Cortical mechanisms of spatial hearing

Kiki van der Heijden[1,*], **Josef P. Rauschecker**[2], **Beatrice de Gelder**[3], **Elia Formisano**[3]

[1.]Department of ENT/Audiology, School for Mental Health and Neuroscience (MHeNS). Faculty of Health, Medicine, and Life Sciences. Maastricht University Medical Center. Maastricht, The Netherlands. [2.]Laboratory of Integrative Neuroscience and Cognition, Department of Neuroscience, Georgetown Medical Center, Washington United States. [3.]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.

Abstract

Humans and other animals use spatial hearing to rapidly localize events in the environment. However, neural encoding of sound location is a complex process involving the computation and integration of multiple spatial cues that are not represented directly in the sensory organ (the cochlea). Our understanding of these mechanisms has increased enormously in the past few years. Current research is focused on the contribution of animal models for understanding human spatial audition, the effects of behavioral demands on neural sound location encoding, the emergence of a cue-independent location representation in the auditory cortex, and the relationship between single-source and concurrent location encoding in complex auditory scenes. Furthermore, computational modeling seeks to unravel how neural representations of sound source locations are derived from the complex binaural waveforms of real-life sounds. In this article, we review and integrate the latest insights from neurophysiological, neuroimaging, and computational modeling studies of mammalian spatial hearing. We propose that the cortical representation of sound location emerges from recurrent processing taking place in a dynamic, adaptive network of early (primary) and higher-order (posterior-dorsal and dorsolateral prefrontal) auditory regions. This cortical network accommodates changing behavioral requirements, and is especially relevant for processing the location of real-life, complex sounds and complex auditory scenes.

Introduction

The position of a sound source reveals vital information about relevant events in the environment, especially for events taking place out of sight (for example, in a crowded visual scene). However, as sound location is not mapped directly onto the sensory epithelium in the cochlea, spatial hearing poses a computational challenge for the auditory system. The computational complexity is increased by the need to integrate information across multiple location cues and sound frequency ranges. For decades, auditory neuroscientists have examined the neuronal mechanisms underlying spatial hearing. This research shows that spatial cues are extracted and processed to a large extent in subcortical structures, but also

*kiki.vanderheijden@maastrichtuniversity.nl.

underscores a crucial role for the auditory cortex in neural sound location encoding (especially lesion studies[1-7]).

This Review brings together the latest insights into the cortical encoding of sound location in the horizontal plane (although of great interest, sound localization in the vertical plane is outside the scope of the present Review). Focusing on the specific contributions of the cortex to spatial hearing (that is, over and above subcortical processing), we discuss the empirical and theoretical work in the context of emergent perceptual representations of sound location. In particular, we describe the growing evidence for the relevance of cortical mechanisms and networks for goal-oriented sound localization, for spatial processing of real-life sounds, and for spatial hearing in complex auditory scenes.

Mammalian spatial hearing

The anatomy of the head, torso and pinna (the part of the ear residing outside of the head) introduces disparities in the time and intensity of sound waves emitted by a sound source when these arrive at the two ears. For mammals, these binaural disparities – the interaural time difference (ITD) and interaural level difference (ILD) – provide information on the spatial position of a sound source in the horizontal plane (FIG. 1A). In the case of periodic sounds (that is, pure tones consisting of a single sine wave), the delay between the sound waves arriving at each ear can also be expressed as the interaural phase difference (IPD, Box 1) instead of as a time difference. Another set of cues are the monaural, spectral cues. These cues are introduced by the shape of the pinnae and contribute to both horizontal and vertical sound localization[8], and to resolving *front–back ambiguities* in the horizontal plane[9].

According to the duplex theory of spatial hearing, the contribution of each binaural cue to sound localization is dependent on sound frequency. Specifically, ITDs are considered most relevant for localization of low frequency sounds (<1.5 kHz), and ILDs for localization of high frequency sounds (>1.5 kHz)[10]. Several psychoacoustic studies support this apparent cue-dichotomy for sound localization[11-15]; however, this theory has also been challenged by other studies showing that each type of binaural cue contributes to sound localization in a wide range of frequencies. For instance, ITDs conveyed by the envelope of high-frequency sounds, as well as the ILDs present in low-frequency sounds, can be used for localization[16-18] (especially in reverberant listening settings[19]). In addition, no clear relationship exists between neural tuning to sound frequency and to ITDs or ILDs. That is, ITDs modulate not only the firing rate of neurons that are tuned to low frequencies, but also those that are tuned to high frequencies[20]. Neurons in the inferior colliculus of the guinea pig that are tuned to low frequencies even respond maximally to ITDs outside of the physiological range[21]. Comparably, neural encoding of ILDs in the chinchilla midbrain is frequency invariant[22]. Thus, binaural cues seem to be relevant for a wider range of frequencies than predicted by the duplex theory, and neurons encode both types of binaural cue irrespective of their frequency tuning.

In terms of localization acuity, psychoacoustic studies show that the resolution of spatial hearing in the horizontal plane is highest around the interaural midline and deteriorates towards the acoustic periphery, and especially the back[9,23,24] (FIG. 1 B). Furthermore,

localization acuity is higher for broadband than for narrowband sounds, mostly because of the presence of monaural, spectral cues in broadband sounds[9,25-27]. However, most sound localization studies use artificial stimuli that listeners do not encounter regularly in daily life, such as tones, clicks and noise bursts[9,11,14,23,25,28-30]. Therefore, little is known about localization of complex, meaningful sounds. A study addressing this gap in psychoacoustic research showed that, in addition to the acoustic features described before, higher-order sound characteristics such as level of behavioral relevance and sound category modulate the localization acuity of complex, meaningful sounds [31]. These results highlight the need for realistic experimental set-ups using real-life sounds in ecologically valid listening settings (for example, with reverberation) in order to develop a more complete understanding of the complex mechanisms involved in mammalian spatial hearing.

Cortical spatial tuning properties

To understand the cortical mechanisms involved in spatial hearing, auditory neuroscientists have first sought to characterize neural spatial tuning properties in the network of densely interconnected primary and higher-order areas that together make up the mammalian auditory cortex[32-35]. In nonhuman primates, neurophysiological measurements have revealed that cortical spatial tuning is generally broad and predominantly contralateral (that is, the majority of neurons responds maximally to sound locations in the contralateral space)[36-38]. However, comparing spatial tuning between primary and higher-order auditory regions in response to complex, behaviorally meaningful sounds, such as conspecific calls, showed that neurons in higher-order caudal belt areas (especially the caudolateral belt area[39]) have a markedly higher spatial selectivity than those in the primary auditory cortex (PAC)[38-41]. By contrast, neural responses in rostral belt fields are less selective for space[39,40].

Similar spatial tuning properties have been observed in the auditory cortex of cats and ferrets. That is, the majority of neurons in the PAC is sensitive to sound location, their tuning is largely contralateral, and they typically have broad spatial receptive fields[42-47]. Moreover, under anesthesia, tuning in the majority of neurons in the PAC is level-dependent such that spatial receptive fields broaden further with increasing sound level[42,44,45] (but this was not confirmed in alert cats[48]). Further, similar to the findings in non-human primates, spatial selectivity in cat PAC was lower than in several higher-order areas, especially the dorsal zone and posterior auditory field[43,46,47].

A comparable cortical organization for spatial processing has been found in humans. Functional MRI (fMRI) research showed that sound location processing activates in particular the posterior auditory cortex, that is, the planum temporale[49-52], and the inferior parietal cortex[53]. Additionally, spatial tuning is broad and mostly contralateral[54-56], and spatial selectivity is higher in the posterior, higher-order planum temporale than in the PAC[57].

Taken together, several neuronal spatial tuning properties seem to be consistent across mammalian species: broad spatial receptive fields, an oversampling of contralateral space, and relatively higher spatial selectivity in posterior-dorsal regions. However, novel empirical

work has revealed that understanding of these properties requires further refinement. For instance, the behavioral state of a listener has been shown to affect neuronal spatial selectivity. In alert and behaving cats and humans, spatial tuning sharpened during goal-oriented sound localization, especially in the PAC[57,58]. This finding emerged only recently because, until now, most studies measured neuronal spatial tuning properties during passive listening[37,40,49-52,54] or even under anesthesia[42-47]. Importantly, these findings emphasize that measurements in alert and behaving subjects are needed to reveal modulatory influences of task performance and attention on neural sound location encoding.

In addition, findings from the past few years emphasize that some caution is required when generalizing neuronal spatial tuning properties across mammalian species owing to the differences in head morphology and size and their effects on neuronal spatial tuning. Specifically, in species with a relatively large head size (such as humans and other primates), the distribution of IPD tuning is not uniformly contralateral but dependent on sound frequency[59]. For instance, in macaque monkeys, the bimodal distribution of IPD preference consisting of two populations of laterally tuned neurons (as described in the previous paragraphs) is only observed for low sound frequencies (<1,000 Hz). For high sound frequencies (>1,000 Hz), the distribution of IPD tuning across the neural population is homogeneous and spans the entire azimuth. Only in mammals with a small head size, does spatial tuning consistently resemble the bimodal distribution of lateral best IPDs across low and high sound frequencies (as measured in subcortical structures)[21,59]. Thus, these findings challenge the notion that neuronal spatial tuning is predominantly contralateral for all mammalian species, and highlight differences between mammals with small heads and mammals with relatively large heads.

Neural coding of sound location

Given the properties of cortical spatial tuning, the question arises as to how sound location is processed in the auditory cortex. Specifically, what is the neural representation of sound location? Two complementary approaches have been used to address this question (FIG. 2). Most attention has been given to 'decoding' approaches, examining which aspects of the neural responses are most informative of sound location, and how this information can be read-out by downstream (that is, frontal) regions. Less attention has been given to the complementary 'encoding' approach, which aims to provide a mechanistic explanation of the transformation from binaural sound wave into neural response. Both approaches provide important insights into cortical spatial auditory processing, which we describe here.

Encoding models of binaural sound

Computational models of the transformation from stimulus to neural representation provide valuable insights into sensory processing. For example, a model based on interdependent spectro-temporal modulation encoding[60] accurately captures the transformation from real-life sound to its neural representation in the auditory cortex of humans[61,62] and macaques[63]. For sound localization, the place code proposed by Jeffress[64] was a first – and influential – step towards an encoding model describing the transformation from binaural sound to neural response. The place code posits that ITDs are encoded through ipsilateral and contralateral

axonal delay lines of varying length, each projecting to *coincidence detectors* at the next stage in the auditory hierarchy[64]. These coincidence detectors are thought to be tuned to a specific ITD to which they respond with the maximum firing rate[64]. Additionally, the coincidence detectors presumably have relatively sharp, level-invariant tuning curves that match the resolution of spatial hearing in behavioral reports of sound localization acuity, sample the azimuth homogenously and are organized topographically (reviewed elsewhere[65]). Thus, according to the place code, sound location is encoded by activating distinct clusters of neurons in a topographic manner through a system of delay lines and coincidence detectors.

The advantage of encoding models such as the place code is that they generate predictions about the neural response (in this case to sound location) that can be compared to actual neural data. For the place code, single-unit recordings confirmed the existence of azimuthal ITD maps in the brain stem of barn owls[66-68] and in multi-sensory structures in the brain stem of mammals[69] (Box 2). However, the broad spatial receptive fields with an oversampling of contralateral locations that are found in subcortical and cortical auditory regions in mammalian species are not in agreement with the topographic, homogenous neuronal sampling of the azimuth predicted by the place code. Moreover, the existence of ITD maps in barn owls can also be explained by other neural coding strategies[59] (see Box 1 and the next section).

Despite these apparent shortcomings of the place code, only a few other encoding models have been proposed. Młynarski[70] described a model that utilizes statistical regularities of real-life sounds as the driving force in the transformation from binaural real-life sounds to neural representations. This model is based on the theory of efficient coding for neural processing[71], which argues that neural stimulus representations are derived in a hierarchical, sparse manner from statistical regularities in the environment[72,73]. Founded on these principles, Mlynarski describes a two-stage model that transforms a binaural sound into a neural representation of spectrotemporal and interaural features. In the first stage, phase and amplitude information in each ear are extracted and separated with complex-valued, *sparse coding* that mimics cochlear filtering processes. In this stage, binaural phase information is also transformed into IPDs (reflecting medial superior olive processing). In the second stage, the model uses sparse coding to jointly encode monaural amplitude information from each ear and the IPDs, resulting in a spectrotemporal representation of the sound combined with spatial cues at the output stage (presumably the cortex). Given binaural real-life sound as input, this model predicts a cortical representation of sound location consisting of two neuronal subpopulations with broad receptive fields, tuned to lateral locations in opposing hemifields.

These model predictions are in close alignment with observations of neuronal spatial tuning in cat auditory cortex[46,70], indicating that the model describes relatively well the transformation from binaural sound to neural representation (at least for cats). This confirmation of the model's predictions also suggests that cortical spatial receptive fields in cats indeed reflect statistical regularities in binaural sound[70]. A similar relationship between statistical regularities in sound and neural encoding has been shown for other sound

attributes. For example, cortical spectrotemporal sensitivity[61,63] reflects the statistical regularities of spectrotemporal modulations in real-life sounds[74].

Along the same lines, Harper and McAlpine[75] introduced the notion of 'optimal coding'. Although this model is a conceptual rather than a formal encoding model (that is, it does not provide a mechanistic explanation for *how* binaural sound is transformed into the predicted neural representation), the optimal coding model is conceptually congruent with Mlynarski's encoding model[70] hence we discuss it here. Harper and McAlpine[75] posit that the neuronal representation of IPDs is organized such that the range of IPDs present in real-life binaural sounds is encoded as accurately as possible. In other words, the neural representation is considered to be a consequence of the sensory input that it receives, similar to the sensory processing theories underlying Mlynarski's model.

More concretely, because the range of IPDs in real-life sounds is dependent on head size and sound frequency (Box 1), the optimal coding model predicts that the neural coding strategy across different mammalian species is also dependent on these factors[75]. That is, for animals with a relatively large head size (such as macaques and humans), the optimal coding model predicts a mixture of neural IPD tuning: two subpopulations that have lateral IPDs (outside of the physiological range) in opposing hemifields for low frequency tones, and a homogeneous distribution of IPDs spanning the entire azimuth for high frequency tones (FIG. 2). For animals with a relatively small head size (such as gerbils), the optimal coding model predicts two subpopulations with lateral IPD tuning even at higher frequencies. As we described previously (see Cortical spatial tuning properties above), these predictions are in line with measurements of mammalian neural spatial tuning.

Interestingly, even though barn owls are considered distinct from mammals in terms of spatial auditory processing[65], the optimal coding model also explains the neural spatial tuning observed in this species. That is, neural encoding of IPDs in barn owls fits within the optimal coding framework by taking into account differences in the temporal resolution of the auditory system across species (Box 1): auditory nerve fibers in barn owls can phase lock to frequencies up to 10 kHz[76,77], enabling the extraction of accurate phase information at much higher sound frequencies than in other mammals (for example, in mammals such as humans, phase locking occurs up to a maximum of 1.5 kHz[78]). Consequently, even though barn owls have a small head, the optimal coding model predicts that for tones >3 kHz the most accurate IPD encoding requires a homogeneous distribution of IPDs across auditory neurons, in line with single-unit measurements of neural spatial tuning in barn owls[66]. Whether or not this homogeneous IPD distribution – either in barn owls or in mammals with a relatively large head size – arises from a delay line organization as proposed by Jeffress[64] is another matter that is not addressed by the optimal coding model.

In summary, the first steps in computational modeling of sound location encoding have been taken, but further developments are needed. The validity of existing models needs to be tested empirically, and comparisons across different models are required to better understand which computational mechanisms explain best the transformation from real-life binaural sound to neural representation. Such comparisons can also be made with fMRI in humans, using model-based analyses[62,79,80] or representational similarity analysis[81]. These methods

can evaluate models in terms of their ability to predict measured fMRI response patterns, and have been used to investigate the neural representations of sensory features in vision[79-81], as well as in audition[61,62,82]. For example, a comparative study testing the accuracy of predictions of various computational models of cortical sound encoding by measuring neuronal responses to naturalistic sounds with fMRI, showed that the cortex contains interdependent, multi-resolution representations of sound spectrograms[61]. Similar fMRI encoding studies can provide important insights into the cortical representational mechanisms of sound location. Importantly, the increased resolution and submillimeter specificity enabled by high-field fMRI (7 Tesla and higher) may be crucial to link the results in humans with those obtained in animal models[83].

Encoding of multiple sound attributes

The spatial position of a sound source is just one of the many attributes of sound. However, sound location can be considered somewhat distinct from other perceptual attributes that have a role in object recognition and categorization (for example, pitch) because it is used most directly for guiding sensorimotor behavior. Nevertheless, questions arise about when and how acoustic, perceptual, and location features are integrated to give rise to a unified neural representation of a localized sound object (reviewed in detail elsewhere[84]).

Interdependent encoding of multiple sound attributes may provide a (partial) answer to these questions. For instance, the encoding model[70] described in the previous section argues for joint encoding of spatial and spectrotemporal sound features. Single-cell recordings in ferrets also support this idea: neurons in the primary and higher-order auditory cortex of ferrets are not only sensitive to sound location but are also co-modulated by perceptual features such as pitch and timbre[85]. In mammals with intermediate or large head sizes, such as macaques, cortical tuning to IPD is a function of sound frequency[59]. Further, reports of a relationship between frequency tuning and ITD tuning have also been reported on a subcortical level[21,86]. However, research directly testing the principles of interdependent encoding for binaural real-life sound is scarce and more extensive studies are required to develop a better understanding of the relationship between the encoding of spatial position and other sound attributes[84].

Decoding models: reading the code.

In the previous sections, we described computational models of the physiological mechanisms underlying the transformation from binaural real-life sound into neural response (encoding models). A complementary question is how this information is subsequently read out by downstream areas (such as the frontal cortex) to give rise to a perception of sound location. That is, which information in the neural response is most informative about location, and how is this read-out by higher-order, non-sensory regions? To answer these questions, researchers have used decoding approaches that reconstruct (decode) the position of a sound source as accurately as possible from measured neural activity. In the next paragraphs, we discuss insights derived from these studies.

A first, much debated concept that has been investigated with decoding approaches is whether the neural representation of sound location is based on *opponent coding*

mechanisms. Theories of opponent coding build on the observation that the majority of spatially sensitive neurons exhibit the greatest response modulation at locations around the midline (both in subcortical structures and in the cortex)[21,37,46,54,87]. At a computational level, integrating information across neuronal populations with such tuning properties leads to the highest spatial acuity in the region around the midline – as evidenced by psychoacoustic studies[9,23,24] (see the section on mammalian spatial hearing). In addition, opponent coding has been considered a possible means to resolve the problem of level-invariance[46,54]. For accurate spatial hearing, the neural representation of sound location needs to be robust to changes in sound level. This can be achieved by an opponent coding mechanism in which the activity in two neuronal populations tuned to opposite hemifields is compared.

On a neural level, opponent coding mechanisms can take different forms: comparing the difference in mean or summed activity between two contralaterally tuned channels across hemispheres[21,46], or between an ipsilateral and contralateral channel within a single hemisphere[46,70]. Other studies have even suggested that a third, frontally tuned channel is involved[88,89]. Decoding studies in cats[46], humans[54] and rabbits[87] confirm that such opponent coding strategies convey information on sound location. Other empirical work – such as neural adaptation experiments and psychoacoustic studies – further support the validity of an opponent population rate code in the human auditory cortex[90-93].

The opponent coding mechanisms described above rely on population activity rate; however, whether population sums or averages capture the richness of information available in the population response has been a matter of debate. Researchers have therefore also explored biologically plausible ways in which the brain can decode sound location from the *pattern* of population activity, focusing on approaches such as population vectors[94,95] and maximum likelihood estimation (MLE)[96]. Although population vector codes based on the activity of broadly direction-sensitive neurons accurately represent the direction of arm movement in the motor cortex, they have been less successful in applications to the representation of sound location[87,97].

More promising are the results of maximum likelihood models, which have provided sound location estimates corresponding to the level of behavioral acuity from spike rate patterns in the caudolateral area of macaque auditory cortex[97], and in rabbit inferior colliculus[87]. Likelihood estimation is a form of template matching in which the observed neural response across a population is compared to a template (that is, the idealized response curve) that is derived from tuning curves of neurons[96,98]. Consequently, the neurons that contribute most to the location estimate are the neurons for which the estimate lies in the region of largest modulation (that is, the slope of the tuning curve), not the neurons that are most active[98]. Thus, MLE models are well suited to extract information from cortical neural response patterns. Furthermore, although these MLE models were initially developed for neural spike rates, they can be adapted to other types of data. For instance, a 2018 study successfully applied a modified version of the MLE model to fMRI activity patterns to decode sound location from the human auditory cortex[57]. Moreover, likelihood estimation does not require an explicit definition of opponent coding as a subtraction mechanism, and can even accommodate a mixed coding strategy in which spatial tuning is dependent on sound

frequency and head size[60]. Thus, likelihood estimation is a good candidate model for the read-out of sound location from population activity patterns by higher-order regions.

Taken together, evidence is growing that – in mammals – population pattern representations are more informative of sound location than rate codes or population vector representations. Additionally, most empirical findings described here indicate that the read-out of sound location from neural responses benefits from the inclusion of opponent mechanisms. The precise form of such opponent coding remains a matter of debate. Furthermore, understanding of where this read-out of spatial information takes place is also limited. In the next paragraphs, we therefore examine sound location processing on a larger scale, that is, within the cortical auditory processing network.

The cortical spatial auditory network

Hierarchical, specialized processing.

A prominent model of auditory processing is the *dual-stream model,* which posits that auditory processing takes place in two functionally specialized pathways[99,100]: A ventral 'what' pathway dedicated to processing sound object identity, and a dorsal 'where' pathway dedicated to spatial processing (analogous to the dual-stream model for the visual system[101,102]). That is, this framework of functionally specialized processing views cortical auditory processing as a hierarchical series of feed-forward analysis stages from sensory (acoustic) processing in the PAC to specialized processing of higher-order sound attributes (such as sound location) in higher-level areas. Thus, in this view, sound localization is a higher-order sound attribute.

Evidence for this dual hierarchical organization of auditory processing comes from animal as well as human studies. Single-cell recordings first identified functionally specialized 'what' and 'where' pathways in rhesus monkeys[39,99,100]. Anatomical studies of corticocortical connections in non-human primates provided further support, showing a dorsal stream originating from caudal belt fields and projecting to the posterior parietal cortex (PPC), to eventually end in the dorsolateral prefrontal cortex (dlPFC[103]). In humans, neuroimaging studies revealed a similar functional dissociation between spatial and object identity processing[104-106]. Anatomically, the human homologue of the dorsal spatial processing pathway projects from the planum temporale to the inferior parietal lobule (IPL), premotor cortex (PMC), and finally dlPFC[107-110] – or its equivalent in the inferior frontal cortex (IFC)[110-112]. Finally, lesion studies provide causal evidence for the notion of dual auditory pathways in the ventral and dorsal streams of humans[113,114] and other mammals[115].

In the past few years, functions associated with the dorsal 'where' stream have been extended beyond spatial auditory processing to include auditory motion processing[116-118], temporal processing[119] and sensorimotor functions (reviewed elsewhere[110]). These functions are closely related to spatial processing, and altogether robust evidence across mammalian species indicates the involvement of posterior-dorsal regions in spatial auditory processing. However, later empirical work suggests that a strictly hierarchical, feed-forward notion of spatial auditory processing may be incomplete. Below, we review these findings and discuss their implications for the cortical processing of sound location.

A recurrent model.

Psychophysical studies show that active spatial listening is dynamic. For instance, attention to specific locations leads to more rapid processing of auditory targets at these locations[120-122], and spatial attention facilitates the understanding of speech in the presence of competing but spatially separated speech streams[123]. According to the hierarchical view of auditory processing, the neural mechanisms supporting such dynamic spatial listening are expected to be found in the posterior-dorsal auditory areas that are functionally specialized for sound location processing. However, research shows that the PAC might sustain this dynamic spatial listening[57,58] (see Cortical spatial tuning properties).

Specifically, single-cell recordings in alert and behaving cats revealed that spatial tuning in the PAC sharpens during task performance[58]. Although this effect was also present in 'spatial' areas of the dorsal zone and posterior auditory field, the effect was stronger in the PAC [124]. Similar sharpening of spatial tuning in the PAC during active, goal-oriented sound localization was demonstrated recently with fMRI in humans[57]. Notably, although these effects may have been expected in the planum temporale – the human 'spatial' auditory area – task performance did not modulate spatial tuning in this region. Another fMRI study in humans did not observe similar changes in ILD and ITD selectivity with task performance[55]. However, this study considered average response functions across the entire auditory cortex, which may have diluted localized effects within the PAC [55].

Thus, the combined evidence from animal and human studies provides a first indication that spatial sensitivity in the PAC is flexible and dependent on behavioral demands, and that the PAC is involved in sound location processing during active, goal-oriented localization. These findings tap into a long-standing debate on the functional role of the PAC in the transformation from acoustic processing to the representation of higher-order 'abstract' sound properties that takes place in the auditory system[84]. Even though comparable modulations of neuronal tuning in the PAC by behavioral demands have been reported for other acoustic attributes, these effects typically concerned low-level (acoustic) features. For example, attending to a particular reference sound can induce adaptive changes in spectrotemporal tuning that facilitate target detection[125-127] (reviewed elsewhere[128]).

In the context of spatial processing, it may be argued that neurons in the PAC are selective for elementary spatial cues such as ITDs and ILDs rather than location per se[129-131] (see section on cue integration).In that case, sharpening of neural tuning in the PAC reflects sharpening of responses at a processing level similar to the sharpening of spectrotemporal tuning with attention. Alternatively, following the idea that sound location is a higher-order attribute, sharpening of spatial receptive fields in the PAC during active sound localization seems to call into question the proposed strictly hierarchical nature of cortical spatial processing.

The observed effects in the PAC might be a consequence of feedback projections from higher-order regions that are initiated or strengthened by behavioral or cognitive demands during goal-oriented sound localization (FIG. 3). Such feedback connections have also been incorporated in models of visual processing (for example, recurrent[132] and integrated[133] models, and the 'reverse hierarchy' model for visual and auditory sensory learning[134]). For

auditory spatial processing, dlPFC is a likely candidate for the source of origin of the top-down modulations[135,136]. This region is assumed to translate sensory representations into task-based representations[137].

However, a more intriguing question is what the targets of the frontal feedback projections are. Possibly, these feedback projections are reciprocal to the feedforward projections within the auditory dorsal stream. In this scenario, feedback projections reach the PAC either directly (although no evidence exists for an anatomical connection between PAC and PFC), or indirectly through the connection between posterior-dorsal regions (that is, the caudal belt in non-human primates and the planum temporale in humans) and the PAC[32,34]. An alternative scenario involves corticofugal projections from PFC to subcortical structures, which in turn affect PAC processing[58]. That is, corticofugal projections from PFC can modulate activity in the thalamic reticular nucleus, as has been demonstrated for visual attention[138]. These effects are then propagated as inhibitory modulations to the medial geniculate body[139], subsequently leading to neural response changes in the PAC[140].

In summary, empirical work indicates a stronger involvement of the PAC in spatial auditory processing during active, goal-oriented localization than assumed within the hierarchical framework of cortical spatial auditory processing. This finding suggests that the hierarchical framework needs to be extended by incorporating recurrent connections to accommodate task-dependent modulations of spatial tuning in primary regions. Which form this dynamic cortical auditory network takes is presently not clear, although several hypotheses can be formulated. To test the validity of these hypotheses, functional connectivity studies in humans and non-human primates are needed to measure the task-dependent modulations of interactions between cortical (and subcortical) regions during sound localization. On a smaller scale, laminar electrophysiology[141-143] and laminar fMRI[144-146] can elucidate the cortical micro-circuitry of feedforward and feedback connections involved in dynamic spatial receptive fields and attentional effects.

Is one hemisphere enough?

Another important debate concerning the cortical network for sound location processing is whether both hemispheres are required for accurate sound localization, or only one. Lesion studies typically report contra-lesional localization impairments following unilateral lesions[3,4,114,147,148], suggesting that one hemisphere is enough to accurately localize sounds in contralateral space. However, location decoding studies typically report more accurate location estimates when combining information across hemispheres[21,37,46,57], indicating that combining information from bilateral auditory cortices is optimal. Possibly, these divergent findings are the result of differences in the behavioral state of the listener between these experimental paradigms; localization impairments in lesion studies are measured during active, goal-oriented sound localization, whereas location decoding is most often performed using measurements of neural activity acquired during passive listening.

Specifically, in previous sections we described empirical work showing that the behavioral state of a listener modulates spatial tuning within the auditory cortex (see the section on cortical spatial tuning properties). Similarly, behavioral demands (that is, engaging in active sound localization) might conceivably strengthen the functional connection between the

bilateral auditory cortices, which in turn affects the neural representation of sound location within each hemisphere. Thus, during passive listening the interhemispheric connection is marginal and each hemisphere contains an isolated, sub-optimal representation of predominantly contralateral sound locations. In this case, a post-hoc combination of the neural activity in each hemisphere by a location decoding approach integrates unique spatial information, resulting in more accurate location estimates. This 'bilateral gain' has been demonstrated in population coding studies that measured neural responses during passive listening[46,54].

By contrast, during active, goal-oriented sound localization, the neural representation of sound location becomes more precise within each hemisphere (that is, sharper spatial tuning)[57,58], resulting in accurate contralateral sound localization (in line with lesion studies[3,4,114,147,148]). In addition, the stronger functional connection between bilateral auditory cortices during active localization leads to interhemispheric information exchange, resulting in comparable spatial representations. A post-hoc combination of neural activity across hemispheres in a decoding approach will therefore not necessarily lead to better location estimates. Importantly, the results of a decoding study in humans confirm this pattern: during a non-localization task each hemisphere contained complementary information, while the information was redundant during an active sound localization task[57].

Thus, the behavioral state of a listener appears to affect the cortical network of auditory spatial processing in multiple ways: in terms of feedback connections from frontal regions to cortical auditory regions, and in terms of interhemispheric connections. We are only beginning to unravel the nature of these dynamic mechanisms in the cortical network for sound location processing, and more research on the functional connections within this network – in varying behavioral conditions – is required.

Adaptive mechanisms

The neural processing of sound location is not only shaped by the behavioral state of the listener, but also by the input to the auditory system. This effect is demonstrated clearly by research into sound localization in unilateral hearing loss (reviewed in detail elsewhere[149]). Specifically, monaural deprivation can have a detrimental effect on spatial hearing[150-152], but behavioral data shows that adaptive mechanisms can mediate the effects of a change in input and – partly – restore spatial hearing[149]. These adaptive mechanisms include the reweighting of spatial cues (resulting in an increased reliance on spectral cues for horizontal localization[153-157]), and the remapping of binaural cues to correct for the altered input to the two ears in case of asymmetric hearing loss (that is, with some remaining input to the deprived ear)[157,158].

Single-cell recordings in ferrets[156-158] and rats[159] indicate that these compensatory mechanisms are the result of neural plasticity in the PAC (although adaptive changes at other sites may also be of relevance[160]). Interestingly, both compensatory mechanisms are represented in the PAC: a subpopulation of neurons exhibits enhanced sensitivity to spectral cues[156], while a separate distinct subpopulation exhibits remapping of ILD sensitivity[157,158]. Importantly, the latter finding suggests that neural tuning to binaural spatial cues is a

function of the input to the auditory system (hence the shift in sensitivity in monaural deprivation), but that the underlying neural coding principles for binaural sound localization (that is opponent coding) are robust to changes in input.

Cue integration

Until now, we have described cortical mechanisms for spatial hearing mostly in terms of processing of 'sound location'. However, to arrive at a perception of sound location, the auditory system has to combine information from different types of spatial cues: ITDs, ILDs and spectral cues (see the section on mammalian spatial hearing). Given that ITDs and ILDs are processed in anatomically distinct pathways in the brainstem, at least until the level of the inferior colliculus[161] (Box 2), such cue integration might conceivably start either at the inferior colliculus or at the cortex. In fact, cue integration might be one of the most important contributions of the auditory cortex to spatial auditory processing. Whereas most of the studies described so far employed stimuli that contained a mixture of spatial cues, other lines of research have focused on unraveling cortical cue integration with artificially spatialized stimuli in which ITDs and ILDs can be manipulated in isolation.

The results of such approaches are equivocal. Some studies report that ITDs and ILDs are processed in overlapping cortical regions[55,56,162]. These findings are indicative of an abstract, cue-independent representation of sound location in the cortex. Further evidence for integrated processing comes from magnetoencephalography measurements demonstrating location-dependent neural adaptation irrespective of the type of spatial cue of the probe and adaptor sound[163]. Moreover, cortical fMRI activity patterns in response to ITDs are similar to those in response to ILDs, even to the extent that location can be decoded by a classifier across cues[55]. However, other studies have shown distinct topographies and time-courses for each binaural cue[129-131], suggesting that ITDs and ILDs are processed in parallel (possibly interacting) cortical networks. The findings of a lesion study in humans also contradict the idea of fully integrated processing: differential patterns of localization errors were observed for sounds spatialized with ITDs or ILDs after brain damage (although considerable variety existed in the extent and onset of lesions between participants)[5].

How can these discrepancies in experimental results be reconciled? Possibly, neural encoding of sound location is a two-stage process that starts from lower-order representations of individual binaural cues that later converge to a generalized representation of sound location. Similar frameworks have been proposed for the encoding of other sound attributes. For example, the auditory system has been proposed to extract pitch information from two different mechanisms (a temporal and a spectral mechanism) that may converge later in a cortical 'pitch center'[164,165] to arrive at a generalized representation[166] (although the notion of a pitch center is debated, reviewed elsewhere[167]). However, if this two-stage process for encoding of sound location is correct, where and how the convergence of ITD and ILD processing occurs in the cortex is unclear. Thus, the cortical contribution to cue integration and the emergence of an abstract perception of sound location are important areas for future research.

Sound localization in complex scenes

In this review, we have – so far – discussed cortical mechanisms for the encoding of absolute sound location. However, in daily life, listeners are typically presented with complex auditory scenes consisting of multiple, spatially separated sound sources (FIG. 4). Such listening environments introduce additional difficulties for absolute sound localization, but simultaneously provide new information to the auditory system in terms of *relative sound location*. In this final section, we review empirical work addressing these topics concerning spatial hearing in complex auditory scenes.

Interfering sounds.

Multi-source listening environments result in distorted and fluctuating binaural cues[20]. Nevertheless, humans are often still able to localize a target accurately[168-171] (although one study showed that the presence of a distractor shifts the perceived location of the target towards the distractor[172]). Although cortical top-down mechanisms such as *spatial selective attention* can help to reduce the confusion between a target and interfering sound sources[173], this mechanism does not solve the inherent ambiguity of the binaural cues. Accordingly, the presentation of an interfering, spatially separated sound distorts azimuth response functions in the auditory pathway of cats[174] and rabbits[20]. Given this ambiguity in the neural response, how does the auditory system derive the spatial position of a target?

One possible strategy – considering the relatively sparse temporal and spectral character of real-life sounds – is to localize the target in the short periods in which one of the sources is present in relative isolation[20,175] (Fig. 4). This approach effectively removes the binaural ambiguity and reduces the computation to that of single-source localization. However, how the brain determines whether one or more sound sources are simultaneously present is not clear. It is possible that the auditory system uses the degree of interaural decorrelation (for example interaural coherence) as a criterion[175,176]. Interaural coherence decreases in the presence of multiple, spatially separated sound sources, and ITD sensitive neurons in the inferior colliculus and auditory cortex of alert rabbits typically exhibit sensitivity to interaural coherence as well[87,177]. However, no direct evidence yet exists to support this hypothesis, and the mechanisms underlying localization in the presence of an interfering, spatially separated sound source are not well understood.

Spatial processing for scene analysis.

A related topic is the neural encoding of relative sound location, which contains useful information for the auditory system. For instance, spatial separation can contribute to the grouping of incoming soundwaves into coherent auditory objects, a process known as auditory scene analysis[178]. In this context, spatial cues seem to be especially relevant for grouping (streaming) of interleaved sound sequences over time[173,179] (although experimental evidence is equivocal[180,181]).

Neurophysiological measurements show that *auditory stream segregation* on the basis of acoustic properties such as sound frequency involves the PAC[182-185]. However, research into spatial stream segregation is very limited. A study in cats showed that spatial stream

segregation for sequential, interleaved streams can be predicted from neural activity in the PAC using a model of neural spatial sensitivity to isolated sound sources, in combination with an attenuation factor representing the observed decrease in neuronal response in the presence of a competing sound source[186]. Thus, the neural representation of spatial stream segregation in cat PAC appeared to be contingent on the representation of absolute sound location (FIG. 4).

By contrast, psychophysical results in humans point in the direction of distinct location and separation processing[187]. These diverging results might – once more – be traced to differences in the behavioral state of the listener: the human listeners were alert and performing a stream segregation task, whereas the single-cell recordings were made in anesthetized cats. Possibly, the representation of spatial separation is more distinct from absolute location processing during active spatial streaming than during passive listening. To test this hypothesis, neurophysiological research examining the relationship between absolute and relative sound location processing in different behavioral conditions is required. Such research would also advance our understanding of the relative contribution of primary and higher-order auditory regions (especially the 'computational hub', that is, the planum temporale[188,189]) to spatial segregation, a question that has not been addressed directly by existing studies.

Conclusions

Novel research paradigms using alert and responsive subjects, real-life sounds, and advanced computational modeling approaches have contributed significantly to our understanding of the complex computational mechanisms underlying spatial hearing. Evidence is growing that the hierarchical model for location processing might need to be extended to include recurrent (and possibly even interhemispheric connections) to accommodate goal-oriented sound localization. Furthermore, the notion of a neural representation of sound location as population pattern code is gaining momentum, even if the transformation from binaural sound to neural response is still poorly understood. Additionally, an important note of caution for cross-species comparisons has emerged from insights into the influence of head size on IPD encoding in mammals. In terms of the role of the auditory cortex in spatial hearing, the research discussed in this Review highlights that the auditory cortical network is especially relevant for spatial processing in the context of behavioral goals, and for spatial processing of complex, real-life sounds and in multi-source auditory scenes.

Amid these advancing insights, several important questions remain unresolved. For example, a mechanistic understanding of the integration of different spatial cues, or the localization of sounds in multi-source settings, is still lacking. Knowledge of the interaction between neural processing of sound location and other sound attributes is equally minimal. For future research directions, the work discussed here emphasizes that experimental set-ups using real-life, complex sounds in ecologically valid listening scenes are required in order to gain a better understanding of the full complexity of cortical sound location processing. Merging computational modeling strategies with neurophysiological measurements will provide significant support to these research efforts.

ACKNOWLEDGEMENTS

GLOSSARY

Front-back ambiguities:
Humans can have difficulty distinguishing whether a sound source is located behind or in front of them because the interaural time and level differences are identical for sound sources at the same angular position with respect to the interaural midline yet located in the front and back

Coincidence detectors:
Neuron whose firing rate is modulated by the time of arrival of input from two lower-level neurons, such that it responds maximally when the input arrives simultaneously

Sparse coding:
A neural coding strategy in which single neurons encode sensory stimuli efficiently by representing the maximal amount of information possible (thereby saving computational resources), and neuronal populations consist of neurons that encode unique information (that is, neural responses are independent)

Opponent coding:
A neural representational mechanism in which sensory stimuli are represented by the integrated activity of two neuronal populations tuned to opposite values of the characteristic under consideration (for sound location: the integrated activity of a population tuned to the left and a population tuned to the right hemifield)

Relative sound location:
In multi-source listening environments, the relative sound location refers to the location of the individual sound sources with respect to each other, that is, the spatial separation

Spatial selective attention:
The attentional focus of a listener on a particular location and the sounds presented at this location, while ignoring sounds at other locations

Auditory stream segregation:
The segregation and grouping of concurrent or interleaved sound streams in multi-source listening environments into their respective sound sources

REFERENCES

1. Heffner H & Masterton B Contribution of auditory cortex to sound localization in the monkey (Macaca mulatta). Journal of Neurophysiology 38, 1340–1358 (1975). [PubMed: 815514]

2. Heffner HE & Heffner RS Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. Journal of Neurophysiology 64, 915–931 (1990). [PubMed: 2230934]

3. Malhotra S, Hall AJ & Lomber SG Cortical control of sound localization in the cat: unilateral cooling deactivation of 19 cerebral areas. Journal of Neurophysiology 92, 1625–1643 (2004). [PubMed: 15331649]

4. Malhotra S, Stecker GC, Middlebrooks JC & Lomber SG Sound localization deficits during reversible deactivation of primary auditory cortex and/or the dorsal zone. Journal of Neurophysiology 99, 1628–1642 (2008). [PubMed: 18199813]

5. Spierer L, Bellmann-Thiran A, Maeder P, Murray MM & Clarke S Hemispheric competence for auditory spatial representation. Brain 132, 1953–1966 (2009). [PubMed: 19477962]

6. Zatorre RJ & Penhune VB Spatial localization after excision of human auditory cortex. Journal of Neuroscience 21, 6321–6328 (2001). [PubMed: 11487655]

7. Zündorf IC, Karnath H-O & Lewald J The effect of brain lesions on sound localization in complex acoustic environments. Brain 137, 1410–1418 (2014). [PubMed: 24618271]

8. Hofman PM, Van Riswick JG & Van Opstal AJ Relearning sound localization with new ears. Nature neuroscience 1, 417 (1998). [PubMed: 10196533]

9. Oldfield SR & Parker SP Acuity of sound localisation: a topography of auditory space. I. Normal hearing conditions. Perception 13, 581–600 (1984). [PubMed: 6535983]

10. Rayleigh Lord. On our perception of sound direction. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 13, 214–232 (1907).

11. Blauert J Sound localization in the median plane. Acta Acustica united with Acustica 22, 205–213 (1969).

12. Mills AW Lateralization of high-frequency tones. The Journal of the Acoustical Society of America 32, 132–134 (1960).

13. Sandel T, Teas D, Feddersen W & Jeffress L Localization of sound from single and paired sources. the Journal of the Acoustical Society of America 27, 842–852 (1955).

14. Stevens SS & Newman EB The localization of actual sources of sound. The American journal of psychology 48, 297–306 (1936).

15. Wightman FL & Kistler DJ The dominant role of low-frequency interaural time differences in sound localization. The Journal of the Acoustical Society of America 91, 1648–1661 (1992). [PubMed: 1564201]

16. Bernstein LR & Trahiotis C Lateralization of sinusoidally amplitude-modulated tones: Effects of spectral locus and temporal variation. The Journal of the Acoustical Society of America 78, 514–523 (1985). [PubMed: 4031250]

17. Hartmann WM, Rakerd B, Crawford ZD & Zhang PX Transaural experiments and a revised duplex theory for the localization of low-frequency tones. The Journal of the Acoustical Society of America 139, 968–985 (2016). [PubMed: 26936576]

18. Henning GB Detectability of interaural delay in high-frequency complex waveforms. The Journal of the Acoustical Society of America 55, 84–90 (1974). [PubMed: 4815755]

19. Rakerd B & Hartmann W Localization of sound in rooms, II: The effects of a single reflecting surface. The Journal of the Acoustical Society of America 78, 524–533 (1985). [PubMed: 4031251]

20. Day ML, Koka K & Delgutte B Neural encoding of sound source location in the presence of a concurrent, spatially separated source. Journal of neurophysiology 108, 2612–2628 (2012). [PubMed: 22914651]

21. McAlpine D, Jiang D & Palmer AR A neural code for low-frequency sound localization in mammals. Nature neuroscience 4, 396–401 (2001). [PubMed: 11276230]

22. Jones HG, Brown AD, Koka K, Thornton JL & Tollin DJ Sound frequency-invariant neural coding of a frequency-dependent cue to sound source location. Journal of neurophysiology 114, 531–539 (2015). [PubMed: 25972580]

23. Makous JC & Middlebrooks JC Two-dimensional sound localization by human listeners. The journal of the Acoustical Society of America 87, 2188–2200 (1990). [PubMed: 2348023]

24. Heffner RS & Heffner HE Sound localization acuity in the cat: effect of azimuth, signal duration, and test procedure. Hearing research 36, 221–232 (1988). [PubMed: 3209494]

25. Butler RA The bandwidth effect on monaural and binaural localization. Hearing research 21, 67–73 (1986). [PubMed: 3957797]

26. Carlile S, Delaney S & Corderoy A The localisation of spectrally restricted sounds by human listeners. Hearing research 128, 175–189 (1999). [PubMed: 10082297]

27. Tollin DJ, Ruhland JL & Yin TC The role of spectral composition of sounds on the localization of sound sources by cats. Journal of Neurophysiology 109, 1658–1668 (2012). [PubMed: 23274314]

28. Musicant AD & Butler RA Influence of monaural spectral cues on binaural localization. The Journal of the Acoustical Society of America 77, 202–208 (1985). [PubMed: 3973215]

29. Van Wanrooij MM & Van Opstal AJ Sound localization under perturbed binaural hearing. Journal of neurophysiology 97, 715–726 (2007). [PubMed: 17065242]

30. Voss P, Tabry V & Zatorre RJ Trade-off in the sound localization abilities of early blind individuals between the horizontal and vertical planes. Journal of Neuroscience 35, 6051–6056 (2015). [PubMed: 25878278]

31. Derey K, Rauschecker JP, Formisano E, Valente G & de Gelder B Localization of complex sounds is modulated by behavioral relevance and sound category. The Journal of the Acoustical Society of America 142, 1757–1773 (2017). [PubMed: 29092572]

32. Hackett TA et al. Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. Frontiers in neuroscience 8, 72 (2014). [PubMed: 24795550]

33. Kaas JH & Hackett TA Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences 97, 11793–11799 (2000).

34. Moerel M, De Martino F & Formisano E An anatomical and functional topography of human auditory cortical areas. Frontiers in neuroscience 8, 225 (2014). [PubMed: 25120426]

35. Rauschecker JP, Tian B & Hauser M Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268, 111–114 (1995). [PubMed: 7701330]

36. Benson D, Hienz R & Goldstein M Jr Single-unit activity in the auditory cortex of monkeys actively localizing sound sources: spatial tuning and behavioral dependency. Brain research 219, 249–267 (1981). [PubMed: 7260632]

37. Ortiz-Rios M et al. Widespread and Opponent fMRI Signals Represent Sound Location in Macaque Auditory Cortex. Neuron 93, 971–983. e4 (2017). [PubMed: 28190642] A macaque neuroimaging study providing evidence for an opponent coding mechanism for the neural representation of sound location.

38. Recanzone GH, Guard DC, Phan ML & Su T-IK Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. Journal of neurophysiology 83, 2723–2739 (2000). [PubMed: 10805672]

39. Tian B, Reser D, Durham A, Kustov A & Rauschecker JP Functional specialization in rhesus monkey auditory cortex. Science 292, 290–293 (2001). [PubMed: 11303104]

40. Woods TM, Lopez SE, Long JH, Rahman JE & Recanzone GH Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. Journal of neurophysiology 96, 3323–3337 (2006). [PubMed: 16943318]

41. Ku mierek P & Rauschecker JP Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. Journal of neurophysiology 111, 1671–1685 (2014). [PubMed: 24501260]

42. Brugge JF, Reale RA & Hind JE The structure of spatial receptive fields of neurons in primary auditory cortex of the cat. Journal of Neuroscience 16, 4420–4437 (1996). [PubMed: 8699253]

43. Harrington IA, Stecker GC, Macpherson EA & Middlebrooks JC Spatial sensitivity of neurons in the anterior, posterior, and primary fields of cat auditory cortex. Hearing research 240, 22–41 (2008). [PubMed: 18359176]

44. Mrsic-Flogel TD, King AJ & Schnupp JW Encoding of virtual acoustic space stimuli by neurons in ferret primary auditory cortex. Journal of neurophysiology 93, 3489–3503 (2005). [PubMed: 15659534]

45. Rajan R, Aitkin L, Irvine D & McKay J Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. Journal of Neurophysiology 64, 872–887 (1990). [PubMed: 2230931]

46. Stecker GC, Harrington IA & Middlebrooks JC Location coding by opponent neural populations in the auditory cortex. PLoS biology 3, e78 (2005). [PubMed: 15736980] Extracellular recordings in cat auditory cortical regions demonstrating the validity of opponent population coding for the neural representation of sound location.

47. Stecker GC & Middlebrooks JC Distributed coding of sound locations in the auditory cortex. Biological cybernetics 89, 341–349 (2003). [PubMed: 14669014]

48. Mickey BJ & Middlebrooks JC Representation of auditory space by cortical neurons in awake cats. Journal of Neuroscience 23, 8649–8663 (2003). [PubMed: 14507964]

49. Brunetti M et al. Human brain activation during passive listening to sounds from different locations: an fMRI and MEG study. Human brain mapping 26, 251–261 (2005). [PubMed: 15954141]

50. Deouell LY, Heller AS, Malach R, D'Esposito M & Knight RT Cerebral responses to change in spatial location of unattended sounds. Neuron 55, 985–996 (2007). [PubMed: 17880900]

51. Van der Zwaag W, Gentile G, Gruetter R, Spierer L & Clarke S Where sound position influences sound object representations: a 7-T fMRI study. Neuroimage 54, 1803–1811 (2011). [PubMed: 20965262]

52. Warren JD & Griffiths TD Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. Journal of Neuroscience 23, 5799–5804 (2003). [PubMed: 12843284]

53. Krumbholz K et al. Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. Cerebral Cortex 15, 317–324 (2004). [PubMed: 15297367]

54. Derey K, Valente G, de Gelder B & Formisano E Opponent Coding of Sound Location (Azimuth) in Planum Temporale is Robust to Sound-Level Variations. Cerebral Cortex 26, 450–464 (2015). [PubMed: 26545618] A neuroimaging study in humans demonstrating the validity of an opponent coding mechanism for sound location encoding in planum temporale.

55. Higgins NC, McLaughlin SA, Rinne T & Stecker GC Evidence for cue-independent spatial representation in the human auditory cortex during active listening. Proceedings of the National Academy of Sciences 114, E7602–E7611 (2017).An fMRI study with an active task paradigm and a multivariate pattern analysis approach, showing that human auditory cortex represents sound location independent of the type of binaural cue (that is, ILD or ITD).

56. McLaughlin SA, Higgins NC & Stecker GC Tuning to binaural cues in human auditory cortex. Journal of the Association for Research in Otolaryngology 17, 37–53 (2016). [PubMed: 26466943]

57. van der Heijden K, Rauschecker JP, Formisano E, Valente G & de Gelder B Active sound localization sharpens spatial tuning in human primary auditory cortex. Journal of Neuroscience 38, 0587–0518 (2018).An fMRI study measuring spatial selectivity in human auditory cortex in varying behavioral conditions, demonstrating sharpening of spatial tuning during active sound localization and accurate decoding of sound location using a maximum likelihood approach.

58. Lee C-C & Middlebrooks JC Auditory cortex spatial sensitivity sharpens during task performance. Nature neuroscience 14, 108 (2011). [PubMed: 21151120] Single- and multi-unit recordins in primary auditory cortex of awake and behaving cats show that neural spatial sensitivity sharpens during task performance.

59. Harper NS, Scott BH, Semple MN & McAlpine D The neural code for auditory space depends on sound frequency and head size in an optimal manner. PloS one 9, e108154 (2014). [PubMed:

25372405] Empirical evidence from neurophysiological data in various species supporting the optimal coding model for sound location based on head size and sound frequency.

60. Chi T, Ru P & Shamma SA Multiresolution spectrotemporal analysis of complex sounds. The Journal of the Acoustical Society of America 118, 887–906 (2005). [PubMed: 16158645]

61. Santoro R et al. Encoding of natural sounds at multiple spectral and temporal resolutions in the human auditory cortex. PLoS computational biology 10, e1003412 (2014). [PubMed: 24391486]

62. Santoro R et al. Reconstructing the spectrotemporal modulations of real-life sounds from fMRI response patterns. Proceedings of the National Academy of Sciences 114, 4799–4804 (2017).

63. Erb J et al. Homology and Specificity of Natural Sound-Encoding in Human and Monkey Auditory Cortex. Cerebral Cortex 10.1093/cercor/bhy243 (2018).

64. Jeffress LA A place theory of sound localization. Journal of comparative and physiological psychology 41, 35 (1948). [PubMed: 18904764]

65. Grothe B, Pecka M & McAlpine D Mechanisms of sound localization in mammals. Physiological reviews 90, 983–1012 (2010). [PubMed: 20664077]

66. Carr C & Konishi M A circuit for detection of interaural time differences in the brain stem of the barn owl. Journal of Neuroscience 10, 3227–3246 (1990). [PubMed: 2213141]

67. Knudsen EI & Konishi M A neural map of auditory space in the owl. Science 200, 795–797 (1978). [PubMed: 644324]

68. Knudsen EI, Konishi M & Pettigrew JD Receptive fields of auditory neurons in the owl. Science 198, 1278–1280 (1977). [PubMed: 929202]

69. Palmer A & King A The representation of auditory space in the mammalian superior colliculus. Nature 299, 248 (1982). [PubMed: 7110344]

70. Młynarski W The opponent channel population code of sound location is an efficient representation of natural binaural sounds. PLoS computational biology 11, e1004294 (2015). [PubMed: 25996373] A computational model describing the transformation of binaural sound into a neural representation of spectrotemporal and interaural features based on sparse coding of statistical regularities in the environment.

71. Barlow HB Possible principles underlying the transformations of sensory messages. Sens. Commun 1, 217 – 234 (1961).

72. Simoncelli EP & Olshausen BA Natural image statistics and neural representation. Annual review of neuroscience 24, 1193–1216 (2001).

73. Olshausen BA & Field DJ Sparse coding of sensory inputs. Current opinion in neurobiology 14, 481–487 (2004). [PubMed: 15321069]

74. Singh NC & Theunissen FE Modulation spectra of natural sounds and ethological theories of auditory processing. The Journal of the Acoustical Society of America 114, 3394–3411 (2003). [PubMed: 14714819]

75. Harper NS & McAlpine D Optimal neural population coding of an auditory spatial cue. Nature 430, 682 (2004). [PubMed: 15295602] A neural modeling study of sound location encoding in species with different head sizes, showing that the optimal coding strategy for IPD depends on head size and sound frequency.

76. Köppl C Phase locking to high frequencies in the auditory nerve and cochlear nucleus magnocellularis of the barn owl, Tyto alba. Journal of Neuroscience 17, 3312–3321 (1997). [PubMed: 9096164]

77. Sullivan W & Konishi M Segregation of stimulus phase and intensity coding in the cochlear nucleus of the barn owl. Journal of Neuroscience 4, 1787–1799 (1984). [PubMed: 6737041]

78. Verschooten E, Desloovere C & Joris PX High-resolution frequency tuning but not temporal coding in the human cochlea. PLoS biology 16, e2005164 (2018). [PubMed: 30321166]

79. Kay KN, Naselaris T, Prenger RJ & Gallant JL Identifying natural images from human brain activity. Nature 452, 352 (2008). [PubMed: 18322462]

80. Miyawaki Y et al. Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 60, 915–929 (2008). [PubMed: 19081384]

81. Kriegeskorte N et al. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60, 1126–1141 (2008). [PubMed: 19109916]

82. Moerel M, De Martino F & Formisano E Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity. Journal of Neuroscience 32, 14205–14216 (2012). [PubMed: 23055490]

83. De Martino F et al. The impact of ultra-high field MRI on cognitive and computational neuroimaging. Neuroimage 168, 366–382 (2018). [PubMed: 28396293]

84. Bizley JK & Cohen YE The what, where and how of auditory-object perception. Nature Reviews Neuroscience 14, 693 (2013). [PubMed: 24052177]

85. Bizley JK, Walker KM, Silverman BW, King AJ & Schnupp JW Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. Journal of neuroscience 29, 2064–2075 (2009). [PubMed: 19228960] Neurophysiological recordings in ferret auditory cortex show that most units are sensitive to multiple perceptual attributes, indicating that interdependent encoding may be an underlying mechanism for the unified representation of acoustic and location features.

86. Hancock KE & Delgutte B A physiologically based model of interaural time difference discrimination. Journal of Neuroscience 24, 7110–7117 (2004). [PubMed: 15306644]

87. Day ML & Delgutte B Decoding sound source location and separation using neural population activity patterns. Journal of Neuroscience 33, 15837–15847 (2013). [PubMed: 24089491]

88. Dingle RN, Hall SE & Phillips DP The three-channel model of sound localization mechanisms: interaural level differences. The Journal of the Acoustical Society of America 131, 4023–4029 (2012). [PubMed: 22559375]

89. Dingle RN, Hall SE & Phillips DP The three-channel model of sound localization mechanisms: Interaural time differences. The Journal of the Acoustical Society of America 133, 417–424 (2013). [PubMed: 23297913]

90. Magezi DA & Krumbholz K Evidence for opponent-channel coding of interaural time differences in human auditory cortex. Journal of neurophysiology 104, 1997–2007 (2010). [PubMed: 20702739]

91. Salminen NH, May PJ, Alku P & Tiitinen H A population rate code of auditory space in the human cortex. PLoS One 4, e7600 (2009). [PubMed: 19855836]

92. Phillips DP A perceptual architecture for sound lateralization in man. Hearing research 238, 124–132 (2008). [PubMed: 17980984]

93. Phillips DP, Carmichael ME & Hall SE Interaction in the perceptual processing of interaural time and level differences. Hearing research 211, 96–102 (2006). [PubMed: 16309863]

94. Georgopoulos AP, Schwartz AB & Kettner RE Neuronal population coding of movement direction. Science 233, 1416–1419 (1986). [PubMed: 3749885]

95. Georgopoulos AP, Kettner RE & Schwartz AB Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. Journal of Neuroscience 8, 2928–2937 (1988). [PubMed: 3411362]

96. Jazayeri M & Movshon JA Optimal representation of sensory information by neural populations. Nature neuroscience 9, 690 (2006). [PubMed: 16617339]

97. Miller LM & Recanzone GH Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. Proceedings of the National Academy of Sciences 106, 5931–5935 (2009). Electrophysiological recordings in macaque auditory cortical areas show that a population code based on maximum likelihood estimation can encode sound location from neural firing rate patterns in the caudolateral region, but not in other auditory cortical regions.

98. Pouget A, Dayan P & Zemel R Information processing with population codes. Nature Reviews Neuroscience 1, 125–132 (2000). [PubMed: 11252775]

99. Rauschecker JP Parallel processing in the auditory cortex of primates. Audiology and Neurotology 3, 86–103 (1998). [PubMed: 9575379]

100. Rauschecker JP & Tian B Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proceedings of the National Academy of Sciences 97, 11800–11806 (2000).

101. Goodale MA & Milner AD Separate visual pathways for perception and action. Trends in neurosciences 15, 20–25 (1992). [PubMed: 1374953]

102. Mishkin M, Ungerleider LG & Macko KA Object vision and spatial vision: two cortical pathways. Trends in neurosciences 6, 414–417 (1983).

103. Romanski LM et al. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nature neuroscience 2, 1131–1136 (1999). [PubMed: 10570492]

104. Alain C, Arnott SR, Hevenor S, Graham S & Grady CL "What" and "where" in the human auditory system. Proceedings of the National Academy of Sciences 98, 12301–12306 (2001).

105. Arnott SR, Binns MA, Grady CL & Alain C Assessing the auditory dual-pathway model in humans. Neuroimage 22, 401–408 (2004). [PubMed: 15110033]

106. Maeder PP et al. Distinct pathways involved in sound recognition and localization: a human fMRI study. Neuroimage 14, 802–816 (2001). [PubMed: 11554799]

107. Lewald J, Riederer KA, Lentz T & Meister IG Processing of sound location in human cortex. European Journal of Neuroscience 27, 1261–1270 (2008). [PubMed: 18364040]

108. Weeks R et al. A positron emission tomographic study of auditory localization in the congenitally blind. Journal of Neuroscience 20, 2664–2672 (2000). [PubMed: 10729347]

109. Zatorre RJ, Bouffard M, Ahad P & Belin P Where is' where'in the human auditory cortex? Nature neuroscience 5, 905–909 (2002). [PubMed: 12195426]

110. Rauschecker JP Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. Cortex 98, 262 – 268 (2018). [PubMed: 29183630]

111. Rauschecker JP & Scott SK Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nature neuroscience 12, 718–724 (2009). [PubMed: 19471271]

112. Bushara KO et al. Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nature neuroscience 2, 759–766 (1999). [PubMed: 10412067] A human neuroimaging approach revealing the functional specialization of lateral prefrontal cortex for spatial auditory processing.

113. Ahveninen J et al. Evidence for distinct human auditory cortex regions for sound location versus identity processing. Nature communications 4, 2585 (2013).

114. Clarke S, Bellmann A, Meuli RA, Assal G & Steck AJ Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. Neuropsychologia 38, 797–807 (2000). [PubMed: 10689055]

115. Lomber SG & Malhotra S Double dissociation of'what'and'where'processing in auditory cortex. Nature neuroscience 11, 609–616 (2008). [PubMed: 18408717] A lesion study using reversible cooling in cats, providing evidence for the existence of a specialized posterior-dorsal processing stream for spatial auditory processing, as well as a ventral stream for object identity processing.

116. Poirier C et al. Auditory motion-specific mechanisms in the primate brain. PLoS biology 15, e2001379 (2017). [PubMed: 28472038]

117. Krumbholz K et al. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. European Journal of Neuroscience 21, 230–238 (2005). [PubMed: 15654860]

118. Griffiths TD et al. Right parietal cortex is involved in the perception of sound movement in humans. Nature neuroscience 1, 74–79 (1998). [PubMed: 10195113]

119. Jasmin K, Lima CF & Scott SK Understanding rostral–caudal auditory cortex contributions to auditory perception. Nature Reviews Neuroscience, 1–10 (2019). [PubMed: 30518965]

120. Mondor TA & Zatorre RJ Shifting and focusing auditory spatial attention. Journal of Experimental Psychology: Human Perception and Performance 21, 387 (1995). [PubMed: 7714479]

121. Rorden C & Driver J Spatial deployment of attention within and across hemifields in an auditory task. Experimental Brain Research 137, 487–496 (2001). [PubMed: 11355393]

122. Spence CJ & Driver J Covert spatial orienting in audition: Exogenous and endogenous mechanisms. Journal of Experimental Psychology: Human Perception and Performance 20, 555 (1994).

123. Ruggles D, Bharadwaj H & Shinn-Cunningham BG Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. Proceedings of the National Academy of Sciences 108, 15516–15521 (2011).

124. Lee C-C & Middlebrooks JC Specialization for sound localization in fields A1, DZ, and PAF of cat auditory cortex. Journal of the Association for Research in Otolaryngology 14, 61–82 (2013). [PubMed: 23180228]

125. Fritz J, Shamma S, Elhilali M & Klein D Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nature neuroscience 6, 1216–1223 (2003). [PubMed: 14583754]

126. Fritz JB, Elhilali M, David SV & Shamma SA Auditory attention—focusing the searchlight on sound. Current opinion in neurobiology 17, 437–455 (2007). [PubMed: 17714933]

127. Atiani S, Elhilali M, David SV, Fritz JB & Shamma SA Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. Neuron 61, 467–480 (2009). [PubMed: 19217382]

128. King AJ, Teki S & Willmore BD Recent advances in understanding the auditory cortex. F1000Research 7 (2018).

129. Schroger E Interaural time and level differences: integrated or separated processing? Hearing research 96, 191–198 (1996). [PubMed: 8817318]

130. Tardif E, Murray MM, Meylan R, Spierer L & Clarke S The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. Brain research 1092, 161–176 (2006). [PubMed: 16684510]

131. Ungan P, Yagcioglu S & Goksoy C Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. Clinical Neurophysiology 112, 485–498 (2001). [PubMed: 11222971]

132. Lamme VA & Roelfsema PR The distinct modes of vision offered by feedforward and recurrent processing. Trends in neurosciences 23, 571–579 (2000). [PubMed: 11074267]

133. Bullier J Integrated model of visual processing. Brain research reviews 36, 96–107 (2001). [PubMed: 11690606]

134. Ahissar M, Nahum M, Nelken I & Hochstein S Reverse hierarchies and sensory learning. Philosophical Transactions of the Royal Society of London B: Biological Sciences 364, 285–299 (2009). [PubMed: 18986968]

135. Fritz JB, David SV, Radtke-Schuller S, Yin P & Shamma SA Adaptive, behaviorally gated, persistent encoding of task-relevant auditory information in ferret frontal cortex. Nature neuroscience 13, 1011–1019 (2010). [PubMed: 20622871]

136. Jiang X, Chevillet MA, Rauschecker JP & Riesenhuber M Training Humans to Categorize Monkey Calls: Auditory Feature-and Category-Selective Neural Tuning Changes. Neuron 98, 405–416. e404 (2018). [PubMed: 29673483]

137. Goldman-Rakic PS, Cools A & Srivastava K The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. Philosophical Transactions: Biological Sciences, 1445–1453 (1996). [PubMed: 8941956]

138. McAlonan K, Cavanaugh J & Wurtz RH Attentional modulation of thalamic reticular neurons. Journal of Neuroscience 26, 4444–4450 (2006). [PubMed: 16624964]

139. Cotillon-Williams N, Huetz C, Hennevin E & Edeline J-M Tonotopic control of auditory thalamus frequency tuning by reticular thalamic neurons. Journal of neurophysiology 99, 1137–1151 (2008). [PubMed: 18160422]

140. Ma X & Suga N Specific and nonspecific plasticity of the primary auditory cortex elicited by thalamic auditory neurons. Journal of Neuroscience 29, 4888–4896 (2009). [PubMed: 19369557]

141. Atencio CA & Schreiner CE Laminar diversity of dynamic sound processing in cat primary auditory cortex. Journal of neurophysiology 103, 192–205 (2009). [PubMed: 19864440]

142. Sakata S & Harris KD Laminar structure of spontaneous and sensory-evoked population activity in auditory cortex. Neuron 64, 404–418 (2009). [PubMed: 19914188]

143. Winkowski DE & Kanold PO Laminar transformation of frequency organization in auditory cortex. Journal of Neuroscience 33, 1498–1508 (2013). [PubMed: 23345224]

144. De Martino F et al. Frequency preference and attention effects across cortical depths in the human primary auditory cortex. Proceedings of the National Academy of Sciences 112, 16036–16041 (2015).

145. Kok P, Bains LJ, van Mourik T, Norris DG & de Lange FP Selective activation of the deep layers of the human primary visual cortex by top-down feedback. Current Biology 26, 371–376 (2016). [PubMed: 26832438]

146. Lawrence SJ, Formisano E, Muckli L & de Lange FP Laminar fMRI: applications for cognitive neuroscience. Neuroimage 197, 785 – 791 (2017 [PubMed: 28687519]

147. Jenkins WM & Masterton RB Sound localization: effects of unilateral lesions in central auditory system. Journal of Neurophysiology 47, 987–1016 (1982). [PubMed: 7108581]

148. Jenkins WM & Merzenich MM Role of cat primary auditory cortex for sound-localization behavior. Journal of Neurophysiology 52, 819–847 (1984). [PubMed: 6512590]

149. Kumpik DP & King AJ A review of the effects of unilateral hearing loss on spatial hearing. Hearing research 372, 17 – 28 (2018). [PubMed: 30143248]

150. Moore DR et al. Conductive hearing loss produces a reversible binaural hearing impairment. Journal of Neuroscience 19, 8704–8711 (1999). [PubMed: 10493771]

151. Wilmington D, Gray L & Jahrsdoerfer R Binaural processing after corrected congenital unilateral conductive hearing loss. Hearing research 74, 99–114 (1994). [PubMed: 8040103]

152. Keating P & King AJ Developmental plasticity of spatial hearing following asymmetric hearing loss: context-dependent cue integration and its clinical implications. Frontiers in systems neuroscience 7, 123 (2013). [PubMed: 24409125]

153. Agterberg MJ, Hol MK, Van Wanrooij MM, Van Opstal AJ & Snik AF Single-sided deafness and directional hearing: contribution of spectral cues and high-frequency hearing loss in the hearing ear. Frontiers in neuroscience 8, 188 (2014). [PubMed: 25071433]

154. Slattery WH III & Middlebrooks JC Monaural sound localization: acute versus chronic unilateral impairment. Hearing research 75, 38–46 (1994). [PubMed: 8071153]

155. Van Wanrooij MM & Van Opstal AJ Contribution of head shadow and pinna cues to chronic monaural sound localization. Journal of Neuroscience 24, 4163–4171 (2004). [PubMed: 15115811]

156. Keating P, Dahmen JC & King AJ Context-specific reweighting of auditory spatial cues following altered experience during development. Current Biology 23, 1291–1299 (2013). [PubMed: 23810532]

157. Keating P, Rosenior-Patten O, Dahmen JC, Bell O & King AJ Behavioral training promotes multiple adaptive processes following acute hearing loss. Elife 5, e12264 (2016). [PubMed: 27008181]

158. Keating P, Dahmen JC & King AJ Complementary adaptive processes contribute to the developmental plasticity of spatial hearing. Nature neuroscience 18, 185–187 (2015). [PubMed: 25581359] Extracellular recordings mapping ILD sensitivity in primary auditory cortex of ferrets reared with one ear plugged, demonstrating plastic remapping of ILD sensitivity in asymmetric hearing loss.

159. Popescu MV & Polley DB Monaural deprivation disrupts development of binaural selectivity in auditory midbrain and cortex. Neuron 65, 718–731 (2010). [PubMed: 20223206]

160. King AJ, Parsons CH & Moore DR Plasticity in the neural coding of auditory space in the mammalian brain. Proceedings of the National Academy of Sciences 97, 11821–11828 (2000).

161. Yin TC in Integrative functions in the mammalian auditory pathway 99–159 (Springer, 2002).

162. Edmonds BA & Krumbholz K Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? Journal of the Association for Research in Otolaryngology 15, 103–114 (2014). [PubMed: 24218332]

163. Salminen NH et al. Integrated processing of spatial cues in human auditory cortex. Hearing research 327, 143–152 (2015). [PubMed: 26074304]

164. Bendor D & Wang X The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161–1165 (2005). [PubMed: 16121182]

165. De Angelis V et al. Cortical processing of pitch: Model-based encoding and decoding of auditory fMRI responses to real-life sounds. Neuroimage 180, 291 – 300 (2017). [PubMed: 29146377]

166. Bendor D, Osmanski MS & Wang X Dual-pitch processing mechanisms in primate auditory cortex. Journal of Neuroscience 32, 16149–16161 (2012). [PubMed: 23152599]

167. Oxenham AJ How we hear: The perception and neural coding of sound. Annual review of psychology 69, 27 – 50 (2018).

168. Blauert J Spatial hearing: the psychophysics of human sound localization. (MIT press, 1997).

169. Hawley ML, Litovsky RY & Colburn HS Speech intelligibility and localization in a multi-source environment. The Journal of the Acoustical Society of America 105, 3436–3448 (1999). [PubMed: 10380667]

170. Langendijk EH, Kistler DJ & Wightman FL Sound localization in the presence of one or two distracters. The Journal of the Acoustical Society of America 109, 2123–2134 (2001). [PubMed: 11386564]

171. Best V, van Schaik A & Carlile S Separation of concurrent broadband sound sources by human listeners. The Journal of the Acoustical Society of America 115, 324–336 (2004). [PubMed: 14759025]

172. Braasch J & Hartung K Localization in the presence of a distracter and reverberation in the frontal horizontal plane. I. Psychoacoustical data. Acta acustica united with Acustica 88, 942–955 (2002).

173. Shinn-Cunningham BG Influences of spatial cues on grouping and understanding sound. Proceedings of the Forum Acusticum. 29, (2005).

174. Furukawa S, Xu L & Middlebrooks JC Coding of sound-source location by ensembles of cortical neurons. Journal of Neuroscience 20, 1216–1228 (2000). [PubMed: 10648726]

175. Faller C & Merimaa J Source localization in complex listening situations: Selection of binaural cues based on interaural coherence. The Journal of the Acoustical Society of America 116, 3075–3089 (2004). [PubMed: 15603153]

176. Dietz M, Ewert SD & Hohmann V Auditory model based direction estimation of concurrent speakers from binaural signals. Speech Communication 53, 592–605 (2011).

177. Coffey CS et al. Detection of interaural correlation by neurons in the superior olivary complex, inferior colliculus and auditory cortex of the unanesthetized rabbit. Hearing research 221, 1–16 (2006). [PubMed: 16978812]

178. Bregman AS Auditory scene analysis: The perceptual organization of sound. (MIT press, 1994).

179. David M, Lavandier M & Grimault N Sequential streaming, binaural cues and lateralization. The Journal of the Acoustical Society of America 138, 3500–3512 (2015). [PubMed: 26723307]

180. Füllgrabe C & Moore BC Objective and subjective measures of pure-tone stream segregation based on interaural time differences. Hearing research 291, 24–33 (2012). [PubMed: 22771780]

181. Stainsby TH, Füllgrabe C, Flanagan HJ, Waldman SK & Moore BC Sequential streaming due to manipulation of interaural time differences. The Journal of the Acoustical Society of America 130, 904–914 (2011). [PubMed: 21877805]

182. Micheyl C et al. The role of auditory cortex in the formation of auditory streams. Hearing research 229, 116–131 (2007). [PubMed: 17307315]

183. Micheyl C, Tian B, Carlyon RP & Rauschecker JP Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 48, 139–148 (2005). [PubMed: 16202714]

184. Gutschalk A, Oxenham AJ, Micheyl C, Wilson EC & Melcher JR Human cortical activity during streaming without spectral cues suggests a general neural substrate for auditory stream segregation. Journal of Neuroscience 27, 13074–13081 (2007). [PubMed: 18045901]

185. Fishman YI, Reser DH, Arezzo JC & Steinschneider M Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey. Hearing research 151, 167–187 (2001). [PubMed: 11124464]

186. Middlebrooks JC & Bremen P Spatial stream segregation by auditory cortical neurons. Journal of Neuroscience 33, 10986–11001 (2013). [PubMed: 23825404]

187. Middlebrooks JC & Onsan ZA Stream segregation with high spatial acuity. The Journal of the Acoustical Society of America 132, 3896–3911 (2012). [PubMed: 23231120]

188. Griffiths TD & Warren JD The planum temporale as a computational hub. Trends in neurosciences 25, 348–353 (2002). [PubMed: 12079762]

189. Deike S, Scheich H & Brechmann A Active stream segregation specifically involves the left human auditory cortex. Hearing research 265, 30–37 (2010). [PubMed: 20233603]

190. Golding NL & Oertel D Synaptic integration in dendrites: exceptional need for speed. The Journal of physiology 590, 5563–5569 (2012). [PubMed: 22930273]

191. Plaušká A, Borst JG & van der Heijden M Predicting binaural responses from monaural responses in the gerbil medial superior olive. Journal of neurophysiology 115, 2950–2963 (2016). [PubMed: 27009164]

192. Yin T & Chan J Interaural time sensitivity in medial superior olive of cat. Journal of neurophysiology 64, 465–488 (1990). [PubMed: 2213127]

193. Franken TP, Bremen P & Joris PX Coincidence detection in the medial superior olive: mechanistic implications of an analysis of input spiking patterns. Frontiers in neural circuits 8, 42 (2014). [PubMed: 24822037]

194. Caird D & Klinke R Processing of binaural stimuli by cat superior olivary complex neurons. Experimental Brain Research 52, 385–399 (1983). [PubMed: 6653700]

195. Park TJ, Grothe B, Pollak GD, Schuller G & Koch U Neural delays shape selectivity to interaural intensity differences in the lateral superior olive. Journal of Neuroscience 16, 6554–6566 (1996). [PubMed: 8815932]

196. Moore JK The human auditory brain stem: a comparative view. Hearing research 29, 1–32 (1987). [PubMed: 3654394]

197. De Martino F et al. Spatial organization of frequency preference and selectivity in the human inferior colliculus. Nature communications 4, 1386 (2013).

198. Moerel M, De Martino F, Uğurbil K, Yacoub E & Formisano E Processing of frequency and location in human subcortical auditory structures. Scientific reports 5, 17048 (2015). [PubMed: 26597173]

Box 1:

The physiology of sound perception and the encoding of interaural delays

The delay between sound waves arriving at the two ears can be described in terms of *timing*, i.e. interaural time differences (ITDs). The physiological range of ITDs is determined by head size: for instance, the maximum ITD in cats is ~300 µs, and in humans ~600 µs. However, for periodic tones the time differences up to half the period of the tone can also be expressed in terms of *phase* shifts, i.e. interaural phase differences (IPDs); note that larger time differences result in phase ambiguities because it is unclear which ear is lagging and which is leading. Cochlear filtering breaks broadband sounds into their respective frequency components such that the delay between the two ears is not (only) processed for the overall sound wave, but for the individual frequency components. These frequency components (essentially pure tones) are periodic and therefore the delay between the ears can also be expressed in terms of IPDs. Similar to ITDs, the physiological range of IPDs is determined by head size. Additionally, the maximum IPD is dependent on sound frequency: low frequency tones have smaller maximal IPDs than high frequency tones (see figure, part **a**). Finally, the upper limit of the physiological range of IPDs is determined by the temporal resolution of the auditory system (that is, the highest frequency to which the firing of auditory nerve fibers and neurons at subsequent stages in the subcortical auditory pathway can phase lock). Taken together, the physiological range of IPDs is a function of sound frequency, head size, and temporal resolution of the auditory system across mammalian species (see figure, panel **b,** see also main text for further details). Part **b** is adapted with permission from REF.[75], Springer Nature Limited.
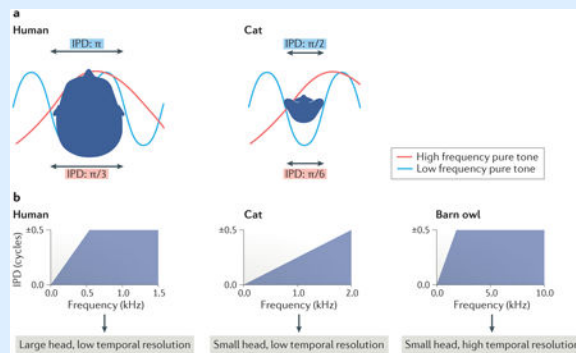
Box 2:

Spatial processing in the subcortical auditory pathway

Binaural spatial cues (ITDs and ILDs) are processed to a large extent, or even completely, before arriving at the auditory cortex (reviewed elsewhere[65]). In brief, sound waves are transduced into action potentials by the hair cells in the cochlea and projected through the auditory nerve fibers to the cochlear nucleus. Subsequently, processing splits into distinct pathways for each type of binaural cue, continuing either into the medial or into the lateral nucleus of the superior olive[161] (MSO and LSO, respectively). In MSO, encoding of ITDs relies on excitatory-excitatory (EE) neurons with Gaussian-shaped tuning curves[65,190-193] while ILDs are encoded by excitatory-inhibitory neurons in LSO, resulting in sigmoidal tuning curves[194,195]. Although research in avian species established the existence of 'auditory space maps' in subcortical structures[66-68], in mammals neither best ILDs nor best ITDs seem to be organized topographically within subcortical structures such as the MSO, LSO, and inferior colliculus[65]. Only multisensory nuclei such as the deep layers of the superior colliculus contain topographic maps of auditory space as well as maps of visual space[69].

Until recently our knowledge of subcortical auditory spatial processing was based solely on single-unit recordings in non-human animals. Owing to the small size of subcortical structures in the human auditory pathway (e.g. the human inferior colliculus has an average width of ~7mm[196]) and the limited spatial resolution of conventional fMRI research at the 3-Tesla level it has heretofore been impossible to measure these processes in humans. However, new developments in ultra-high field MRI (at magnetic field strengths of 7 Tesla or higher) now enable fMRI measurements at sub-millimeter resolution, opening up the functional organization of the human subcortical auditory pathway for non-invasive research. The potential of these technological advances in neuroimaging is demonstrated by studies revealing, for the first time, the tonotopic organization of human IC[197] (consisting of a dorso-lateral to ventro-medial low to high frequency tuning gradient, see figure) and medial geniculate body (MGB)[198], the joint encoding of frequency and sound location in MGB, and the use of tonotopic mapping in MGB to distinguish the different sub nuclei in vivo (i.e. the dorsal [MGd] and ventral division [MGv])[198].
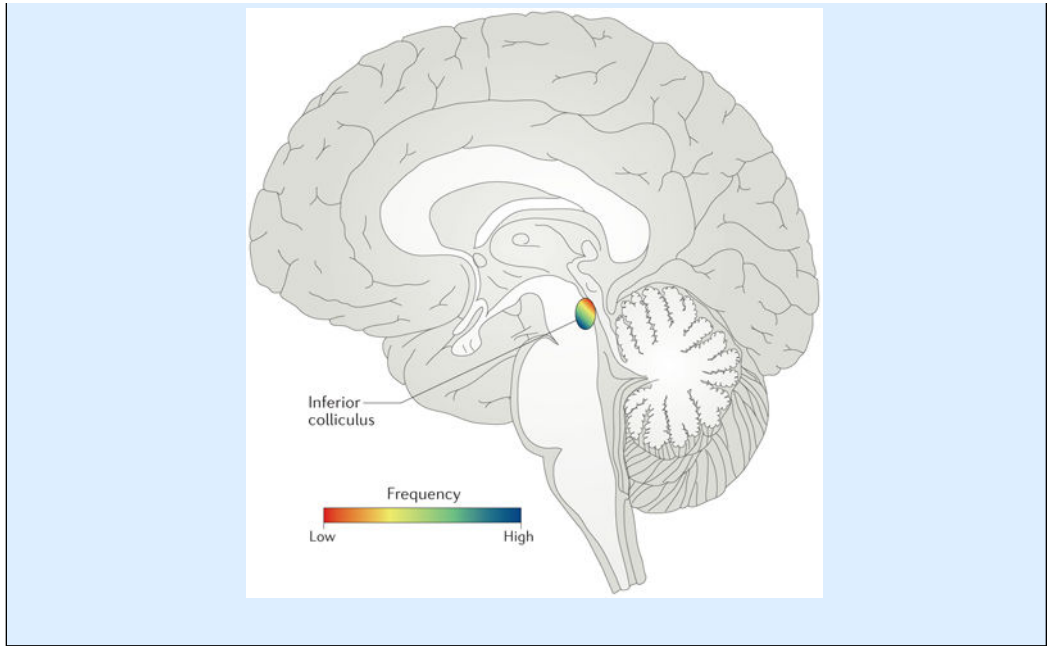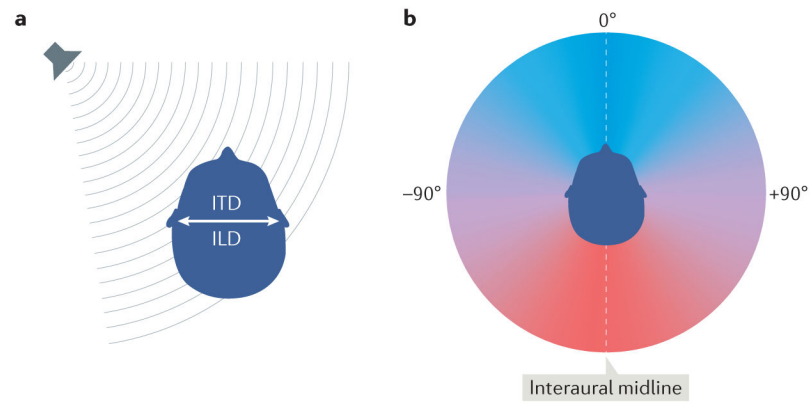
Figure 1. Sound localization in humans.
(A) Schematic representation of binaural disparity cues: Interaural time and level differences (ITD and ILD, respectively). A sound wave emitted by a source at a lateral position is delayed and attenuated in one ear compared to the other ear as a result of the head being in between. (B) Human localization acuity in the horizontal plane. Acuity is best at frontal locations around the interaural midline (at 0°) and deteriorates towards the periphery and back, indicated here by the green to red color gradient
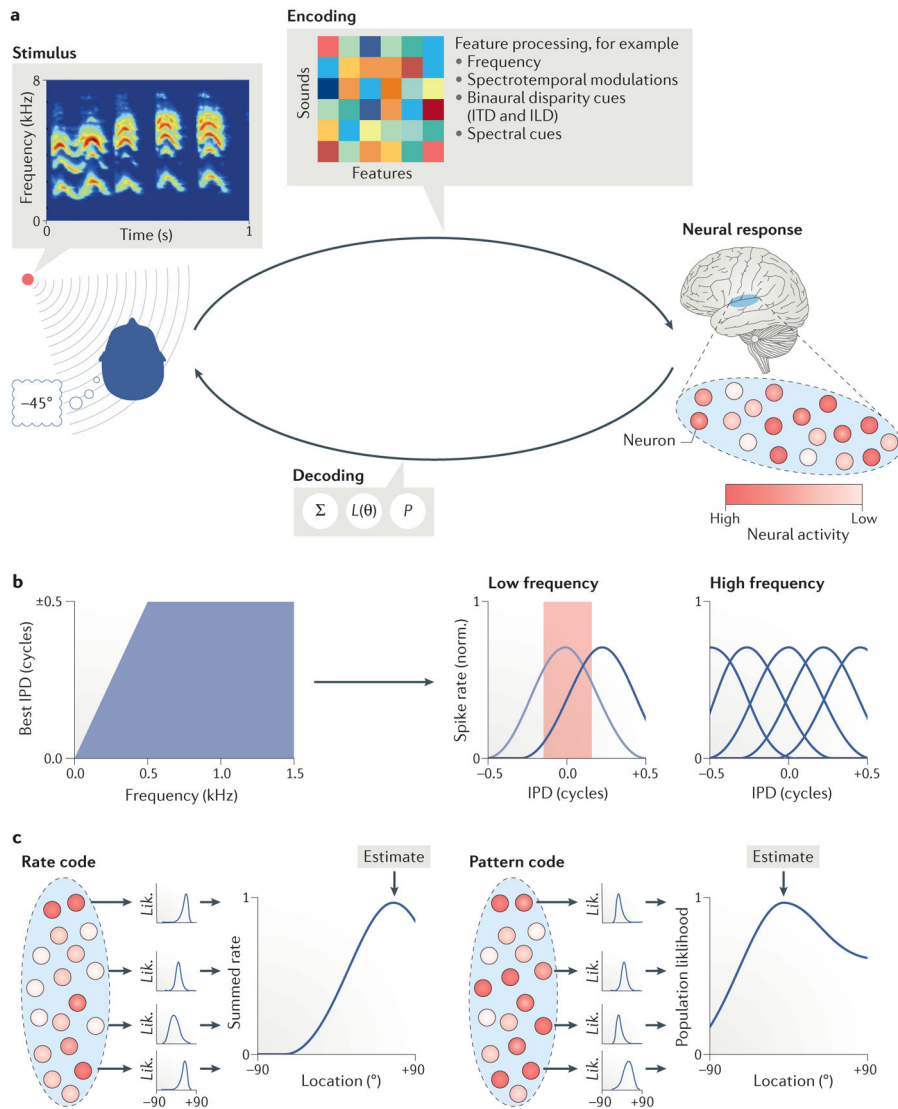
Figure 2. From binaural real-life sound to location percept: Encoding and decoding.
(A) The complex sound wave of a real-life sound at a lateral position arrives at the two ears, which generates different binaural disparity cues as a function of location and ultimately the listener's perception of sound location. Sound location (and the perception thereof) is coded by the response patterns of neuronal populations in the auditory cortex. Circles represent neurons and the intensity of red reflects degree of activity (that is, higher intensity corresponds to higher activity). Encoding refers to the computations required to transform the binaural presentation of a real-life sound into a neural representation of location. Computational models are used to examine the processing mechanisms and understand how relevant features are analyzed and combined. Decoding refers to the read-out of (perceived) sound source location from the neuronal population signal. Possible decoding strategies include a summed rate code ( ), maximum likelihood estimation ($L(\theta)$), and others (P). (B) Illustration of the optimal coding model[75] in humans. The left panel shows the physiological range of IPDs for binaural sound as a function of sound frequency. The right panel shows examples of the optimal distribution – given the physiological range of IPDs displayed on

the left – of neural tuning to IPD to encode a low frequency sound (middle) and a high frequency sound (right). Each curve represents a tuning curve of a neuron to IPD. Specifically, the plot illustrates that given the narrow range of IPDs for low frequencies (indicated by the red rectangle in middle panel), neurons with peak responses to small IPDs (grey curve) would be modulated little by small changes in IPD, while neurons with peak responses outside of the physiological range (black curve) would be modulated maximally by small changes in IPD. In contrast, given that for high frequencies the physiological range of IPDs encompasses the entire phase (blue rectangle in panel on the right), a homogenous distribution of IPDs leads to more accurate encoding of all possible IPDs. (C) Schematic representations of a maximum likelihood population pattern code for sound location. The left and right panels each represent the transformation from neural responses to a specific sound location across a population to a likelihood function (Lik.). Specifically, the activity of each neuron (indicated by the intensity of red) in the population is multiplied (that is, weighted) by the logarithm of the neuron's tuning curve. For each neuron, the resulting likelihood function (Lik.) reflects the probability that the observed neural response was elicited by a sound at a given location (the graphs in the middle display likelihood functions for a few neurons in the population). The likelihood functions of all neurons are subsequently pooled to arrive at a population likelihood function. The peak of the population likelihood function reflects the estimated location (approximately 80° and −10° in the left and right panel, respectively). ILD, interaural level difference; ITD, interaural time difference; norm., normalized. Part **b** adapted from REF.[75], Springer Nature Limited.
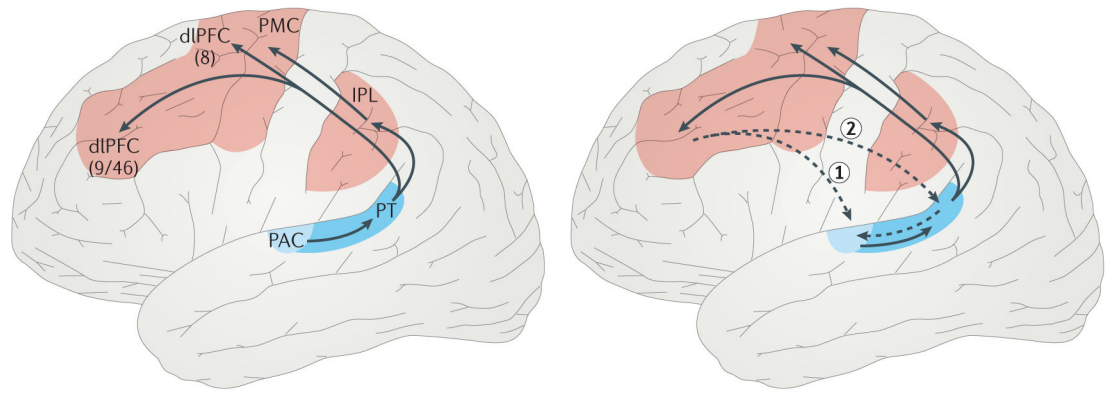
Figure 3. Cortical network of sound location processing.
Schematic representation of the different stages of sound location processing in the cortical auditory pathway during passive sound localization (left), and active, goal-oriented localization (right). Blue areas indicate auditory cortical regions, with darker blue representing increased spatial sensitivity. The red areas are sensorimotor regions of the auditory dorsal stream. Black arrows signal functional cortico-cortical (feedforward) connections. The dashed arrows indicate two potential routes for feedback connections to explain the sharpening of spatial tuning in the primary auditory cortex (PAC) during active sound localization. Arrow 1 reflects a direct feedback route for top-down modulations of activity in PAC by dorsolateral prefrontal cortex (dlPFC). Arrow 2 reflects an indirect feedback route in which top-down modulations of PAC activity by dlPFC are mediated by planum temporale (PT), the area that is traditionally implicated in spatial processing The numbers 8, 9, and 46 refer to Brodmann areas.
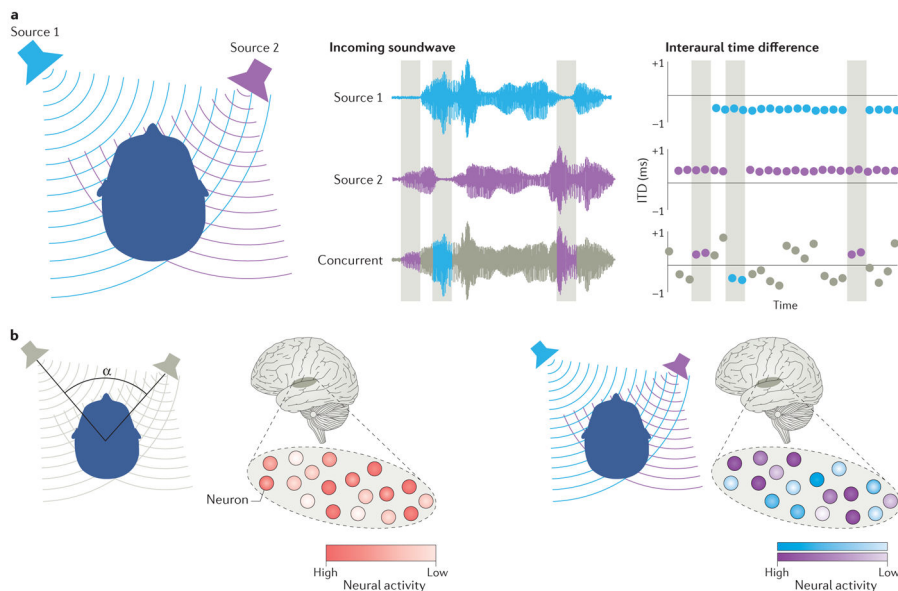
Figure 4. Sound localization in complex auditory scenes.
(A) The distortion of interaural time differences (ITDS) in complex (multi-source) listening scenes. Left panel: the spatial lay-out of a complex auditory scene with two sound sources. Middle panel: the sound wave of the listening scene if source 1 is present in isolation (top row), if source 2 is present in isolation (middle row) or if source 1 and 2 are concurrent (bottom row); owing to the sparse character of natural sounds, concurrent sound streams contain time instants in which only one sound source is present (indicated by the gray rectangles). Right panel: a schematic representation of ITDs to be expected if source 1 is present in isolation (top row), if source 2 is present in isolation (middle row), and if source 1 and 2 are concurrent (bottom row). ITDs for concurrent sounds are distorted and fluctuate over time (bottom row) but accurately reflect the position of a single source in time instants when only this sound source is present (gray rectangles). The auditory system is hypothesized to use the ITDs in these time instants for sound localization. (B) Two potential neural codes for spatial separation. The left panel illustrates a neural representation of relative sound location that is independent of the absolute location of the individual sound sources. That is, the pattern of neural activity in the population is determined by the angle of spatial separation between the individual sound sources (α). Importantly, an identical perceptual spatial separation (that is, identical α) will result in the same pattern of neural activity even when the locations of the sound sources differ from those depicted here. By contrast, the right panel shows a neural representation of spatial separation that emerges from the absolute location of individual sound sources. Specifically, the pattern of activity in the neuronal population is determined by the location of the individual sound sources. Thus, an identical angle of spatial separation (α) will result in a different pattern of neural activity when the two sound sources are at locations different from those depicted here.