral position judgements of tones within and across frequency channels. *Hearing research* **224**, 93–100 (2007).
- Phillips, D. P., Mew, E. J. & Hall, S. E. Selective adaptation in sound lateralization is not due to a repulsion effect. *The Journal of the Acoustical Society of America* **136**, EL424 (2014).
- Smith, P. H., Joris, P. X. & Yin, T. C. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *The Journal of comparative neurology* **331**, 245–260 (1993).
- Beckius, G. E., Batra, R. & Oliver, D. L. Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **19**, 3146–3161 (1999).
- McAlpine, D. & Grothe, B. Sound localization and delay lines—do mammals fit the model? *Trends Neurosci* **26**, 347–350 (2003).
- McAlpine, D. Creating a sense of auditory space. *J Physiol* **566**, 21–28 (2005).
- Ito, M., van Adel, B. & Kelly, J. B. Sound localization after transection of the commissure of Probst in the albino rat. *J Neurophysiol* **76**, 3493–3502 (1996).
- Carlile, S., Hyams, S. & Delaney, S. Systematic distortions of auditory space perception following prolonged exposure to broadband noise. *The Journal of the Acoustical Society of America* **110**, 416–424 (2001).
- Kopco, N., Best, V. & Shinn-Cunningham, B. G. Sound localization with a preceding distractor. *The Journal of the Acoustical Society of America* **121**, 420–432 (2007).
- Kopco, N., Andrejkova, G., Best, V. & Shinn-Cunningham, B. Streaming and sound localization with a preceding distractor. *The Journal of the Acoustical Society of America* **141**, EL331 (2017).

43. Kawashima, T. & Sato, T. Adaptation in sound localization processing induced by interaural time difference in amplitude envelope at high frequencies. *PLoS one* **7**, e41328 (2012).
44. Phillips, D. P. A perceptual architecture for sound lateralization in man. *Hearing research* **238**, 124–132 (2008).
45. Carr, C. *et al.* Maps of ITD in the nucleus laminaris of the barn owl. *Adv Exp Med Biol* **787**, 215–222 (2013).
46. Carr, C. E., Soares, D., Smolders, J. & Simon, J. Z. Detection of interaural time differences in the alligator. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **29**, 7978–7990 (2009).
47. Wagner, H., Kettler, L., Orłowski, J. & Tellers, P. Neuroethology of prey capture in the barn owl (*Tyto alba* L.). *J Physiol Paris* **107**, 51–61 (2013).
48. Carr, C. E. & Christensen-Dalsgaard, J. Sound Localization Strategies in Three Predators. *Brain Behav Evol* **86**, 17–27 (2015).
49. Ewert S. D. AFC - A modular framework for running psychoacoustic experiments and computational perception models. *Proceedings of the International Conference on Acoustics AIA-DAGA 2013 in Merano, Italy*, 1326–1329 (2013).
50. Békésy, vG. Zur Theorie des Hörens. Über das Richtungshören bei einer Zeitdifferenz oder Lautstärkeungleichheit der beidseitigen Schalleinwirkungen. *Phys Z* **31**(824–835), 857–868 (1930).
51. Harper, N. S. & McAlpine, D. Optimal neural population coding of an auditory spatial cue. *Nature* **430**, 682–686 (2004).
52. Luling, H., Siveke, I., Grothe, B. & Leibold, C. Frequency-invariant representation of interaural time differences in mammals. *PLoS Comput Biol* **7**, e1002013 (2011).
53. Harper, N. S., Scott, B. H., Semple, M. N. & McAlpine, D. The neural code for auditory space depends on sound frequency and head size in an optimal manner. *PLoS one* **9**, e108154 (2014).
54. Day, M. L. & Semple, M. N. Frequency-dependent interaural delays in the medial superior olive: implications for interaural cochlear delays. *J Neurophysiol* **106**, 1985–1999 (2011).
55. Goodman, D. F., Benichoux, V. & Brette, R. Decoding neural responses to temporal cues for sound localization. *Elife* **2**, e01312 (2013).
56. Siveke, I., Ewert, S. D., Grothe, B. & Wiegrefe, L. Psychophysical and physiological evidence for fast binaural processing. *J Neurosci* **28**, 2043–2052 (2008).

## Acknowledgements

This work was supported by the German Center for Vertigo and Balance Disorders (IFB, TRFII-3) of the Bundesministerium für Bildung und Forschung (BMBF), the collaborative research grant CRC 870 (TP B01 and B02) of the German Research Foundation (DFG), and the Bernstein Center for Computational Neuroscience. We thank Vanessa Kratzer and Jakob Ulcik for their support during data acquisition.

## Author Contributions

Conceptualization B.G. and M.P.; Methodology A.L., C.L. and M.P.; Software A.L. and C.L.; Investigation A.L. and C.L.; Writing – Original Draft A.L., B.G., C.L. and M.P.; Writing – Review and Editing A.L., B.G., C.L. and M.P.; Visualization A.L., C.L. and M.P.; Supervision B.G.; Project Administration A.L. and M.P.; Funding Acquisition B.G., C.L.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-26690-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018