The shape of ears to come: dynamic coding of auditory space

Andrew J. King, Jan W.H. Schnupp and Timothy P. Doubell

In order to pinpoint the location of a sound source, we make use of a variety of spatial cues that arise from the direction-dependent manner in which sounds interact with the head, torso and external ears. Accurate sound localization relies on the neural discrimination of tiny differences in the values of these cues and requires that the brain circuits involved be calibrated to the cues experienced by each individual. There is growing evidence that the capacity for recalibrating auditory localization continues well into adult life. Many details of how the brain represents auditory space and of how those representations are shaped by learning and experience remain elusive. However, it is becoming increasingly clear that the task of processing auditory spatial information is distributed over different regions of the brain, some working hierarchically, others independently and in parallel, and each apparently using different strategies for encoding sound source location.

Although vision might be our dominant sense, we rely on hearing as our only panoramic, long-range sensory system. The ability not only to detect and identify a sound, but also to pinpoint swiftly and accurately the location of its source can bring substantial advantages. This applies equally to a predator stalking its prey in the wild and to the modern day pedestrian negotiating a busy crossroads. But although determining the location of a visual or tactile stimulus is relatively trivial (it can be read off directly from the receptor array in the retina or skin), localizing a sound source is a highly complex computational task. The brain has to infer sound source locations from acoustical cues generated by the physical properties of the head and external ears (see Box 1). Some of these cues vary over ranges that tax the resolution of our sensory system to the limits. For example, humans can discriminate between two nearby sound sources by detecting differences as small as 10 µs in the time it takes each sound to reach each ear. Moreover, each set of auditory localization cues is, to some extent, ambiguous, and might have to be processed in an environment full of distracting noises and echoes. To arrive at consistent, and usually highly accurate, estimates of sound source location, the brain has to pool information from different auditory processing channels. Matters are further complicated by the fact that the association between sound locations and their corresponding cue values must be learned and then recalibrated if these spatial relationships change. This applies especially, though not exclusively, during development, as the head grows.

Although auditory distance perception in humans and other primates has been the subject of recent studies1-3, most psychophysical and neurophysiological research on sound localization has focused on the basis for determining the horizontal and vertical direction of a sound source. Much is now known about the sensitivity of neurones in the central auditory system to individual localization cues. Nevertheless, many details of the neural circuitry involved in processing auditory spatial information remain obscure. Indeed, the

---

References:

58 Miyamoto, H. et al. (1996) A Kendama learning robot based on bidirectional theory. Neural Netw. 9, 1281–1302

---

Andrew J. King
Jan W.H. Schnupp
Timothy P. Doubell
University Laboratory of Physiology, Parks Road, Oxford, UK. OX1 3PT.
*e-mail: andrew.king@physiol.ox.ac.uk

http://tics.trends.com 1364-6613/01/$ – see front matter © 2001 Elsevier Science Ltd. All rights reserved. PII: S1364-6613(00)01660-0
The number of brain areas implicated in this task continues to grow, as further candidate areas are identified in studies using functional imaging, neuroanatomical methods or electrophysiological recordings. With the availability of powerful new approaches, including virtual auditory space (VAS) techniques (see Box 2) and multi-electrode arrays, as well as a growing recognition of the need to combine behavioural, neurophysiological and computational studies, it is likely that considerable progress will be made in this field in the next few years. In this article, we highlight recent advances in our understanding of the nature of the neural code for auditory space at different levels of the brain and then consider the extent to which these representations are shaped by experience.

Extraction of auditory localization cues in the brainstem

The neural processing of different auditory localization cues begins in largely separate pathways in the brainstem. Auditory nerve fibres leaving the cochlea terminate on neurons in different subdivisions of the cochlear nucleus. Cells in the ventral subdivisions project, either directly or indirectly, to the superior olivary complex (SOC) on both sides of the brainstem. The SOC is therefore the principal site of binaural convergence, and it is at this level that neuronal sensitivity to interaural time differences (ITDs) and interaural level differences (ILDs) arises. Although the neural basis for detecting binaural localization cues has been studied extensively, comparatively little is known about the steps involved in the extraction of spectral cues for sound location. However, recording and behavioural studies have implicated the dorsal cochlear nucleus in this process.

The primary target of these brainstem structures is the central nucleus of the inferior colliculus (ICC). The spatial selectivity of ICC neurons reflects both the interactions that take place in the brainstem and further processing in this midbrain nucleus. The ICC provides an essential relay in the transmission, via the thalamus, of information to the primary
localize stimuli with a narrow bandwidth\(^b\). Additional cues are therefore required in order to determine unambiguously the location of the sound source.

ILDs are more complex than ITDs in that they vary much more with sound frequency. Low-frequency (long wavelength) sounds easily travel around the head and produce negligible ILDs (Fig. IIa). ILD values produced at higher frequencies (Fig. IIb) are larger, and are increasingly influenced by the filter properties of each external ear, which imposes peaks and notches on the sound spectrum reaching the eardrum (Fig. III). Instead of being centred on the inter-aural axis, cones of confusion associated with particular ILD values have different shapes for each sound frequency\(^c–e\). Consequently, pooling information across frequency channels makes it possible to pin down the sound direction to the intersection points of the ILD contours. Sounds with broad frequency content can thus be localized unambiguously.

In addition to helping to resolve front-back confusions, the spectral cues arising from the direction-dependent filtering by the external ear are responsible for vertical localization\(^a\), and even allow sounds to be localized using one ear alone\(^f,g\). It is, however, important to remember that these spectral pattern cues are superimposed on the spectrum of the sound to be localized. This source spectrum might be unknown, or might not contain much energy in the high frequency regions where pinna cues are most prominent. Monaural localization cues are therefore unambiguous only when the sounds to be localized are familiar and broadband\(^h\).

**References**


**Fig. III.** Monaural spectral cues for sound location. Direction-dependent filtering of sounds by the head and external ear of a human subject. Broadband sounds were presented in front of the subject at different elevations. Changes in amplitude (gains) are shown as a function of sound frequency and elevation.

**Mapping auditory space in the midbrain**

Free-field studies have shown that a map of auditory space exists within the deeper layers of the SC (Ref. 11). This is defined by an orderly relationship between the preferred sound directions of auditory neurones and their location within the nucleus (Fig. I). In contrast to most other relay stations between the cochlea and the cortex, the SC is not ‘tonotopically’ organized. Instead, spatial information is combined across different frequency-specific channels in the brainstem to produce neurones that are tuned to specific locations in space.

The precision of the auditory space map varies between different species, reaching its zenith in the barn owl’s optic tectum, the avian homologue of the SC (Ref. 11). In each case, however, the range of preferred sound directions is superimposed on the representation of the visual field, which co-exists in this multisensory nucleus. Visual and auditory inputs, together with those from the body surface, are translated by the SC into motor commands for controlling orienting movements that re-direct attention toward novel events, irrespective of their modality. Aligning the sensory maps is also crucially important for making sense of the multisensory interactions exhibited by many SC neurones, which depend on the spatial and temporal relationship between different sensory stimuli\(^3\).

The reference frames used for representing space vary between different sensory systems. For example, the coordinates of visual space are centred on the retina, whereas those of auditory space are centred on the head and ears. Crossmodal integration of spatial information must therefore take into account the
Box 2. Virtual auditory space

Psychophysical and neurophysiological studies of sound localization have traditionally employed two approaches for delivering stimuli to the subject. In free-field studies, sounds are presented from loudspeakers positioned around the subject’s head. These stimuli interact naturally with the head, torso and external ears, giving rise to various spatial cues depending on the direction and frequency composition of the sound. Dichotic presentation of sounds over headphones allows the experimenter to measure the sensitivity of human listeners or single neurones recorded in animals to interaural time differences (ITDs) and interaural level differences (ILDs). However, when sounds are delivered over headphones, the filtering characteristics of the external ears are typically eliminated. As a result, the sound is ‘lateralized’ (i.e. judged to come from the left or right) on the basis of the binaural cues, but perceived as if it comes from inside the head or a position very close to the ear. Signals presented in this way will be ‘externalized’, i.e. sound as if they originate from a real sound source out in space, only if the spectral characteristics of the external ears are faithfully replicated in the headphone signals\(^a,b\).

Virtual acoustic space (VAS) stimulation employs digital signal processing techniques to synthesize headphone signals that replicate all the localization cues associated with a real sound source. This gives rise to the percept of an externalized source that can be localized just as accurately as a comparable free-field source\(^c\). Once the acoustic cues have been measured at a large number of spatial locations, VAS stimulation can be used to generate stationary or moving sound images, without the physical constraints imposed by using loudspeakers. By manipulating the signals presented to each ear, this approach can be used to assess the relative contributions of ITDs, ILDs and spectral cues to human auditory localization\(^d,e\), and is being used increasingly to study spatial coding in animals\(^f\).

References


Alternative coding strategies in the cortex

Recent studies of the neural processing of auditory space have focused increasingly on the cortex. It has been known for a long time that temporal lobe damage can lead to impaired sound localization\(^17,18\). But lesion studies are not always easy to interpret,
Fig. 1. Representation of auditory space in the deeper layers of the superior colliculus.

(a) Experiments using free-field stimulation under anechoic conditions have been used to determine the spatial receptive fields of auditory neurones. In this case, broadband noise bursts were presented from a single loudspeaker that was positioned at different directions with respect to the head of an anaesthetized ferret. (b) Spatial response profile showing how the spike firing rate of one superior colliculus (SC) neurone varies with the azimuthal direction of the sound source. Although the neurones have large receptive fields, they typically respond best to sounds presented from a single region of space. (c) Plotting the value of the auditory best position for each neurone against the location of its recording site in the SC reveals that sound azimuth is mapped along the rostrocaudal axis of this midbrain nucleus. Similarly, best elevations are represented topographically across the mediolateral axis. The region of space covered by each dimension of this neural representation conforms to visual and somatosensory maps that co-exist in the SC. Adapted from Refs 10,43.
Auditory localization is individually calibrated

The size and shape of the head and external ears, and therefore the values of the auditory localization cues corresponding to particular sound directions, can vary substantially between individuals35–37. These cue values also undergo changes within a subject as the head and ears grow (Fig. 3). Accurate localization therefore requires that the neural code for auditory space be calibrated to these individual characteristics. Evidence for this has been provided in human psychophysical studies in which VAS techniques have been used to allow listeners to hear through someone else’s ears. The localization of virtual sound sources is more accurate when these stimuli are based on head-related transfer function (HRTF) measurements of the subject’s own ears than on those derived from the ears of other individuals36,38,39. Much of this difference in performance can be eliminated by frequency scaling the transfer functions of the ‘foreign’ ears, so that the main spectral features match those of the subject’s own ears more closely38. A possible neural correlate for these observations is provided by the finding that switching from individualized to non-individualized HRTFs can significantly change the spatial tuning of A1 neurones40,41 (Fig. 4).

Although these studies further highlight the importance of spectral cues in spatial hearing, we must remember that many features of the external ear’s directional filtering occur at the high frequency end of the hearing range (Fig. 3, Box 1). However, the frequency resolution of the auditory system decreases as sound frequency increases, so we would expect the detailed spectral pattern to be poorly resolved in the central auditory system. This appears to have been confirmed by a study showing that spectral cues can be subjected to considerable smoothing without appreciable loss in localization performance42. The process of calibrating the auditory system to individual pinna cues is
original connections remain in place but are suppressed (broken lines). The tectum. This involves growth of novel projections from the ICC to the ICX (unbroken lines); the prism-reared owls, thereby re-establishing the alignment with the optically displaced visual map in the tectum. The auditory space maps in the ICX and tectum gradually shift by an equivalent amount in front of their eyes. The same visual stimulus now activates a different set of neurones in the optic nucleus of the inferior colliculus (ICC). The ICX map of auditory space is then conveyed to the optic tectum, where it is superimposed on a map of visual space. (b) From around 40 days after hatching, auditory cortex of an adult ferret. The diameter of the neurone in the primary fields obtained from a different animal. Data from Ref. 40.

therefore, presumably, based more on the principal peaks and valleys in the HRTFs than on detailed spectral structure.

Experience shapes the developing representation of auditory space
Most of the neurophysiological evidence for experience-driven adjustments in sound localization has come from developmental studies of the auditory space map in the SC. Experiments in which animals have been reared with modified auditory or visual inputs have revealed substantial plasticity in this representation.43,44

The potential for the sound localization pathway to be shaped by experience of the acoustic localization cues provided by the listener’s own ears is illustrated by the compensatory changes that can be induced in the auditory space map by raising ferrets45,46 or barn owls47 with experimentally altered binaural cues. These adjustments in auditory spatial tuning tend to preserve the alignment with the visual representation. In owls, they are brought about by frequency-specific shifts in neuronal tuning to ITDs and ILDs (Ref. 47), although, as discussed below, the basis for this adaptive plasticity in mammals might be different.

Experiments in which animals with normal hearing are raised with altered visual cues suggest that vision plays a guiding role in aligning the different representations of space in the SC (Ref. 43,44). For example, because barn owls are incapable of large eye movements, it is possible to shift their visual world by mounting prisms in front of the eyes. The resulting misalignment between the visual and auditory maps is overcome in prism-reared owls, by retuning neurones in the optic tectum and ICX to ITDs that correspond to the optically displaced visual receptive fields of the tectal neurones. The shift in the auditory space map in the ICX is brought about by the growth of new connections from the ICC that convey sensitivity to the new ITD values, and by the suppression, through GABA-mediated inhibition, of existing connections (Fig. 5). The recent discovery of a projection from the tectum to the ICX (Ref. 51),52 suggests a possible basis by which either visual guidance cues or visual–auditory misalignment signals are conveyed to the site of plasticity. A similar situation is likely to exist in mammals, as neonatal lesions of the superficial visual layers of the ferret SC, which project both to the deeper layers and to the principal source of auditory input in the nBIC (Ref. 53), disrupt the development of the auditory space map in the SC (Ref. 54).

Plasticity of sound localization is not restricted to development
There is considerable evidence that neural circuits, and the behaviours to which they contribute, are particularly dependent on sensory experience during a ‘critical period’ of early life. The plasticity of the auditory space map in the SC is no exception to this, as the spatial tuning of these neurones is more susceptible to altered sensory inputs during infancy than later in life.43,44. This makes good sense, as developmental plasticity is clearly necessary in order to adjust neuronal sensitivity to auditory spatial cues during the growth period when the association...
between those cues and the locations in space to which they correspond is changing.

But recent behavioural and neurophysiological findings indicate that auditory localization can, at least under certain conditions, show substantial plasticity well into adulthood. In owls, the capacity of the auditory system to adjust to, and recover from, abnormal visual experience depends not only on the age of the animal, but also on the richness of the environment in which it is raised. In particular, the tectal map of auditory space and the accuracy of head orienting responses in prism-reared owls will shift back to normal after prism removal at any age if the owls are kept in large aviaries, whereas this capacity for recovery is restricted to the period leading up to sexual maturity in birds that are housed in individual cages. Moreover, although the initial adjustment to abnormal visual cues is developmentally regulated, the ITD tuning of tectal neurones will gradually re-adapt, following an intervening period of normal experience, to a second exposure to the same prisms in adulthood. No adjustments are made at this stage, however, to prisms that shift the visual field by different magnitudes or in other directions. In both cases, these late adaptive responses can be understood in terms of the anatomical changes induced by prism rearing in infancy (Fig. 5): circuits for both normal and abnormal auditory maps are likely to be present, and either set might be expressed in later life according to the prevailing visual inputs.

Psychophysical studies of sound localization in mammals also reveal a capacity for plasticity in adulthood. By using modified VAS stimuli, Shinn-Cunningham and colleagues have shown that human listeners can be rapidly trained to reinterpret the relationship between auditory localization cues and directions in space. More extensive adaptation has been observed by altering spectral cues over a longer time period. To achieve this, Hofman and colleagues fitted moulds to each external ear, which, as expected, disrupted the ability of their subjects to make elevation judgments (Fig. 6). These errors gradually disappeared over a period of several weeks, indicating that the subjects had learned to associate new spectral cues with each direction in space. Substantial recovery in the accuracy of sound localization in the horizontal plane has also been observed in adult ferrets after chronically plugging one ear.

In contrast to the effects of prisms on auditory space processing in owls, the adult plasticity demonstrated in these studies does not appear to depend on previous experience of the abnormal cues during development. Intriguing insights into the underlying mechanisms that might be involved are provided by the consequences of returning to the subjects’ unmodified ears. For example, only small after-effects are observed following earplug removal in ferrets that have learned to localize accurately, either in infancy or adulthood, with one ear occluded.
Questions for future research

- How are the various schemes that have been proposed for coding space in the auditory cortex affected by the presence of multiple sound sources, or under conditions in which other features of the sound source vary? Are these coding schemes emergent properties of the cortex or do they originate in the subcortical nuclei that feed into A1?
- Studies that emphasize the importance of temporal discharge patterns in auditory spatial coding have mostly been carried out on anaesthetized animals. What is the effect of anaesthesia and behavioural state on the spatiotemporal properties of cortical neurones?
- Are different coding strategies employed in different cortical fields? Is there a specialized space-processing pathway that originates with a subset of neurones in A1?
- What are the relative contributions of midbrain and cortex to experience-driven plasticity of sound localization behaviour? If the cortex represents auditory space by the spatiotemporal distribution of activity, how is this affected by sensory experience?
- It is well established that the circuits leading to the synthesis of a map of auditory space in the midbrain are refined by sensory experience during development. What is the basis for the visual calibration of the developing map of auditory space in the midbrain? Does this involve template matching between the visual and auditory maps based on a Hebbian process of synaptic plasticity or some other mechanism? Do the changes induced by experimentally altered inputs involve different mechanisms from those responsible for the initial matching of visual and auditory maps during normal development?
- What is the neurophysiological basis for the improvement in auditory localization reported in blind subjects? An expansion of the cortical areas involved in auditory processing and a sharpening of spatial tuning have been reported. Does this also involve a change in the temporal response properties of cortical neurones?

Concluding remarks

The neural computations involved in determining the location of a sound source are established in the brainstem, but the manner in which this information is encoded changes at higher levels of the auditory system. The SC contains a topographic representation of auditory space, which contributes to the multisensory guidance of orienting behaviour. The steps leading to the formation of this neural map and its alignment with other sensory representations have provided a valuable and accessible substrate for investigating plasticity in the neural coding of auditory space. However, various regions of the cortex have been implicated in localization behaviours that entail more than just reflexive orienting responses. A different form of coding appears to operate at this level, based on the spatiotemporal distribution of activity across populations of neurones. Although some studies suggest that refinements in the auditory spatial tuning of cortical neurones might contribute to the plasticity of localization behaviour, others place greater emphasis on changes in the temporal correlation of discharges among groups of neurones. The lack of topography makes investigating the cortical basis for encoding sound source location and its adaptive plasticity more challenging than in the midbrain, but this is necessary if we are to understand the full capacity of the auditory system to be shaped by experience.

Acknowledgements

Our research is supported by the Wellcome Trust (Senior Research Fellowship to A. J. K.) and by Defeating Deafness (Dunhill Medical Trust Fellowship to J. W. H. S.). We are grateful to Tom Mrsic-Flogel and to Doris Kistler for their assistance with recording some of the data illustrated in this paper, and to David Moore, Tom Mrsic-Flogel and Carl Parsons for their comments on an earlier draft of the manuscript.