Relearning sound localization with new ears

Paul M. Hofman, Jos G.A. Van Riswick and A. John Van Opstal

University of Nijmegen, Department of Medical Physics and Biophysics, Geert Grooteplein 21, NL-6525 EZ Nijmegen, The Netherlands

Correspondence should be addressed to J.V.O (johnvo@mbfys.kun.nl)

Because the inner ear is not organized spatially, sound localization relies on the neural processing of implicit acoustic cues. To determine a sound's position, the brain must learn and calibrate these cues, using accurate spatial feedback from other sensorimotor systems. Experimental evidence for such a system has been demonstrated in barn owls, but not in humans. Here, we demonstrate the existence of ongoing spatial calibration in the adult human auditory system. The spectral elevation cues of human subjects were disrupted by modifying their outer ears (pinnae) with molds. Although localization of sound elevation was dramatically degraded immediately after the modification, accurate performance was steadily reacquired. Interestingly, learning the new spectral cues did not interfere with the neural representation of the original cues, as subjects could localize sounds with both normal and modified pinnae.

Human auditory localization is remarkably accurate, even in complete darkness under open-loop conditions[1–4] (Fig. 2; see also Methods). This performance may rival that of the barn owl, a well studied nocturnal hunter that relies heavily on its auditory system to capture prey. Yet, quite different localization mechanisms are used by each species. In barn owls, sound azimuth (horizontal direction) is derived from interaural timing differences, whereas sound elevation cues are provided by the interaural level differences that result from the up–down asymmetry of its facial ruff[5]. In humans, however, both interaural cues only relate to the sound-source azimuth. In contrast, sound elevation and front-back direction are determined on the basis of spectral cues generated by the direction-dependent filtering of the pinnae[6–9]. Sounds entering the ear via the pinna aperture have some frequencies amplified and some attenuated, with an effect that can be described mathematically by linear transfer functions ('pinna filters', Fig. 1)[10]. Sound elevation detection in humans and in many other mammalian species may therefore be considered as a spectral pattern-recognition problem.

The importance of the pinnae in sound localization has been demonstrated in both pinna-occlusion experiments[11], and in narrow-band sound localization studies[1]. Moreover, when localization is attempted 'through another person's ears' (using virtual sound-source synthesis techniques), localization errors increase dramatically[12]. However, accurate localization on the basis of spectral cues poses constraints on the sound spectrum. A sound needs to be broad-band in order to yield sufficient spectral shape information, and the acoustic signal at the eardrum comprises the original source spectrum as modified by the linear pinna filter. Both spectral functions however, are *a priori* unknown to the auditory localization system, and the extraction of sound elevation and front–back direction is therefore not trivial[1,2,13]. It has been suggested that the auditory system may resolve this problem by assuming that real-life sounds do not contain the prominent peaks and notches of the different pinna filters. (Fig. 1a)[1,2,13].

Because sound localization relies on implicit physical cues, the auditory system must somehow transform the binaural differences and monaural spectral pinna cues into consistent spatial information. It is thought that the auditory system acquires these spatial relations through learning, and that the visual system may train and calibrate the acoustic localization process by providing accurate spatial feedback. Indeed, behavioral experiments with young barn owls have shown that the integrity of the visual system guides acoustic localization performance[14,15]: When reared with prisms, the owl's auditory localization response shifts in the same direction as the altered visual representation, although the acoustic cues remain unchanged. Likewise, acoustic perceptual shifts that are induced by one-sided ear plugs are resolved by the availability of visual feedback, although some adjustment of the optic tectum's spatial map also occurs after eyelid suture[16].

In addition, the formation of the auditory space map in the owl's inferior colliculus has been shown to rely mostly on visual experience in early life[17]. Comparable results have been obtained for the spatial representation of sound in the midbrain superior colliculus of newborn mammals[18,19]. Conceivably, the human auditory localization system may develop by a similar learning process, as the subtle acoustic cues vary substantially during growth. However, no data are available that clearly demonstrate an adaptive capability of the human auditory localization system[20]. We therefore tested whether human subjects would be able to adapt to a consistent change in the spectral localization cues.

Results

To investigate the processes underlying the formation of the acoustic spatial percept in humans, four adult subjects continuously wore well fitting, custom-made molds within the concha of both ears for a period of up to six weeks. Subjects did not receive any specific localization training during this period. Although the molds dramatically altered the subject's spectral shape cues, they still provided consistent spectral information about stimulus elevation (Fig. 1a). The undisturbed ear of this example subject contained prominent spectral peaks and notches at different frequency bands for each sound elevation. For example, at elevation -40 degrees, there was a profound notch at about 6 kHz and a peak near 12 kHz. The position and presence of such peaks and notches varies in a systematic, albeit complex, manner with elevation[6–9]. There were marked changes after inserting the mold (Fig. 1b).

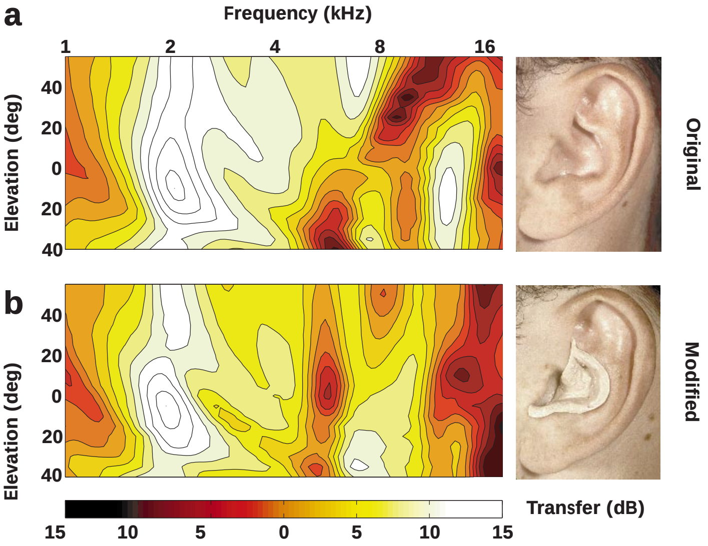

Fig. 1. Effect of molds on pinna acoustic transfer. The human pinna filters the acoustic spectrum in a direction-dependent way. **(a)** Normal pinna transfer functions of the right ear of one subject. **(b)** Pinna transfer functions of the same ear after application of the mold (see also photograph). The linear acoustic transfer functions are shown as a function of frequency (ordinate) and sound direction (abscissa) in the midsagittal plane (azimuth zero degrees; elevation from -40 to +50 degrees). Color encodes the amplitude (in dB) of the transfer function. A value of zero dB indicates that the presence of the head and pinna does not change the sound pressure amplitude of a tone at that particular frequency and elevation. Light colors correspond to sound amplification; dark areas refer to sound attenuation.

During the adaptation period, each subject's localization performance in two dimensions (2D) was quantified several times a week under open-loop conditions. Because of the method used to measure localization performance, the target domain was confined to positions within the oculomotor range (± 40 degrees straight ahead in all directions)[13]. We determined baseline behavior in response to white-noise sounds presented at random locations for all four subjects (Fig. 2a). Both azimuth and elevation components of the measured eye-movement vectors correlate well with the actual stimulus directions, and the first-saccade end points accurately capture the spatial structure defined by the stimulus locations. Perfect behavior would require that the averaged saccade responses align with the stimulus matrix. The data were also quantified by determining linear regression lines between target and response coordinates, for elevation and azimuth components, respectively. The fitted slopes, together with their standard deviations, are provided in Fig. 3. Here, the range for the mean absolute errors across target positions (Δ_ε for elevation, and Δ_α for azimuth components) as well as Pearson's linear correlation coefficients (r_ε and r_α, respectively) are provided for each row. Both the azimuth and elevation components of saccadic responses are quite accurate for all four subjects before application of the molds (Δ_ε, 4.2–7.7 degrees; Δ_α, 2.9–6.6 degrees; r_ε, 0.92–0.96; r_α, 0.97–0.98) (Fig. 2a).

Immediately after application of the molds, however, localization performance was dramatically disrupted, as far as the detection of sound elevation was concerned. All subjects perceived the sound at a roughly fixed, eccentric elevation angle (given by the offset value of the regression line, which fell below -20 degrees for subjects PH, JO and JR, and was about +20 degrees for MZ), regardless of the actual stimulus elevation (Fig. 2b). (Saccade responses scatter around an approximately horizontal line.) The azimuth component of the responses, however, seemed to be almost unaffected and still accurate (Δ_ε, 16.3–23.0 degrees; Δ_α, 3.8–8.3 degrees; r_ε, 0.0–-0.3; r_α, 0.97–0.98). This result underlines the importance of the spectral cues in sound elevation localization in humans; it also demonstrates the existence of independent neural mechanisms for the detection of sound source azimuth and elevation.

After wearing the molds for several days, localization accuracy steadily improved over time in all subjects (Fig. 2c). There was some stimulus-related structure in the elevation responses, which is apparent from significant values of the gains and decreasing offset values for the regression lines (Δ_ε, 11.1–-16.5 degrees; Δ_α, 3.3–8.3 degrees; r_ε, 0.63–0.82; r_α, 0.97–0.98). This improvement, depicted by the systematic 'unfolding' of the response matrices, continued for about three to six weeks, after which the learning process seemed to stabilize (Δ_ε, 7.4–12.6 degrees; Δ_α, 4.2–7.0 degrees; r_ε, 0.78–0.89; r_α, 0.98) (Fig. 2d).

After localization performance had reached a stable level, the molds were removed and the subject's behavior was tested without the molds. Interestingly, immediately after adaptation, the subject's localization accuracy with undisturbed ears was still as high as before the start of the experiment (Δ_ε, 4.5–9.6 degrees; Δ_α, 4.5–5.6 degrees; r_ε, 0.95–0.96; r_α, 0.98) (Fig. 2e). The largest absolute elevation errors in the control conditions are for subject PH, who displayed a systematic downward response bias. Apparently, the auditory system had acquired a new representation of the pinna transfer functions, without interfering with the old set.

We made adaptation curves (azimuth and elevation gain as a function of time) for all four subjects (Fig. 3). Qualitatively, the results are very similar, although idiosyncratic differences in the learning behavior, as well as in the overall performance accuracy may also be noted. As long as several days after the adaptation session was concluded, subjects could still localize adequately with the molds, although the overall elevation gain slowly decreased with time (tested in three subjects, data not shown). However, we cannot exclude any contribution of subtle differences in the spectral localization cues provided by the individual molds.

Discussion

These experiments show that the adult human auditory system is capable of considerable adaptation in response to altered spectral cues. This finding corroborates earlier results from barn owls in which the ability to adjust the auditory space map in the optic tectum in response to altered acoustic cues (induced by cutting the facial ruff) persists into adulthood[21]. It differs qualitatively from findings showing that changes in the auditory space map in response to prisms are strongly limited by age in this species[22].

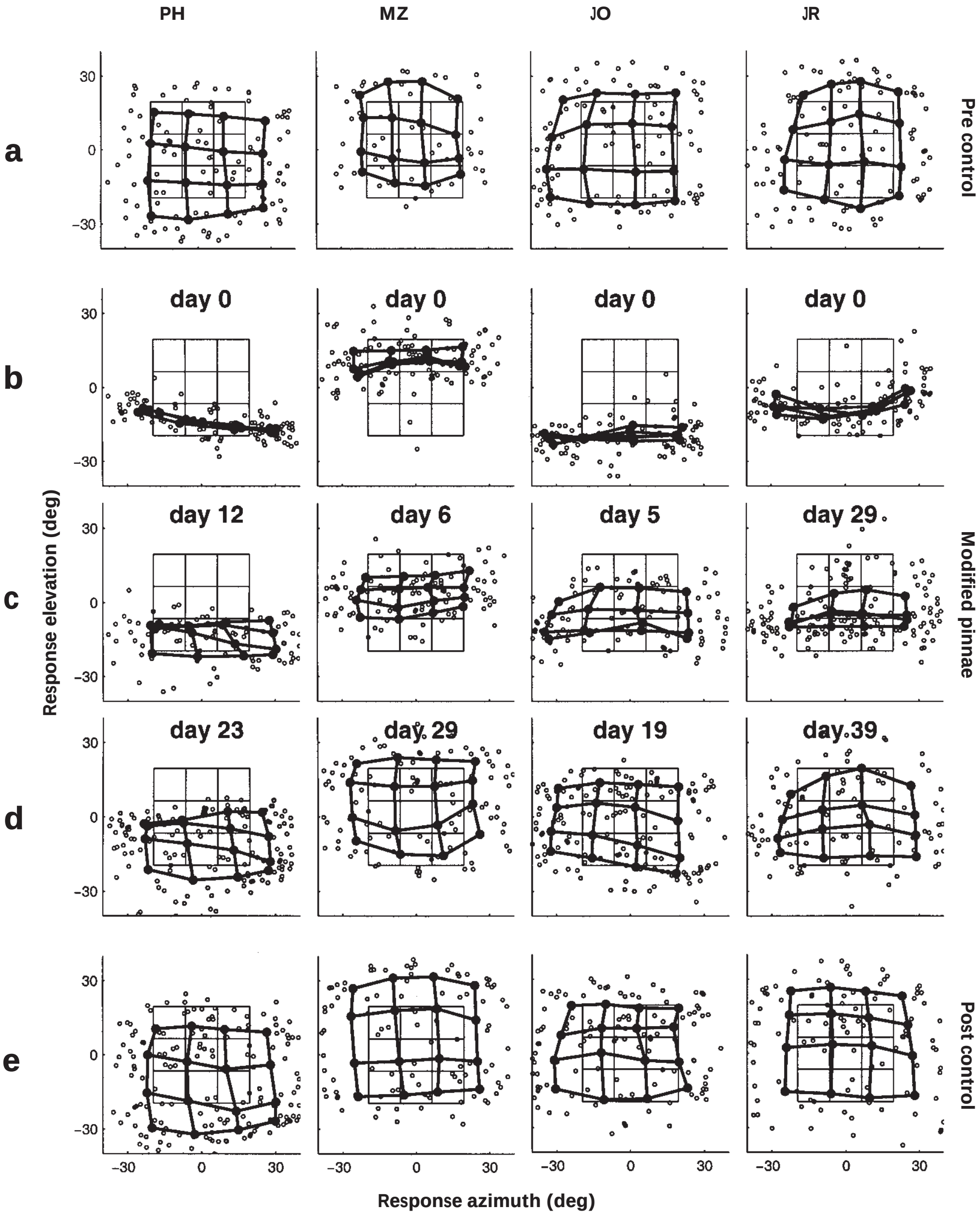

Fig. 2. Adaptation to altered spectral cues. Localization behavior of all four subjects (from left to right) before, during and immediately after the adaptation period. Day zero marks the start of the adaptation experiment. The panels show, for each subject, the individual saccade vector endpoints in the azimuth–elevation plane (symbol ∘). In addition, the saccade vectors were also averaged for targets belonging to similar directions by dividing the target space into sixteen half-overlapping sectors. Averaged data points (symbol •) from neighboring stimulus sectors are connected by thick lines. In this way, a regular response matrix indicates that the subject's saccade endpoints capture the actual spatial distribution of the applied target positions. The target matrix, computed in the same way as the saccade matrix, has been included for comparison (thin lines). **(a)** Results of the pre-adaptation control experiment on day zero, immediately preceding the application of the molds. **(b)** Localization responses immediately after inserting the molds (day 0). Note the dramatic deficit in elevation responses for all subjects. **(c)** Results during the adaptation period after twelve (PH), six (MZ), five (JO) and 29 (JR) days of continuously wearing the ear molds. **(d)** Results near the end of the adaptation period. Stable and reasonably accurate localization behavior has been established in all subjects. **(e)** Results of the control condition, immediately after removal of the molds. All subjects localized sounds with their original ears equally well as before the start of the experiment several weeks earlier.

The existence of a critical period for the formation of a spatial auditory map in the mammalian superior colliculus has also been demonstrated for newborn guinea pigs and ferrets[23]. Our finding that the molds remained effective after their removal indicates that the newly acquired pinna representations may have a semi-permanent basis. Interestingly, the presence of experience-induced auditory 'memory traces' has recently been demonstrated in the adult barn owl localization system[24].

Another finding is the ability of the human sound localization system to cope with different representations of the spectral pinna cues. We believe that this property does not reflect the involvement of higher cognitive processes, as subjects never received feedback about their performance during the recording sessions. Moreover, response latencies were typically well below 300 milliseconds, and subjects were not aware of a difference in perceived sound quality for the mold and no-mold conditions.

Apparently, both filter sets (see Fig. 1) are simultaneously represented within the human auditory system. The learning of a new set of pinna transfer functions therefore resembles more the acquisition of a new language than other forms of sensory adaptation. As a consequence, both neural representations will always be activated by sounds, and the question then is how the auditory system could select the correct pinna filter. It is possible to show, however, that when the two pinna sets are sufficiently disjunct from each other (that is, the correlations between different filter functions within and among the two sets are close

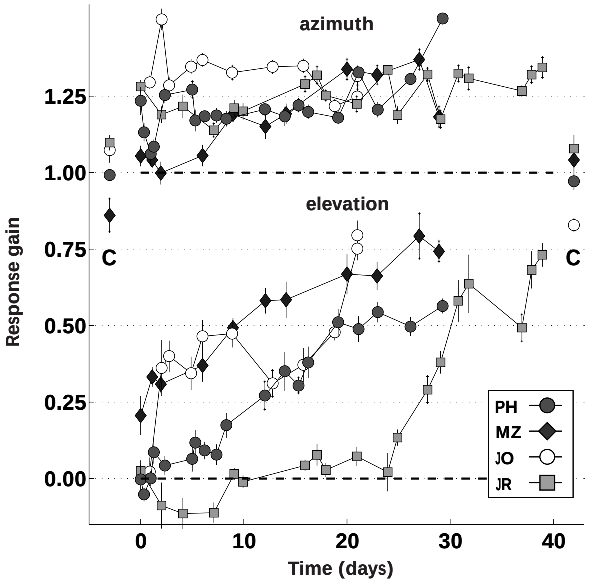

Fig. 3. Summary of the results for all four subjects. Adaptation curves for azimuth and elevation gain (defined as the slope of the best-fit regression line through the data of each recording day), as a function of time from the start of the adaptation experiment (in days). Standard deviations in the gains were obtained by bootstrapping the data 100 times[30,31]. Results for the pre- and post-adaptation control conditions are also shown for comparison (symbols C, elevation data only, for clarity; control azimuth gains fell in the range 1.0–1.2 pre- and 1.2–1.3 post-adaptation).

to zero[1,2,13]), a broad-band white-noise sound will typically yield a high response for only one particular filter[13]. Not only does this filter belong to the actually applied pinna set, it also corresponds to the correct target elevation[13]. If the sound localization percept is ultimately produced by a 'winner-take-all' mechanism, this maximally activated filter determines the actual localization response. In this way, the subject may rely entirely on the available acoustic input, provided the sound-source spectrum itself is both broad-band and unrelated to any of the stored pinna filters of either set[1,2,13].

The adaptation mechanism reported in this study is very different from adaptation to prisms in humans, for which a strong after-effect is usually obtained, resulting in systematic localization errors immediately after prism removal. Also in owls, adaptation to prisms, as well as to monaural ear plugs, results in a remapping of the auditory spatial map such that, when the prism or earplug is removed, the owl exhibits systematic errors in sound localization. These errors are resolved only by providing adequate (visual) feedback[14].

All four subjects were unaware of the anomaly in their acoustic localization ability while wearing the molds in daily life. This emphasizes once more the power of visual spatial information over the perception of auditory location (also known as the 'ventriloquist effect'[25]) and may have provided the driving force underlying the learning response.

We propose that the adaptive capability of the human auditory localization system is contingent on the availability of a sufficiently rich set of spectral cues, as well as on visual feedback about actual performance in daily life. Although not tested here, active head movements may also have contributed to the calibration of the auditory localization system, especially for those spatial regions where vision has a poor resolution (that is, in the far retinal periphery) or is even absent (for example, for rear stimulus positions and in darkness). Indeed, recent evidence indicates that active head movements are necessary to resolve front–back ambiguities in localization[26].

At present, it is not known which pathways of the human auditory system are responsible for the spectral analysis in elevation detection, nor where the learning in the human auditory system may take place. Future experiments will aim to determine the sensorimotor mechanisms and sites that are involved in these processes.

Methods

SUBJECTS AND EAR MOLDS. Subjects were four adults (the authors and one naïve subject), ages 22–40. Pinna filter functions were measured by recording the sound pressure level in the ear canal, at a location 1–2 mm from the eardrum, with a thin silicone tube attached to a miniature microphone (Knowles, EA1842). Sounds were presented at many different locations surrounding the subject (speaker positioned at five-degree intervals, both in azimuth [-90, +90] degrees, and in elevation [-40, +75] degrees). The sound used to measure the pinna filters was a minimum-peak broad-band FM sweep (0.2-20 kHz, flat amplitude spectrum, Schröder phase[10,13,27]).

Concha molds were prepared by making a negative imprint of both ears by filling both pinnae with plaster. From the hardened negative plaster images, silicone positive replicas of both ears were manufactured. Subsequently, the concha molds were precisely shaped by applying a thin polyester layer (about 0.5–1 mm in thickness) within the concha replicas. Then, a thin layer of wax (about 2 mm) was applied, finished with skin-colored paint (see Fig. 1).

We verified that the modified ear still received specific elevation-dependent spectral features by training a two-layer feedforward neural network to map the pinna filter functions of the modified ear (input layer) onto the elevation domain (one output unit). The trained network yielded a high correlation between required and predicted target elevation.

EXPERIMENTAL CONDITIONS. Auditory spatial locations, as well as first-saccade vector endpoints are described in a double-pole azimuth (α) and elevation (ε) coordinate system[5,13]; α and ε are the directions relative to the vertical median plane and horizontal plane respectively. Auditory stimuli used in the localization experiments consisted of broad-band white-noise sound bursts (0.2–20 kHz, 500 ms, 60 dB SPL), and were presented in complete darkness in an echo-free room at randomly chosen locations within the 2D oculomotor range ([α, ε] within [-30, 30] degrees). The speaker was moved by a two-link robot system that could rapidly position the stimulus at any point on a virtual hemisphere surrounding the subject at a radius of 0.9 m (refs 4,13). Subjects kept their head immobile against a head rest and were instructed to generate a rapid and accurate saccadic eye movement from an initial, centrally presented light-emitting diode toward the perceived location of the sound. No feedback was given about performance. Eye movements were measured with the magnetic search coil technique[28,29]. Earlier studies from our group[4,13] have indicated excellent localization behavior under these open-loop testing conditions (no dynamic spatial cues, as the head was static; see also Fig. 2a).

Acknowledgements

This research was supported by the Dutch Foundation for the Life Sciences (SLW 805-33.705-P; PMH), the University of Nijmegen, the Netherlands (AJVO), and the Human Frontiers Science Program (RG0174/1998-B; AJVO).

RECEIVED 28 MAY: ACCEPTED 7 JULY 1998

1. Middlebrooks, J. C. Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am*. 61, 2607–2624 (1992).

2. Zakarouskas, P. & Cynader, M. S. A computational theory of spectral cue localization. *J. Acoust. Soc. Am.* 94, 1323–1331 (1993).
3. Oldfield, S. R. & Parker, S. P. Acuity of sound localization: a topography of auditory space. I. Normal hearing conditions. *Perception* 13, 581–600 (1994).
4. Frens, M. A. & Van Opstal, A. J. A quantitative study of auditory-evoked saccadic eye movements in two dimensions. *Exp. Brain Res.* 107, 103–117 (1995).
5. Knudsen, E. I. & Konishi, M. Mechanism of sound localization in the barn owl (Tyto alba). *J. Comp. Physiol. A* 133, 13–21 (1979).
6. Batteau, D. W. The role of pinna in human localization. *Proc. R. Soc. Lond. B* 168, 158–180 (1967).
7. Blauert, J. *Spatial Hearing. The Psychophysics of Human Sound Localization.* (MIT Press, Cambridge, Massachusetts, 1996).
8. Teranishi, R. & Shaw, E. A. G. External-ear acoustic models with simple geometry. *J. Acoust. Soc. Am.* 44, 257–263 (1968).
9. Lopez-Poveda, E. A. & Meddis, R. A physical model of sound diffraction and reflections in the human concha. *J. Acoust. Soc. Am.* 100, 3248–3259 (1996).
10. Wightman, F. L. & Kistler, D. J. Headphone simulation of free-field listening. I. Stimulus synthesis. *J. Acoust. Soc. Am.* 85, 858–867 (1989).
11. Oldfield, S. R. & Parker, S. P. Acuity of sound localization: a topography of auditory space. II. Pinna cues absent. *Perception* 13, 601–617 (1984).
12. Wenzel, E. M., Arruda, M., Kistler, D. J. & Wightman, F. L. Localization using nonindividualized head-related transfer functions. *J. Acoust. Soc. Am.* 94, 111–123 (1993).
13. Hofman, P. M. & Van Opstal, A. J. Spectro-temporal factors in two-dimensional human sound localization. *J. Acoust. Soc. Am.* 103, 2634–2648 (1998).
14. Knudsen, E. I. & Knudsen, P. F. Vision guides the adjustment of auditory localization in young barn owls. *Science* 230, 545–548 (1985).
15. Knudsen, E. I. & Knudsen, P. F. Vision calibrates sound localization in developing barn owl. *J. Neurosci.* 9, 3306–3313 (1989).
16. Knudsen, E. I. & Mogdans, J. Vision-independent adjustment of unit tuning to sound localization cues in response to monaural occlusion in developing owl optic tectum. *J. Neurosci.* 12, 3485–3493 (1992).
17. Brainard, M. S. & Knudsen, E. I. Experience-dependent plasticity in the inferior colliculus: a site for visual calibration of the neural representation of auditory space in the barn owl. *J. Neurosci.* 13, 4590–4608 (1993).
18. King, A. J., Hutchings, M. E., Moore, D. R. & Blakemore, C. Developmental plasticity in the visual and auditory representation in the mammalian superior colliculus. *Nature* 332, 73–76 (1988).
19. Withington-Wray D. J., Binns, K. E. & Keating, M. J. The maturation of the superior collicular map of auditory space in the guinea pig is disrupted by developmental visual deprivation. *Eur. J. Neurosci.* 2, 682–692 (1990).
20. Javer, A. R. & Schwarz, D. W. F. Plasticity in human directional hearing. *J. Otolaryngol.* 24, 111–117 (1995).
21. Knudsen, E. I., Esterly, S. D. & Olsen, J. F. Adaptive plasticity of the auditory space map in the optic tectum of adult and baby barn owls in response to external ear modification. *J. Neurophysiol.* 71, 79–94 (1994).
22. Knudsen, E. I. & Knudsen, P. F. The sensitive period for auditory localization in barn owls is limited by age, not by experience. *J. Neurosci.* 6, 1918–1924 (1986).
23. King, A. J. & Moore, D. R. Plasticity of auditory maps in the brain. *Trends Neurosci.* 14, 31–37 (1991).
24. Knudsen, E. I. Capacity for plasticity in the adult owl auditory system expanded by juvenile experience. *Science* 279, 1531–1533 (1998).
25. Stein, B. E. & Meredith, A. M. *The Merging of the Senses* (MIT Press, Cambridge, Massachusetts, 1993).
26. Perrett, S. & Noble, W. The contribution of head motion cues to localization of low pass-noise. *Percept. Psychophys.* 59, 1018–1026 (1997).
27. Schröder, M. R. Synthesis of low-peak factor signals and binary sequences with low autocorrelation. *IEEE Trans. Inform. Theory* 16, 85–89 (1970).
28. Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans. Biomed. Eng.* 10, 137–145 (1963).
29. Collewijn, H., Van der Mark, F. & Janssen, T. J. Precise recording of human eye movements. *Vision Res.* 15, 447–450 (1975).
30. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vettering, W. T. *Numerical Recipes in C*, 2nd edn (Cambridge Univ. Press, Cambridge, 1992).
31. Efron, B. & Tibshirani, R. Statistical analysis in the computer age. *Science Wash. D.C.* 253, 390–395 (1991).