Published in final edited form as:

Hear Res. 2010 February ; 260(1-2): 96. doi:10.1016/j.heares.2009.12.001.

Sound Localization Cues in the Marmoset Monkey

Sean J. Slee and Eric D. Young

Abstract

The most important acoustic cues available to the brain for sound localization are produced by the interaction of sound with the animal's head and external ears. As a first step in understanding the relation between these cues and their neural representation in a vocal new-world primate, we measured head related transfer functions (HRTFs) across frequency for a wide range of sound locations in three anesthetized marmoset monkeys. The HRTF magnitude spectrum has a broad resonance peak at 6-12 kHz that coincides with the frequency range of the major call types of this species. A prominent first spectral notch (FN) in the HRTF magnitude above this resonance was observed at most source locations. The center frequency of the FN increased monotonically from ~12-26 kHz with increases in elevation in the lateral field. In the frontal field FN frequency changed in a less orderly fashion with source position. From the HRTFs we derived interaural time (ITDs) and level differences (ILDs). ITDs and ILDs (below 12 kHz) varied as a function of azimuth between $\pm 250 \mu\text{s}$ and $\pm 20 \text{ dB}$, respectively. A reflexive orienting behavioral paradigm was used to confirm that marmosets can orient to sound sources.

Keywords

ITD; ILD; spectral notch; sound localization; HRTF; marmoset behavior

1. Introduction

The sound localization cues available to the brain are produced by the interaction of sound waves with the head and external ears. Previous studies have shown that the most important cues are interaural level difference (ILD), interaural time difference (ITD), and spectral shape (SS) provided by directionally-dependent filtering of the pinnae. When these cues are imposed on broadband sounds and presented with earphones, they provide sufficient information to reproduce the perception of the direction of an auditory event (Wightman and Kistler, 1989; Middlebrooks and Green, 1991; Middlebrooks, 1999).

Sound localization cues have been measured in several species including human (Harrison and Downey, 1970; Shaw, 1982; Middlebrooks et al., 1989; Wightman and Kistler, 1989; Middlebrooks and Green, 1990), macaque monkey (Spezio et al., 2000), cat (Roth et al., 1980; Musicant et al., 1990; Rice et al., 1992; Tollin and Koka, 2009), marmoset monkey (Aitkin and Park, 1993), ferret (Carlile, 1990; Schnupp et al., 2003), tamar wallaby (Coles and Guppy, 1986), several species of bat (Jen and Chen, 1988; Obrist et al., 1993; Wotton et al., 1995; Fuzessery, 1996; Firzlaiff and Schuller, 2003; Aytekin et al., 2004), rat (Koka et al.,

---

Corresponding Author: Sean Slee, Postdoctoral Fellow, Biomedical Engineering Dept., Johns Hopkins University, 505 Traylor Research Building, 720 Rutland Avenue, Baltimore, MD 21205, phone: 443-287-6186, sslee1@jhmi.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

2008), gerbil (Maki and Furukawa, 2005), guinea pig (Carlile and Pettigrew, 1987; Sterbing et al., 2003), mouse (Chen et al., 1995), and barn owl (Moiseff, 1989; Keller et al., 1998). Here we present additional measurements of sound localization cues in the common marmoset (*Callithrix jacchus*), a vocal new world primate. The marmoset is gaining popularity as a model system for unanesthetized studies in the auditory pathway including the inferior colliculus (Nelson et al., 2009), thalamus (Bartlett and Wang, 2007), and cortex (Lu et al., 2001; Bendor and Wang, 2006; Wang, 2007).

The general characteristics of marmoset head-related transfer functions (HRTFs) are similar to those seen in other animals. Specifically, the greatest variation in both ILD and ITD occur with azimuth and the spectral shape of the HRTF magnitude is directionally dependent. The frequency of the first spectral notch (local minimum) in the HRTF magnitude changes with both source azimuth and elevation. The most orderly variation in the frequency of the first spectral notch occurs with changes in the elevation of lateral source positions.

Because there are no reports of sound localization behavior in marmosets, we measured reflexive orienting movements to sounds varying in azimuth and elevation. The tests were informal using untrained animals and were designed only to verify that marmosets orient to sounds. The behavioral responses suggest that marmosets can perceive the location of sound sources.

2. Methods

2.1. Sound source and stimulus presentation

All experiments were conducted in a double-walled, acoustically-isolated sound chamber (internal dimensions $2.5\text{ m} \times 2.6\text{ m} \times 2\text{ m}$) with 7 cm thick sound absorbing foam (Pinta Acoustics) covering the walls, ceiling, and floor. Sound was presented over a dual speaker system with a high frequency tweeter (Fostex, model FT28D) and a mid range woofer (Morel, model CAW 428) enclosed in a custom built wooden box ($21 \times 20 \times 10\text{ cm}$). The speakers were connected by a 4th order Linkwitz-Riley crossover network (crossover frequency = 2 kHz, -24 dB/octave, Dayton Audio model 260-140) that provided a wideband stimulus over the 0.2-40 kHz range of the marmoset's audiogram. The source was attached to an arc-shaped stand that placed the speakers 1 m from the monkey's head. The tweeter was centered along the arc while the woofer was laterally offset by 8.6 degrees.

The arc was first aligned vertically (relative to the floor) using a level and then bolted to the ceiling of the chamber. A bob on a string was used to align the tweeter at 90° EL, the position of the center of the animal's head, and the center of rotation of the rotary table holding the animal. The center of the tweeter at 0° EL and the position of the eardrum were aligned in the horizontal plane by measuring their heights above the floor with a tape measure and level. At the end of this procedure it was verified that the position corresponding to the center of the animal's head was the same distance from the face of the tweeter at all elevations. During the experiment the animal's head was aligned in the median plane (yaw) visually relative to the speaker arc.

Source locations were in reference to a single-pole coordinate system centered at the monkey's head. Azimuthal (AZ) angles of $\pm 90^\circ$ correspond to the source position directly lateral to the right/left ear and 0° corresponds to the source position directly in front. Angles of elevation (EL) below the horizontal plane are negative and those above are positive. An EL of 90° corresponds to the source position directly overhead. To change the AZ, the monkey was turned on a rotary table connected to a computer-controlled stepper motor with a resolution 0.01 degrees. The EL of the source was adjusted by manually moving the speakers along the arc.

The stimuli were presented from 150° (ipsilateral to the right ear) to -150° AZ in steps of 7.5° or 15°. ELs ranged from -30° to 90° in 17 steps of 7.5°.

The acoustic stimulus was a pair of 8192-point Golay codes (Zhou et al., 1992), sampled at 97.656 kHz, and converted to analog with 12 bit resolution (Tucker-Davis Technologies RP2.1). Experimental stimulus blocks consisted of the 84 ms Golay stimulus repeated 20 times. The Golay stimulus has a spectrum that varies less than 1 dB from 0 to 48 kHz.

2.2. Recording system and signal processing

The pressure waveforms near both tympanic membranes (TMs) were transduced into an electrical signal using hearing aid microphones (Knowles Electronics model FG-23329-C05). In two animals (77T and 9N) each microphone was coupled to the ear through a stainless steel probe tube (1 mm outer diameter, 18 mm length). This probe tube adds to the measurements a first order low-pass frequency response and a peak at around 3 kHz. In the third animal (67S) the microphones were not coupled to probe tubes, but were placed directly in the ear canal. The frequency responses of the microphones in this case were first order low pass at frequencies above ~15 kHz. (See section 2.3 for details of the microphone placement.) The microphone signal was amplified (34 dB with a custom built amplifier), filtered (0.2-40 kHz, -24 dB/octave with a Krohn-Hite model 3100 filter), and digitized (12 bits at a 97.656 kHz sampling rate using a TDT model RP2.1).

“Free-field” measurements (no monkey, with the microphone at the position of the animal's ear) were made in about half the source positions to allow the properties of the speaker, microphone, probe tube (if present), and room to be eliminated from the HRTF. The spectrum of the free-field signal, computed as described in the next paragraph, is shown in Fig. 1C by the gray line. It shows a roll off at low frequencies (<0.15 kHz, not shown) and at high frequencies (>15 kHz) because of the properties of the microphone and electronics. These free-field spectra were very similar across source positions, varying by less than ~3 dB in most cases. For this reason, we averaged the free-field spectra measured between -15° and 30° EL and 135° AZ to -105° AZ and used the averaged free-field function for all calculations except those involving phase. Phases are more sensitive than magnitude to the location and orientation of the microphone in the room, so free-field measurements were taken separately at each position for phase calculations. Free field spectra below -15° EL were affected by low frequency (~700 Hz) reflections caused by the proximity of the woofer to the floor of the chamber. In addition, spectra measured between -105° and -150° AZ contained reflections from the stand holding the animals for some ELs. For these positions the free field spectrum was attenuated by as much as 15 dB at some frequencies (> 10 kHz). HRTFs measured at these positions contain artifacts (not shown).

The signal processing steps are illustrated in Fig. 1. The last 19 of 20 responses to each Golay code were averaged. The average responses to the two codes were processed with the Golay algorithm to calculate the impulse responses of the free-field and ear-canal signals (Rife and Vanderkooy, 1989; Zhou et al., 1992). Figure 1 shows the impulse response with the sound source positioned at 0° AZ, 0° EL. Figure 1A shows the free field response beginning at about 2.7 ms. Room reflections are indicated by the arrow. The reflections induce peaks and valleys in the spectra (Fig. 1C,D); these can be eliminated from the transfer functions by windowing the impulse responses with a 512 point Hanning window centered at 2.56 ms, as shown in Figs. 1A,B. Figure 1B shows the windowed impulse measured in the ear canal, along with the window. The unwindowed impulses transformed with an 8192 point FFT (frequency resolution of 12 Hz) are shown in Fig. 1C, as the magnitudes of the spectra. The HRTF magnitude is the dB difference of the ear canal and the free field spectra, shown by the gray line in Fig. 1D. The small and rapid peaks and valleys are largely caused by the room reflections. The HRTF magnitude computed with the windowed impulse responses is shown by the black line.

Figure 1C also shows the free field spectrum with the probe tube blocked with modeling clay. The separation between the blocked and unblocked free-field spectra is at least 25 dB below 20 kHz and 20 dB below 40 kHz. This level is the noise floor and implies that the minimum gain in the HRTF magnitude that can be measured is about -20 dB.

The HRTFs in both ears were measured for each source position. The ILD at each frequency was determined by subtracting (in decibels) the HRTF magnitude measured in the left ear from that measured in the right ear. Calculation of ITDs is described in connection with Fig. 12. The data presented in Figs. 7, 9, and 12 are plotted on a sphere using a Mollweide equal area projection. The data were interpolated by a factor of 4 before plotting.

2.3. Marmoset preparation

Measurements taken from three marmosets are reported in this study. The animals were young adults (2-5 year old) with healthy external ears. The TMs of each animal were examined with an otoscope and were clear and translucent. Two of the animals were used in an acute protocol while the other was used in a chronic protocol.

The two acute animals were initially anesthetized with intramuscular injections of ketamine (20 mg/kg) and acepromazine (1 mg/kg). A venous catheter was placed in the femoral vein to allow administration of pentobarbital (1-5 mg/kg diluted in 2-5 ml of lactated Ringer with 5% dextrose) as necessary to maintain an areflexive level of anesthesia throughout the course of the experiment (~36 hours). The animals appeared stable in terms of anesthetic use. The animals were given intramuscular injections of atropine sulfate (0.1 mg/12 hr) to reduce mucous and glandular secretions and penicillin-G (30,000 units/24 hr) to reduce the onset of bacterial infections. A tracheotomy was performed and a tracheal tube was implanted. The marmoset's temperature was maintained at 39°C throughout the rest of the experiment with an automatic heating pad system with feedback provided from a rectal probe. Finally, three EKG leads were placed subcutaneously to monitor heart rate.

In the acute animals, a head-post was implanted on the skull 15 mm anterior of the stereotaxic AP 0. In doing this surgery, care was taken not to damage tissue connected to the pinna, to prevent them from sagging out of position. Guide cannulae were inserted subcutaneously into the ear canal through its dorsal wall and held in place with a wire attached to the head-post implant. The probe microphone was inserted through the guide cannula. The final position of the tip of the probe tube was verified with an otoscope. It was about 1 mm inside of the ear canal and a few mm away from the TM (see Fig. 2B). The marmoset was removed from the stereotaxic frame, wrapped in a small blanket and suspended from a ring stand with Velcro straps. The marmoset's head was held in a horizontal orientation at the center of the speaker arc by fixing the head post to a metal rod connected to the ring stand.

Measurements were first taken at the lowest EL (-30°) and over AZs (135° to -135°) at a 15° resolution. The EL was subsequently increased by 7.5° and the next set of measurements was made across AZ. This protocol was repeated at all 17 ELs. The entire procedure took 3-4 hours. In two animals (77T and 9N), the measurements were also made at azimuths ranging from 127.5° to -142.5° at a 15° resolution. In one marmoset (9N), measurements were also made behind the animal (-90° to 90°) at a 7.5° resolution. At the end of the acute experiments, the marmoset was euthanized with an overdose of pentobarbital (150 mg/kg) and transcardially perfused with saline followed by a saline solution containing 0.5% formaldehyde.

Several measurements were repeated in 77T and 9N over the course of the experiment, which lasted about 36 hours. The HRTFs were stable and only varied by a 1-2 dB, except near some notch centers. In 77T the minimum level of a notch centered at 35 kHz increased by 15 dB over the course of the experiment. In 9N the minimum level of a notch centered at 20 kHz

increased by 7 dB over the course of the experiment. The center frequencies of the notches did not change. Middle-ear pressure buildup could potentially make the TM more rigid over time which would tend to increase the magnitude of canal resonances (Guinan and Peake, 1967). This behavior was not observed in the measurements.

The other marmoset (67S) was first implanted with a modified version of the standard head cap used in chronic neurophysiological studies (Lu et al., 2001; Nelson et al., 2009). The standard implant surgery requires partial removal of the temporalis muscles and causes the pinnae to sag downward. Here we used a smaller implant that keeps the temporalis muscles intact and the pinnae closer to their natural position. Following the implant surgery the animal was allowed to recover over several weeks.

To measure HRTFs, the marmoset was first anesthetized with ketamine (30 mg/kg) plus acepromazine (1 mg/kg) and then placed in the acoustic apparatus described above. One additional dose of ketamine (10-30 mg/kg) was administered to maintain a light level of anesthesia. Microphones were placed in the ear directly in front of the TM. The microphones are 2.5 mm long which positions the face of the microphone ~3-4 mm in front of the TM (see Fig. 2B) The spatial cues in the frontal field (-150° AZ to 150° AZ) of the marmoset were measured as described above. Measurements at a few spatial positions were repeated to verify temporal stability. The entire procedure took just over 3 hours. At the end of the procedure the animal was recovered and used subsequently for single-neuron recording experiments.

2.5. Marmoset behavioral methods

Ten marmoset monkeys took part in a behavioral, reflexive sound localization experiment. The monkeys were untrained. The experiments were conducted in the same acoustic chamber described above. Monkeys of both sexes were placed in a semi-restraint primate chair in the center of the chamber. The marmoset primate chair was similar to that described elsewhere (Lu et al., 2001; Nelson et al., 2009). Briefly, the chair is constructed from a tube of Plexiglas with a plate covering the top. The monkey sits comfortably within the tube with its head sticking out of a slit cut in the plate. The Plexiglas plate is a 16 cm wide and 11 cm long. The monkey's chin is a few cm from the plate. The animal is free to turn its head and body but its shoulders, torso, and lower body are restrained beneath the plate. The chair was attached to a support stand that was adjusted in height to align the ear canal with the face of two speakers (Fostex, model FT28D) located at 0° EL and at $\pm 60^{\circ}$ AZ (1.3 m away). A third speaker was placed at 0° AZ (directly in front of the animal) and 60° EL (0.8 m away). The speakers were hidden from view behind acoustically invisible drapes.

The acoustic stimulus was a single noise burst of 100 or 200 ms gated on and off with 5 ms raised cosine ramps. The noise was Gaussian and lowpass filtered at 8 or 30 kHz (Butterworth, -12 dB/octave). The stimulus was presented over one of the three speakers in random order at a level of ~90 dB SPL.

The monkey's head responses were measured with a calibrated accelerometer (for EL; Dimension Engineering, model DE-ACCM3D) temporarily affixed to the top of the monkey's head with hydrophilic vinyl polysiloxane (Examix, GC America, Inc.) (See Fig. 13A). The accelerometer is a tiny microchip that weighs 1.2 grams and is 10 by 21 mm. There were no signs that the accelerometer inhibited the monkeys' orienting responses (head dragging, etc). The AZ of the head orientation was recorded with a video camera (30 frames/sec) mounted above the animal that measured the angle of a line marked on the accelerometer. The output voltage of the accelerometer was digitized at 20 kHz with a TDT A/D converter. The accelerometer output reported the EL when the animal was stationary. During the orienting behavior, the accelerometer reports acceleration due to gravity as well as that of the animal's

movements. For this reason the EL must be recorded during fixation following the completion of the movement.

Each session consisted of 3-4 stimulus trials presented in random order with an inter-stimulus interval of 8-15 minutes. The monkey first sat quietly in the chair in the sealed chamber for ~10 minutes. The chamber had low intensity lighting provided by an LED array. As soon as the monkey appeared to have its gaze focused forward, a stimulus was presented. An infrared LED was illuminated at the stimulus onset to synchronize it to the video record. Both video and accelerometer recordings were collected from 1 second before to 9 seconds after the stimulus. Each animal participated in 3-5 sessions with at least 3 days between sessions. During the second preceding sound presentation, the mean and standard deviation of the accelerometer output were computed (baseline). Movement on a trial was counted if the accelerometer output deviated from the baseline mean by more than 6 times the baseline standard deviation. Head movements were observed in 42 of 80 trials.

3. Results

3.1. The external ear of the marmoset

A scaled drawing of the profile of marmoset 9N is shown in Fig. 2A. Tufts of hair that protrude from the side of the head have been removed to give a clear view of left pinna. The marmoset pinna contains many complicated folds. In addition, the antitragus is particularly prominent. Compared to humans, the marmoset pinna is quite large relative to the animal's head size. For example, the ratio of the pinna width (2.0 cm, measured in the horizontal plane from the top of the tragus to the most lateral part of the helix) to the binaural intertragus distance (3.8 cm) measured on marmoset 67S was 53%. For one of the authors (SJS) the ratio of pinna width (2.5 cm) to the intertragus distance (11.3 cm) was 22%.

A tracing of the cross section of the right ear of marmoset 77T is shown in Fig. 2B. The cross section was cut in the horizontal plane at approximately the level indicated by the gray bar in Fig. 2A. The ear canal has a slight curve and is about 3 mm wide near the tympanic membrane. The white asterisk indicates the approximate position of the probe tube tip.

3.2. Comparison of head related transfer functions (HRTFs) to directional transfer functions (DTFs)

The directional transfer function (DTFs) has been used instead of the HRTF in previous acoustical studies (Middlebrooks et al., 1989; Koka et al., 2008). DTFs are computed from HRTFs by subtracting the average HRTF (computed across source positions) from each HRTF. The average HRTF contains nondirectional components such as the ear canal resonance and any resonances associated with the placement of the microphone relative to the TM. Thus DTFs do not contain these features.

DTFs were computed from the HRTFs measured in 9N. Figure 3A shows the average HRTF in the right ear (green) computed for all source elevations (-30° to 90°) and azimuths ranging from -135° to 135° . The average HRTF has a resonance peaking near 9 kHz and a notch with a minimum around 20 kHz. Inspection of the remaining figures in this paper shows that the peak is universally present, whereas the 20 kHz notch is not, suggesting that the former is an ear canal resonance whereas the latter is created by the pinna and is not due to the microphone placement in the ear canal. Instead, the 20 kHz notch seems to be a common directional component of the HRTF that does not move much with source direction.

Figure 3 shows examples of DTFs (blue) for two sound directions, computed by subtracting the average HRTF (green) from the raw HRTF (red). Both DTFs have a smaller resonance between 6-12 kHz, consistent with the universal character of this resonance. In one case (Fig.

3A, at 0° AZ and 0° EL) the notch at 20 kHz is present in both HRTF and DTF, but is less deep in the DTF; in the other (Fig. 3B, 120° AZ and -7.5° EL), the notch in the HRTF is at a frequency slightly below 20 kHz, so that a peak is created in the DTF by subtracting the average HRTF.

Because of the uncertainties in interpreting the features of the DTF that are apparent from these examples, we show HRTFs in the remainder of this paper. HRTFs are likely to be closer to the actual transfer functions of the ear and therefore more informative about spectral cues for sound localization.

3.3. Head related transfer function show variation across source positions

Head related transfer functions (HRTFs) from one marmoset (9N) are displayed in Fig. 4. The HRTF magnitudes measured in both ears are shown for 15 positions in the frontal field. Several points are worth noting. 1. The gain is near 0 dB for frequencies below 1 kHz in both ears and all source positions. 2. In most locations the HRTF magnitude has a broad resonance in the range of 6 to 12 kHz. The frequency range of this resonance is relatively independent of source position but its overall level is not. 3. For frequencies above 5 kHz the spectral shape of the HRTF magnitude can be described in terms of peaks, notches, and plateaus that vary with source position. A prominent spectral notch is seen in both ears at 20 kHz for many source positions. This notch has large changes in conformation and depth but only small changes in center frequency across the source positions shown. 4. The transfer functions are seen to change in overall level if specific spectral features like notches are ignored. The overall level increases as the source moves ipsilaterally and decreases as the source moves contralaterally. The interaural level differences, evident by comparing spectra in the two ears, are more pronounced at higher frequencies. At higher frequencies, the spectral notches make this comparison more ambiguous. 5. At the same relative position, HRTF magnitudes measured in the two ears are slightly different. For example, for source positions at 0° AZ and all ELs, the peak of the resonance in the left ear is slightly higher than in the right. Both transfer functions have notches at 20 and ~31 kHz but the notch depths are different.

3.4. Transfer functions behave differently in three frequency regions

Figure 5 shows three sets of transfer functions at a fixed AZ measured in the right ear of one marmoset (9N). Each set contains transfer functions with three ELs separated by 15° and are located in front (Fig. 5A), above (Fig. 5B), and behind the marmoset (Fig. 5C). We define three frequencies regions of interest to facilitate comparison between the marmoset transfer functions and previous results in the cat (Rice et al., 1992). The first is the level difference or ΔL region from 1-12 kHz, which is characterized by a broad resonance between 6-12 kHz. This region is likely important for encoding of AZ by ILD as will be shown in the next section. The second region is called the first notch or FN region and occupies the frequency range from 12-24 kHz. It is characterized by the first (lowest frequency) spectral notch, a sharp minimum in the HRTF magnitude often found at 20 kHz as in Fig. 5A. Prominent notches are present in most positions (Fig. 4, 5). Finally, the high frequency or HF region is defined as frequencies above 24 kHz. This region is characterized by high variability in spectral shape with numerous notches and peaks and also high variability in level. As will be shown, the shape of the transfer functions in the HF region varies with location, but also between the two ears in one animal as well as across different animals. A more thorough description of the three regions follows.

3.5. ILDs in the ΔL region vary most strongly with source AZ in the frontal field

The ΔL region is characterized by a broad resonant peak and no spectral notches. Figures 6A and B display HRTF magnitudes measured in increments of 15° AZ within the frontal field of the horizontal plane (0° EL) in the right ears of two monkeys. For frequencies in the ΔL region, the gain of the HRTF decreases as the source moves contralateral to the ear. In addition, the level and peak frequency of the resonance decreases.

The ILD for each source position was calculated by subtracting the transfer function magnitudes (in dB) of the left ear from the right ear (Fig. 6C,D). The ILD at each position tends to increase with frequency up to ~12 kHz where the notches begin. There is a general trend for the ILD to decrease monotonically as the source moves from ipsi- to contralateral to the reference ear. The change in ILD as the source moves from 0° to 30° (frontal field) is greater than the change in ILD as the source moves from 60° to 90° (lateral field). These results are consistent with theoretical as well as experimental results for plane waves incident on a solid sphere (Strutt, 1904; Strutt, 1945; Wiener, 1947; Duda and Martens, 1998). The ΔL region is generally suited to encoding AZ in the classical view that each ILD in a frequency band is directly related to the angle away from the midline. However, at some frequencies, especially above 5 kHz, the ILD functions are more complicated since they do not increase monotonically with AZ.

The ILD at the ΔL frequencies is a weaker cue for EL and for very lateral AZs. Figure 7 plots the ILD in each spatial position averaged across 3 different frequency bands (A: 3-5 kHz, B: 7-9 kHz, C: 17-19 kHz), the first two in the ΔL frequency region. The ILD changes with AZ in an orderly fashion in the frontal field (-60° to 60° AZ) and remains mostly constant with EL. Note that the largest ILD is at -30° for the 7-9 kHz band in Fig. 7B. This behavior was observed in all three marmosets and results from a small notch in the left ear present in this frequency range. At peripheral angles and high frequencies (Fig. 7C), ILD shows substantial change with EL. In this frequency region the appearance of spectral peaks and notches complicates the behavior of ILD. In addition, at the greatest ELs, the ILD is near zero because AZ changes the source position on the sphere (and hence the acoustics) only slightly. These results are consistent with the traditional view that ILD is primarily a cue for AZ.

3.6. Directional dependence of spectral notches in the FN region is most orderly for lateral source positions

The FN region (12-24 kHz) contains the first (lowest frequency) large minimum in the magnitude of the transfer function. In this region both the notch frequency and overall sound level change for some, but not all locations. Figure 8 plots four sets of transfer functions for variation of EL and AZ in both the frontal and ipsilateral regions of space for one marmoset (67S). The insets show the transfer functions on an expanded frequency scale (10-30 kHz). Figure 8A plots five transfer functions for a source fixed at 30° AZ with ELs ranging from -30° to 30°. The FN frequency changes with EL between about 14-20 kHz but not in a smooth and systematic way. Figure 8B shows HRTFs for AZ from -30° to 30° with the EL fixed at 0°. Here, the notches have similar center frequencies. Figure 8C plots transfer functions with the same EL changes as in Fig. 8A but at an AZ of 120° (ipsilateral and behind the marmoset). Here the notch frequency and depth change systematically with increasing source EL. The frequency increases from 14 to 18 kHz while the notch center decreases from -2 to -13 dB. Figure 8D shows transfer functions for changes in AZ from 120° to 60° at an EL of 0°. As in the frontal field, the notch frequency is relatively constant over a large range of AZs with the exception of the transfer function measured at 60°. This behavior was similar at some but not all ELs.

Spatial maps of FN frequency measured in the right ears of all three monkeys are plotted in Fig. 9. The maps are shown for all ELs at AZs ranging from 135° to -30°. The following procedure was used to find and measure FN frequencies. First, a candidate FN was detected by finding the first local minimum in the HRTF magnitude spectrum. Second, the quality factor (FN frequency / bandwidth) of the notch was measured 5 dB above the minimum. The FN was counted if the quality factor was greater than 5. If not, the next local minimum in the HRTF was considered. Some HRTFs did not contain spectral notches that met this criterion; these are shown as white regions in the figure. As in Fig. 8C, there is a monotonic increase in FN

frequency with EL, for lateral source positions (AZ larger than 60°-80°); FN frequency is relatively constant for changes in AZ in this region. In the frontal field and on the contralateral side, the FN frequency is often relatively constant (Fig. 9A,C), but can show disorderly, almost random, variations with AZ and EL (Fig. 9B). The same general behavior was seen in the left ears (see Fig. 11) and in the DTFs from marmoset 67S. The map of FN frequencies suggests that first-notch frequency provides the most unambiguous information about the EL of lateral sources on the ipsilateral side.

3.7. High frequency region

The HF region (>24 kHz) is characterized by a large variation in spectral level and shape. In some source locations the magnitudes of the transfer functions may be smooth (Fig. 8B) but usually there are many peaks and notches (Fig. 8C,D). In some regions there are systematic changes in notch frequency (Fig. 8A) or overall level (Fig. 8D). However, these changes only occur in limited regions of space. In other regions there are dramatic changes in shape and level with only small changes in source position. Furthermore, HRTF magnitudes in the HF region at the same location often have large variation between the ears of one animal or across ears from different animals (Fig. 11).

3.8. Transfer functions on the median plane

Spectral shape cues can help to disambiguate source locations for which binaural cues are similar. In theory, ambiguous cues are found along a “cone of confusion” (Blauert, 1997), but Fig. 7 shows that the idea of a cone is not usually accurate for the marmoset ear. One situation where the ambiguity of the binaural cues is considerable is in the median plane, where ILDs are near zero for frequencies below ~15 kHz (Figs. 6C, D and Fig. 7); in a subsequent section, we show that ITDs are also ~0 near the median plane (Fig. 12). To extend the analysis of the median plane, transfer functions were measured for source positions behind one animal (9N). Figure 10 shows transfer functions measured in the right ear for locations on the median plane. Considerable differences are seen in the spectral shape cues at frequencies above 18 kHz. For example the position and/or level of the first spectral notch is unique to each location. In addition, over some frequencies in the HF region there is a decrease in overall level as the source moves from directly ahead, to above, to directly behind. In rear positions, this may result from diffraction of sound off the back of the head and pinnae. Thus both notch frequency and overall level in the HF region may provide cues for localization in the median plane.

3.9. Comparison of data across individuals and ears

In this study data are presented from both ears of 3 marmosets. Figure 11 plots transfer function magnitudes measured in all 6 ears overlaid for a few source positions. In general, the spectra are similar between animals and ears in that notches are located at roughly the same frequencies and the overall shapes are similar. However, there are many differences in details. For example the first-notches are consistently at lower frequencies for 67S (red) compared to 77T (blue) and 9N (green). This could be the result of the smaller pinna size of 67S (length, L=2.5 cm, width, W=2.0 cm) compared to 77T (L=2.5 cm, W=2.2 cm) and 9N (L=2.7 cm, W=2.1 cm).

In front of the animal (Fig. 11A and B), the sound is coming from the same direction in the two ears and the HRTFs should be identical. As noted in Fig. 4 and 6, this is not exactly so. At lateral and rear source positions (Fig. 11C,D) prominent FN features are clear in all 6 ears and all increase in frequency as EL increases, consistent with the trend discussed previously (Fig. 9). At most source positions, there is more variability in the FN region for frontal sources than for lateral ones.

3.10. Interaural time differences vary most strongly with source AZ

ITD for each source location was calculated from the phases of the impulse responses measured in the two ears. The interaural phase difference (IPD) at each frequency was computed by subtracting the phase in the left ear from the right. The result was unwrapped to produce a phase difference versus frequency function (Fig. 12A for marmoset 67S). The functions were approximately linear over the range of 0.2-10 kHz, consistent with a simple time delay. At very low frequencies (<0.2 kHz) the phase becomes noisy because of the dropoff of energy in the speaker outputs at low frequencies. At higher frequencies, there are irregularities (steps) in the phase functions having to do with the rapid phase changes near resonances in the HRTFs. This can be seen by comparing the frequencies at which steps in the phase functions occur in Fig. 12A with the frequencies of notches in the HRTF magnitudes in previous figures.

The ITD at each AZ was estimated as the slope of a linear fit to the phase versus frequency function over the range 0.2 - 10 kHz. This provides an estimate of the interaural group delay, but given the linear nature of the phase functions over this range, the answer would be similar if phase delay were estimated as phase divided by frequency. When examined in detail, there are small irregularities in the phase functions shown in Fig. 12A. However, there is no systematic variation of ITD with frequency that is apparent in these data (for example if the slopes are estimated over subsets of the full frequency range).

Figure 12B shows ITDs measured in the three marmosets for 0° EL and AZs in the horizontal plane ranging from -90° to 90°. The ITD averaged over all the three animals decreases in an orderly fashion from about +188 to -183 μsec as the source moves from the ipsi- to contralateral side. The black curve is the theoretical prediction of ITDs from a geometric model of sound incident on a sphere (Woodworth, 1938):

$$ITD = (r/c) * (\theta + \sin\theta)$$

The sphere radius r is 2.5 cm and c is the speed of sound (343 m/s). This radius was larger than the radii measured on the marmosets, taken as half the inter-tragus distance (67S = 1.9 cm, 9N = 1.9 cm, 77T = 1.8 cm). This result shows that the empirical range of ITD is larger than the theoretical range based on the measured head size. Both the head shape and the pinnae may contribute to the deviation from the theory. The shape of a marmoset head is closer to an ellipsoid than to a sphere. (Marmoset 9N: major axis (face to the back of the head) = 5.1 cm, minor axis (inter-tragus distance) = 3.8 cm). In addition, a recent study suggests that the pinnae contribute to ITDs (Koka et al., 2008).

Figure 12C plots a spatial map of ITD, in the same monkey as in Fig. 12A (67S). As discussed above, ITD decreases as the source moves in AZ from the ipsi- to contralateral sides at most ELs. In the frontal field, ITD is mostly constant with EL at a fixed AZ. In addition, ITD changes more rapidly for sources in the frontal than in the lateral field. The spatial pattern of ITD is similar to that of ILD (Fig. 7). ITD thus likely provides a cue primarily for the AZ of a sound source.

3.11. Behavioral orienting to lateral and elevated sound sources

The data presented thus far suggest that the three primary sound localization cues are available to the marmoset. However, sound localization behavior in the marmoset has not been reported. Therefore we measured reflexive sound localization behavior in 10 marmoset monkeys. Reflexive orienting of untrained animals was measured with head tracking as described in the Methods section. The 3 sound sources were located at +/-60° AZ, 0° EL and 0° AZ, 60° EL. Fig. 13A plots the noise stimulus, located at 0° AZ and 60° EL, as well as the EL of the monkey's

orienting response recorded with the accelerometer. A small response is present a few hundred ms after stimulus onset, however the initial orientation has a latency of over a second. Here the final EL was taken as the average over the time interval from 5-6 sec (41.2°) where the head was stationary.

In total 80 stimuli from the 3 locations were presented to the 10 monkeys. Of those, 42 resulted in a detectable behavioral response as measured by a change in the output signal from the accelerometer (see Methods). In general the animals responded on 1-2 trials of the 3-4 trials presented during each session. The monkeys quickly habituated: most animals did not respond at all during the 4th or 5th sessions. The response latencies varied widely with a range from 35 ms to 2.79 seconds (mean $\pm$ standard error of the mean = 500 ± 95 ms). On some trials the orienting was both fast and accurate, reflexive in quality. On other trials, the response appeared more like a slow and casual glance in the direction of the perceived stimulus. Always the first movement occurring within 3 s of the stimulus was taken as the response. The behavior between the stimulus presentations suggested that the animals' level of attention was highly variable

Figure 13B summarizes the detectable behavioral responses to stimuli presented at $\pm 60^\circ$ AZ, 0° EL. The noise stimulus to the right of the monkey was low-pass filtered at 30 kHz (red) while that to the left was low-pass filtered at 8 kHz (blue). This was done in the hopes of demonstrating an effect of the availability of spectral cues, but the low-pass filtering made no differences in behavior. Each arrow and pair of points corresponds to the initial and final gaze positions. The extent of the responses varied greatly but the AZ component of the response was always in the correct direction. In most cases, the final gaze position fell short of the target. Some responses also had a change in EL associated with them and these were usually inaccurate. The inset of Fig. 13B shows the absolute change in EL versus the change in AZ normalized by the initial distance to the target. Points near 1 on the abscissa correspond to accurate horizontal gaze shifts in this plot. In all cases except one, the AZ of the gaze-shift covered 40-100% of the distance to the target while EL shifts were scattered (normalization of the ordinate by the initial EL offset is not shown because of the large scatter in those data and the often very small initial EL offsets). These results demonstrate that marmosets are able to lateralize sound sources to the correct side with both types of stimuli.

Behavioral responses to the source at 0° AZ and 60° EL are summarized in Fig. 13C. Again, the stimuli were either low-pass filtered at 8 (blue) or 30 kHz (red). Orienting responses to this location were in general less accurate than the responses to AZ in Fig. 13B. All of the responses fell short of the target, although some of the responses did have a large vertical component and appeared close to accurate. For example the response shown in Fig. 13A ends at 35° AZ and 41.2° EL. However, other responses were completely inaccurate and largely composed of a horizontal component (e.g. the trajectory ending at -30° AZ and -36° EL). The inset of Fig. 13C shows the change in EL normalized by the initial distance to the target plotted against the change in AZ. The inset shows EL responses scattered between -20 and 100% of the target offset, a larger range than for AZ in Fig. 13B. Although the horizontal component of the responses to the elevated source were inaccurate and both too short or too long, they were usually directed in the correct initial direction (22/25).

4. Discussion

Validity of HRTF measurements

HRTF measurements are made with the assumption that a point pressure measurement near the TM is indicative of the pressure across the TM and therefore the input to the auditory system. In the human or cat ear canal this is likely to be a good approximation (discussed by Rice et al. 1992). The cylindrical canal supports a plane wave over its central portion and propagates the sound passing through the external ear, including the directionally-dependent

filtering of the sound by the pinna. Reflections from the complex shape of the pinna produce evanescent multidimensional modes that theoretically do not propagate down the ear canal, for canals with the dimension of the cat or human ear (Rabbitt and Holmes 1988). Because the marmoset canal (Fig. 2B) is smaller than the cat or human canal, the theory also predicts that evanescent modes do not propagate in the marmoset canal at frequencies of interest. Thus, the sound field should be uniform across the marmoset canal making it likely that pressure measurements in the canal are a good indication of the input to the cochlea.

Comparison of HRTF magnitudes measured in the 6 ears of 3 marmosets (Fig. 11) reveals that they are similar, especially for ipsilateral sources and frequencies below ~24 kHz. These were measured with implanted probe tubes (two animals) or with microphones placed in front of the TM (one animal). The similarity of the results with different measuring arrangements suggests that signals that depend on the probe placement are not a dominant component of our data set and lends some confidence that our results are meaningful.

The HRTF measurements were made with anesthetized animals, which presumably caused relaxation of the musculature supporting the pinnae. This introduces the possibility of differences in the acoustics compared to the awake animal. These differences are probably small because the marmosets' ears did not sag noticeably under anesthesia and repeated measurements over the course of the experiments were stable.

Comparison with previous studies

There is a resonance in the marmoset HRTF peaking between 6-12 kHz. The frequency of this peak is higher than in cats and humans, presumably because of the shorter length of the marmoset ear canal (Shaw, 1974; Musicant et al., 1990; Rice et al., 1992). The marmoset HRTF resonant frequency is adjacent to the region of spectral notches; this may explain the apparent lowering in the peak frequency as the sound source moves contralaterally and the small notch-like irregularities that appear in the HRTF at frequencies just above the resonance (e.g. Fig. 6). This behavior is different from humans and cats where the ear-canal resonance is at a lower frequency, giving a nondirectional resonant peak (Shaw, 1982).

Nevertheless, it is interesting that the resonance in the HRTF magnitude spectrum peaks in a range where several marmoset vocalizations have a high energy content (DiMattina and Wang, 2006). The resulting head-shadowing effect could help improve the signal to noise ratio of species-specific vocalizations in noisy listening environments for the ear ipsilateral to the signal source.

ILDs show the most variation with the AZ of a sound source. This result is similar to that described in other species, with the exception of the barn owl where the facial ruff creates ILDs that vary most strongly with EL (Moiseff, 1989). The range of ILDs in the ΔL region is $\sim \pm 20$ dB. This range is similar to that reported in the marmoset monkey previously (Aitkin and Park, 1993).

The first spectral notch in the HRTF magnitude of the marmoset occurs between 12-24 kHz. This range is higher than in human (Shaw and Teranishi, 1968; Middlebrooks et al., 1989) and cat (Rice et al., 1992) which is expected from previous work that demonstrated a linear scaling of first spectral notch frequencies with the dimensions of the pinna (Middlebrooks, 1999). The FN center frequency increases with source EL in the lateral field, but changes only gradually with AZ. This behavior is similar to humans (Shaw and Teranishi, 1968) but is in stark contrast to the cat, which shows notch progression for a change in AZ or EL in the frontal field (Rice et al., 1992). In the cat, measurements were made with the ears in the relaxed position with the opening of the pinna flange facing about 45° AZ. The fixed position of the marmoset pinna is

closer to 90° AZ, which might explain the difference in the foveal regions reported in the two studies.

HRTF magnitudes measured in the frontal field of the marmoset have the same FN center frequency (~20 kHz) for multiple locations (Figs 4-6, 8, and 9). Therefore other information besides FN center frequency must be utilized to remove ambiguity among these locations. Binaural ITDs and ILDs are likely cues for AZ discrimination. Possibilities for EL discrimination include the depth, edge, or slope of the first notch or changes in spectral cues in the HF region.

The ITDs measured in the marmoset range between ± 250 μ sec and provide a robust localization cue for AZ. Marmosets should be able to use this cue, because their hearing extends down to low frequencies, about 100 Hz (Seiden, 1958). The range of ITDs is small, consistent with other animals with a small head size. However, the range of ITDs is greater than the theoretical prediction for a sphere the size of the marmoset head. This discrepancy may be due to the influences of the pinnae, which have been shown to increase ITD in other species (Koka et al., 2008). Previous theoretical and empirical evidence reveals that ITDs on a sphere are frequency dependent. Specifically, at some AZs the ITD at low frequencies is about 50% greater than the ITD at higher frequencies with a transition region in between (Kuhn, 1977). We did not detect this type of frequency dependence of ITD in our measurements. It is possible that a spherical model is not a good approximation for marmosets since their pinnae are very large relative to their heads.

ITDs at frequencies near spectral notches are complicated because of the rapid changes in the group delay. Humans do not detect ITDs in the carrier frequencies above about 1.6 kHz (Blauert, 1997). However, the ITD of envelopes imposed on a high frequency carrier can be used for localization, presumably by detecting delays in the envelope of the signal (Henning, 1974). Interpretation of this cue is complicated by the rapid interaural phase changes produced by ear resonances at high frequencies (Fig. 12A). However, envelope delays are primarily dependent on group delays, the slope of the phase-frequency function. Group delays retain their correspondence to AZ at frequencies away from the external ear resonances.

Reflexive Sound Localization in Marmosets

The behavioral results suggest that marmosets can localize sound sources at least crudely in both AZ and EL. The animals were untrained and thus the results reflect their natural abilities based only on experience with sounds in their environment. The responses often fell short of the targets. A possible explanation for this could be that the eye also moved within the head so that the final eye position was on target while the gaze position fell short of the target. The responses to lateral targets were more accurate than responses to elevated targets. However, it is evident from the responses to the elevated sources (Fig. 13C) that the marmosets were usually able to detect that the source was elevated above the horizontal plane for both low-pass filters. Neither ILD (Fig. 7), nor ITD (Fig. 12) provide an accurate cue for the EL of a sound source. Thus it is likely that the animals had to use spectral shape cues (Figs. 4-11) to locate the elevated source. The responses may be less accurate than for the lateral sources because the short duration of the stimulus (100 or 200 ms) does not allow enough time to accurately estimate the source spectrum. The result that a few responses to the stimulus filtered above 8 kHz were as accurate as the responses to the 30 kHz stimulus is intriguing. The clearest spectral cues occur in the FN region from 12-24 kHz (Figs. 8,9). For this reason one might expect that lowpass filtered sounds, with little energy in this region, might be difficult to localize in EL (as in highly trained cats) (Huang and May, 1996). However, psychophysical results suggest that humans can use very subtle spectral information in highly filtered sounds to make accurate localization judgments, even for sounds with little high-frequency energy (Blauert, 1997). It may be that

the relatively non-selective (-12 dB/oct) low-pass filters used here left enough energy in the stimulus to allow the marmosets to estimate elevation, at least to the extent of up versus down.

Acknowledgments

We thank Dr. Bradford May for assistance with surgical design and the behavioral experiments, Ron Atkinson, Jay Burns, Phyllis Taylor, and Qian Gao for technical assistance, and Jenny Estes and Judy Cook for assistance with animal care. This work was supported by grants DC00115 and DC00023 from NIDCD.

References

- Aitkin L, Park V. Audition and the auditory pathway of a vocal New World primate, the common marmoset. *Prog Neurobiol* 1993;41(3):345–67. [PubMed: 8210411]
- Aytekin M, Grassi E, et al. The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation. *J Acoust Soc Am* 2004;116(6):3594–605. [PubMed: 15658710]
- Bartlett EL, Wang X. Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *J Neurophysiol* 2007;97(2):1005–17. [PubMed: 17050830]
- Bendor D, Wang X. Cortical representations of pitch in monkeys and humans. *Curr Opin Neurobiol* 2006;16(4):391–9. [PubMed: 16842992]
- Blauert, J. Spatial hearing. Cambridge: MIT; 1997.
- Carlile S. The auditory periphery of the ferret. II: The spectral transformations of the external ear and their implications for sound localization. *J Acoust Soc Am* 1990;88(5):2196–204. [PubMed: 2269735]
- Carlile S, Pettigrew AG. Directional properties of the auditory periphery in the guinea pig. *Hear Res* 1987;31(2):111–22. [PubMed: 3446669]
- Chen QC, Cain D, et al. Sound pressure transformation at the pinna of *Mus domesticus*. *J Exp Biol* 1995;198(Pt 9):2007–23. [PubMed: 7595162]
- Coles RB, Guppy A. Biophysical aspects of directional hearing in the tammar wallaby, *Macropus eugenii*. *J Exp Biol* 1986;121:371–394.
- DiMattina C, Wang X. Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations. *J Neurophysiol* 2006;95(2):1244–62. [PubMed: 16207780]
- Duda RO, Martens WL. Range dependence of the response of a spherical head model. *J Acoust Soc Am* 1998;104:3049–3058.
- Firzlaff U, Schuller G. Spectral directionality of the external ear of the lesser spear-nosed bat, *Phyllostomus discolor*. *Hear Res* 2003;185(12):110–22. [PubMed: 14599698]
- Fuzessery ZM. Monaural and binaural spectral cues created by the external ears of the pallid bat. *Hear Res* 1996;95(12):1–17. [PubMed: 8793503]
- Guinan JJ Jr, Peake WT. Middle-ear characteristics of anesthetized cats. *J Acoust Soc Am* 1967;41(5):1237–61. [PubMed: 6074788]
- Harrison JM, Downey P. Intensity changes at the ear as a function of the azimuth of a tone source: a comparative study. *J Acoust Soc Am* 1970;47(6):1509–18. [PubMed: 4987621]
- Henning GB. Detectability of interaural delay in high-frequency complex waveforms. *J Acoust Soc Am* 1974;55(1):84–90. [PubMed: 4815755]
- Huang AY, May BJ. Spectral cues for sound localization in cats: effects of frequency domain on minimum audible angles in the median and horizontal planes. *J Acoust Soc Am* 1996;100(4 Pt 1):2341–8. [PubMed: 8865641]
- Jen PH, Chen DM. Directionality of sound pressure transformation at the pinna of echolocating bats. *Hear Res* 1988;34(2):101–17. [PubMed: 3170353]
- Keller CH, Hartung K, et al. Head-related transfer functions of the barn owl: measurement and neural responses. *Hear Res* 1998;118(12):13–34. [PubMed: 9606058]
- Koka K, Read HL, et al. The acoustical cues to sound location in the rat: measurements of directional transfer functions. *J Acoust Soc Am* 2008;123(6):4297–309. [PubMed: 18537381]
- Kuhn GF. Model for the interaural time differences in the horizontal plane. *J Acoust Soc Am* 1977;62:157–167.

- Lu T, Liang L, et al. Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *J Neurophysiol* 2001;85(6):2364–80. [PubMed: 11387383]
- Lu T, Liang L, et al. Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci* 2001;4(11):1131–8. [PubMed: 11593234]
- Maki K, Furukawa S. Acoustical cues for sound localization by the Mongolian gerbil, *Meriones unguiculatus*. *J Acoust Soc Am* 2005;118(2):872–86. [PubMed: 16158644]
- Middlebrooks JC. Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency. *J Acoust Soc Am* 1999;106(3 Pt 1):1493–510. [PubMed: 10489706]
- Middlebrooks JC, Green DM. Directional dependence of interaural envelope delays. *J Acoust Soc Am* 1990;87(5):2149–62. [PubMed: 2348020]
- Middlebrooks JC, Green DM. Sound localization by human listeners. *Annu Rev Psychol* 1991;42:135–59. [PubMed: 2018391]
- Middlebrooks JC, Makous JC, et al. Directional sensitivity of sound-pressure levels in the human ear canal. *J Acoust Soc Am* 1989;86(1):89–108. [PubMed: 2754111]
- Moiseff A. Binaural disparity cues available to the barn owl for sound localization. *J Comp Physiol [A]* 1989;164(5):629–36.
- Musicant AD, Chan JC, et al. Direction-dependent spectral properties of cat external ear: new data and cross-species comparisons. *J Acoust Soc Am* 1990;87(2):757–81. [PubMed: 2307774]
- Nelson PC, Smith ZM, et al. Wide-Dynamic-Range Forward Suppression in Marmoset Inferior Colliculus Neurons Is Generated Centrally and Accounts for Perceptual Masking. *J Neurosci* 2009;29:2553–2562. [PubMed: 19244530]
- Obrist MK, Fenton MB, et al. What ears do for bats: a comparative study of pinna sound pressure transformation in chiroptera. *J Exp Biol* 1993;180:119–52. [PubMed: 8371084]
- Rice JJ, May BJ, et al. Pinna-based spectral cues for sound localization in cat. *Hear Res* 1992;58(2):132–52. [PubMed: 1568936]
- Rife DD, Vanderkooy J. Transfer-function measurement with maximum-length sequences. *J Audio Eng Soc* 1989;37:419–444.
- Roth GL, Kochhar RK, et al. Interaural time differences: implications regarding the neurophysiology of sound localization. *J Acoust Soc Am* 1980;68(6):1643–51. [PubMed: 7462463]
- Schnupp JW, Booth J, et al. Modeling individual differences in ferret external ear transfer functions. *J Acoust Soc Am* 2003;113(4 Pt 1):2021–30. [PubMed: 12703713]
- Seiden, HR. PhD dissertation. Princeton NJ: Princeton University; 1958. Auditory acuity of the marmoset monkey (*hapale jacchus*).
- Shaw EA. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *J Acoust Soc Am* 1974;56(6):1848–61. [PubMed: 4443484]
- Shaw EA, Teranishi R. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *J Acoust Soc Am* 1968;44(1):240–9. [PubMed: 5659838]
- Shaw, EAG. External ear response and sound localization. In: Gatehouse, RW., editor. *Localization of sound: theory and applications*. Groton: Amphora Press; 1982.
- Spezio ML, Keller CH, et al. Head-related transfer functions of the Rhesus monkey. *Hear Res* 2000;144(12):73–88. [PubMed: 10831867]
- Sterbing SJ, Hartung K, et al. Spatial tuning to virtual sounds in the inferior colliculus of the guinea pig. *J Neurophysiol* 2003;90(4):2648–59. [PubMed: 12840079]
- Strutt JWLR. On the acoustic shadow of a sphere. *Philos Trans R Soc Lond Ser A* 1904;203:87–89.
- Strutt, JWLR. *The theory of sound*. New York: Dover; 1945.
- Tollin DJ, Koka K. Postnatal development of sound pressure transformations by the head and pinnae of the cat: monaural characteristics. *J Acoust Soc Am* 2009;125(2):980–994. [PubMed: 19206874]
- Wang X. Neural coding strategies in auditory cortex. *Hear Res*. 2007
- Wiener FM. Sound diffraction by rigid spheres and circular cylinders. *J Acoust Soc Am* 1947;19:444–451.
- Wightman FL, Kistler DJ. Headphone simulation of free-field listening. II: Psychophysical validation. *J Acoust Soc Am* 1989;85(2):868–78. [PubMed: 2926001]

Woodworth, RS. Experimental psychology. New York: Holt; 1938.

Wotton JM, Haresign T, et al. Spatially dependent acoustic cues generated by the external ear of the big brown bat, *Eptesicus fuscus*. J Acoust Soc Am 1995;98(3):1423–45. [PubMed: 7560511]

Zhou B, Green DM, et al. Characterization of external ear impulse responses using Golay codes. J Acoust Soc Am 1992;92(2 Pt 1):1169–71. [PubMed: 1506522]

Abbreviations

HRTF	head-related transfer function
DTF	directional transfer function
AZ	azimuth
EL	elevation
ITD	interaural time difference
ILD	interaural level difference
FN	first notch
HF	high frequency
dB	decibel
SS	spectral shape
TM	tympanic membrane
EKG	electrocardiogram
sec	second
hr	hour
mg	milligram
kg	kilogram
cm	centimeter
mm	millimeter
L	length
W	width
r	radius
Fig.	figure
ml	milliliter
kHz	kilohertz
ms	millisecond
A/D	analog to digital

Figure 1.

Signal processing steps to compute the HRTF. A. Impulse response computed using the Golay method of a speaker located at 0° AZ and 0° EL in free field. The microphone and probe tube were placed at approximately the same position in the room as if in the ear canal, but no animal was present. Reflections from the wall are evident. The Hanning window used to smooth the signals is shown. B. Impulse response for the same speaker and microphone placement, except recorded in the ear canal. The impulse response is shown after the window was applied. C. Spectra of the un-windowed impulse responses in the ear canal and free field, as the magnitude of the Fourier transforms of the signals. The dB scale is arbitrary, with 0 dB at the peak amplitude. The noise floor was measured with the probe tube filled with modeling clay. D. Magnitude of the HRTF for the signals in C unwindowed (gray line) and windowed (black line).

Figure 2.

The external ear of the marmoset. A. A scale sketch of the left profile of marmoset 9N. Hair tufts have been removed to expose the pinna. B. A tracing of a horizontal section through the right ear (marmoset 77T), looking downward from above. The middle and inner ear are also shown. The approximate height of the section is indicated by the gray line in Fig. 1A. The white asterisk in the ear canal indicates the approximate position of the probe tube during the measurements. (A = anterior, M = medial.)

Figure 3.

Comparison between HRTFs and DTFs. A. The individual HRTF (red), average HRTF (green), and individual DTF (blue) measured in the right ear at 0° AZ and 0° EL in marmoset 9N. B. The individual HRTF (red), average HRTF (green), and individual DTF (blue) measured in the right ear at 120° AZ and -7.5° EL.

Figure 4. Magnitudes of HRTFs from 15 speaker positions, as indicated in the figure. HRTFs are shown for both ears, the right ear as the black line (AZ is positive to the right) and the left ear as the red line.

Figure 5.

Magnitudes of HRTFs from three spatial regions, in front of the animal (A), above the animal (B), and behind the animals (C). In each case a notch-free ILD region is present at frequencies below 12 kHz, a first-notch (FN) region between 12-24 kHz, and a high frequency (HF) region above 24 kHz.

Figure 6.

A, B. Magnitudes of HRTFs for a range of AZs, identified in the legend, at 0° EL. Data from two animals are shown. C, D. ILDs computed from the data in A and B as the difference of the HRTFs in A and B at corresponding left and right AZs.

A. Marmoset 67S, 3-5 kHz

B. Marmoset 67S, 7-9 kHz

C. Marmoset 67S, 17-19 kHz

Figure 7.

ILDs averaged across three frequency bands are shown as a function of AZ and EL plotted on a Mollweide equal area projection. The dots indicate the location of the ear canals. A. B. Results for frequencies in the ΔL region. C. Results from the FN frequency region.

Figure 8.

Examples of spectral cues in four regions of space. The main plots show magnitudes of HRTFs measured at locations in the frontal field (A,B) and the lateral field (C,D) for changes in EL (A,C) and AZ (B,D). The FN region is expanded in the insets.

Figure 9.

Spherical contour plots of FN frequency as a function of AZ and EL for the three animals. These are for HRTFs in the right ear (dot). The regions without reliable FNs, as defined in the text, are blank. Some regions contain high frequency FNs that saturate the color scale (red).

Figure 10.

Magnitudes of HRTFs for five positions in the median plane, showing the spectral cues that accompany the mostly zero binaural cues.

Figure 11.

Similarity of HRTFs across different animals. Magnitudes of HRTFs are shown for four source positions in three different animals, coded by color. Data for the left (dotted lines) and right (solid lines) ears are shown. In Fig. 11C,D the azimuth ($\pm 105^\circ$) is ipsilateral to the ear.

Figure 12.

ITD measurements. A. IPD versus frequency measured at different AZs in the frontal field (0° EL). ITDs are the slopes of the lines. Note the discontinuities in the FN region. B. ITD versus AZ in the frontal field (0° EL) measured in 3 monkeys (colored lines). The black curve is a theoretical prediction from the Woodward model for a head radius of 2.5 cm. C. A contour plot of ITD at different locations relative to the ears (dots).

Figure 13.

Reflexive sound localization behavior in marmosets. A, Example of the elevation change measured during a gaze shift in response to a noise stimulus (top) located at 0° AZ and 60° EL. The green box indicates the movement and the red line indicates the period over which the response was calculated. B, Behavioral gaze shifts to sources located at 0° EL and +60° (red) and -60° (blue) AZ. The noise stimulus was lowpass filtered at 8 (blue) or 30 kHz (red). The arrows indicate the movement direction between the initial and final position (dots). The inset shows ΔEL versus ΔAZ normalized by the initial distance to the target. C, Behavioral gaze shifts to sources located at 60° EL and 0° AZ. The noise stimulus was lowpass filtered at 8 (blue) or 30 kHz (red). The arrows indicate the movement direction between the initial and

final position (dots). The inset shows the ΔEL normalized by the initial distance to the target versus ΔAZ . One response is not shown in the inset. The response was to the noise lowpass filtered at 30 kHz and had a $\Delta EL/\text{distance to target} = -1.27$ and $\Delta AZ = 50$.