

Sound Localization in Noise by Gerbils and Humans

ANDREA LINGNER¹, LUTZ WIEGREBE¹, AND BENEDIKT GROTHE¹

¹*Division of Neurobiology, Department Biology II, Ludwig-Maximilians-Universitaet Munich, Großhaderner Str. 2-4, 82152 Martinsried-Planegg, Germany*

Received: 10 March 2011; Accepted: 25 October 2011; Online publication: 14 January 2012

ABSTRACT

Detection and localization of a target sound in the presence of concurrent, spatially distributed masking sounds is one of the most challenging tasks for the mammalian auditory system. Previous studies demonstrated that the ability to localize signals is decreased by interfering noise. In order to directly compare the behavioral performance in a signal processing task in noise between gerbils and humans in the free sound field, we quantified their localization ability for a low-frequency signal in the presence of six masking noise sources surrounding the subject. Thresholds were measured both for masking noises that were correlated or uncorrelated across the masking sources. Overall, the gerbils required a higher signal/noise ratio to detect the low-frequency signal than the humans; that is, the behavioral performance of the gerbils was considerably worse than that of the humans. Moreover, switching from maskers that were uncorrelated across the masking sources to correlated maskers resulted in more masking in gerbils but in a release from masking in humans. These results would suggest that the gerbil may not be a good animal model for binaural processing. However, simulations of the localization thresholds in a numerical model of binaural processing in gerbils and humans reveal that both the inferior overall performance in gerbils and the opposite effect of masker correlation on the detection thresholds can be attributed to the smaller head size and the wider peripheral auditory filters in gerbils. Thus, the current data indicate that the binaural processor itself (i.e., the evaluation of signals

coming from the two ears) is equally sensitive in gerbils and humans. However, the physical limitations imposed by the small head prevent the gerbil from performing equally well in the current paradigm.

Keywords: unmasking, BMLD, binaural hearing, signal detection, binaural modeling

INTRODUCTION

All creatures that rely on hearing have to deal with the detection and localization of vitally important sounds in a noisy environment. Coping with this task is one of the biggest challenges for the auditory system. A phenomenon called the “cocktail party effect,” first introduced by Cherry (1953), describes the fact that humans are capable of focusing on a single sound source and thus separating a single person talking from a mixture of interfering voices. Cherry suggested various potential factors responsible for this ability. Among others, he proposed that the spatial position of the signal and interferers could play an important role in localizing a signal in a noisy environment. Psychophysical studies have investigated the ability of human listeners to detect a signal embedded in masking noise, by presenting both stimuli via two loudspeakers. These loudspeakers used to have either the same or different positions in azimuthal space. A review of the cocktail party problem can be found in Bronkhorst (2000). Summarized, these studies showed that the detection threshold for spatially separated sounds is about 6–10 dB lower than the detection threshold for a signal co-located with the masker source.

Binaural signal detection in the presence of a masker has also been extensively studied under head-

Correspondence to: Andrea Lingner · Division of Neurobiology, Department Biology II · Ludwig-Maximilians-Universitaet Munich · Großhaderner Str. 2-4, 82152 Martinsried-Planegg, Germany. Telephone: +49-89-218074369; fax: +49-89-218074304; email: lingner@zi.biologie.uni-muenchen.de

phones. Here, Licklider (1948) and Hirsh (1948) were the first to show that phase differences between the signals at the two ears can lead to a detection improvement compared to a situation where the signal at the two ears is in phase. This phenomenon was later termed binaural masking level difference (BMLD). In humans, BMLDs due to the phase variation of either signal or masker can be as big as 9–15 dB.

Both spatial release from masking and BMLDs have been demonstrated in non-human mammals and birds (Wakeford and Robinson 1974; Hine et al. 1994; Dent et al. 1997; Ison and Agrawal 1998).

However, signal detection in noise is only one part of the binaural signal processing. In order to react appropriately to a sound (i.e., an enemy), the correct localization of a target sound is important. Several studies have shown decreased localization ability of human subjects at very low signal/noise ratios (SNRs) (Good and Gilkey 1996; Lorenzi et al. 1999; Braasch and Hartung 2002). The same is true for the localization of a target alone presented at low intensity levels (Inoue 2001).

Because of their well-developed low-frequency hearing (Ryan 1976) and the possibility of conducting physiological and behavioral approaches, gerbils are a commonly used animal model (Brand et al. 2002; Carney et al. 2011). However, there have been no previous comparisons of the behavioral performance of humans and gerbils in an identical binaural experiment.

Here, we quantified the localization ability under free-field conditions of a low-pass filtered noise signal masked by spatially distributed noise maskers in both gerbils and humans. The empirical work is supplemented with a quantitative simulation to explain the behavioral performance differences between gerbils and humans.

MATERIALS AND METHODS

Gerbil psychoacoustics

Subjects

Experiments were performed with five male Mongolian gerbils (*Meriones unguiculatus*). To examine the effect of noise rearing on the localization ability (Kapfer et al. 2002), three of the five gerbils were reared in omni-directional white noise (see section on “Noise rearing” under “Materials and Methods”). The remaining two were reared under normal laboratory conditions, serving as a control group. Each animal group was housed in a separate 71×46×31 cm (length × width × height) cage, containing wooden chipping as bedding, a sleeping house, and paper

towels for nesting. Gerbils were kept under consistent laboratory conditions with a 12-h day/night rhythm, a temperature of 23°C, and a humidity of 55%. Gerbils were trained 5 days a week, followed by a break of 2 days. They had unrestricted access to water all the time. During training days, gerbils were food deprived, receiving 20 mg pellets (Dustless Precision Pellets Product number F0071; BIO-SERV; Frenchtown, NJ, USA) as a reward for correct decisions in the experiment. On days without training, they had unrestricted access to rodent dry food (ssniff Gerbil; ssniff Spezialdiäten GmbH; Soest, Germany). Body weight was monitored daily during the whole training period. In case of a weight loss, animals received additional food supply.

Noise rearing

Three out of the five subjects were exposed to omni-directional white noise between postnatal day (P) 10 and 25. Details for the noise rearing are given in Kapfer et al. (2002). Briefly, white noise was generated using two independent, analogue generators (Bruel and Kjaer; Germany) and presented over four speakers on each side of the noisebox. The noise level was approximately 80 dB sound pressure level (SPL). All experiments were approved according to the German Tierschutzgesetz (AZ 55.2-1-54-2531-58-5).

Apparatus

Behavioral experiments were performed in a double-walled sound attenuating chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Walls, ceiling, and floor were covered with foam wedges (Industrial Acoustics Company GmbH, Niederkrüchten, Germany). Foam wedges were 40 cm in depth, highly reducing echoes for frequencies of more than approximately 250 Hz. Training took place in a circular arena placed in the chamber. A schematic diagram of the setup is depicted in Figure 1A. The arena had a diameter of 94 cm, enclosed by a wire mesh with a height of 29 cm; the floor was covered by carpet. A platform (3 cm in height and a diameter of 9 cm), with a small ring arranged in front of it was placed in the center of the arena, serving as the starting position. Six custom-made arms were mounted on a rail around the arena. Each arm consisted of a loudspeaker (Aura Sound, NSW1-205-8A, Santa Fe, CA, USA), a foot-switch, and a feeder to deliver pellets for correct decisions. A calibration routine was used to assure a flat frequency and phase response from approximately 200 Hz to 10 kHz for each loudspeaker. Loudspeakers were positioned at $\pm 17.5^\circ$, $\pm 52.5^\circ$ and $\pm 87.5^\circ$ off midline. This resulted in a separation of 35°

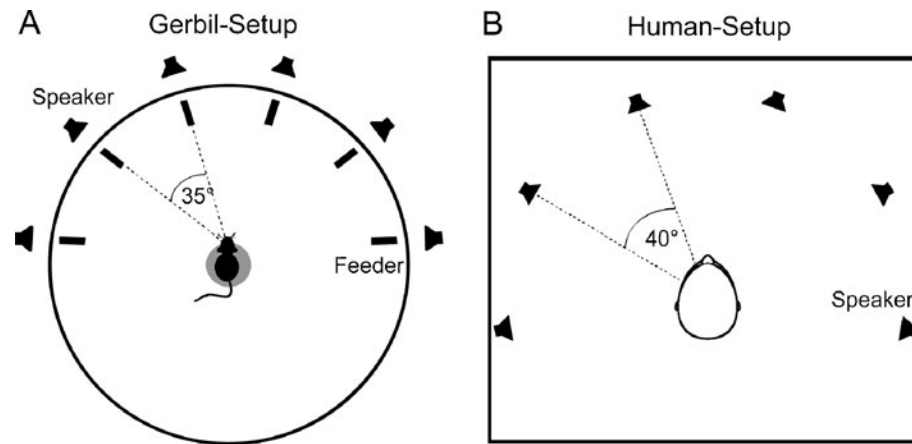


FIG. 1. Experimental setup. **A** Gerbil experimental setup. A platform, serving as starting position, with a little ring arranged in front of it is located in the middle of the setup. Six loudspeakers are positioned at $\pm 17.5^\circ$, $\pm 52.5^\circ$, and $\pm 87.5^\circ$ off midline around the setup, resulting in a separation of 35° for neighboring speakers. In front of each loudspeaker is a foot-switch and feeder delivering

pellets for correct decisions. **B** Human experimental setup. The setup consists of a centered chair with six surrounding loudspeakers positioned at $\pm 20^\circ$, $\pm 60^\circ$, and $\pm 100^\circ$ off midline, resulting in a separation of 40° for neighboring speakers. In the human experiment, correct and incorrect decisions were indicated visually via a touch screen (not shown here).

for neighboring speakers and a separation of 175° for the two outer speakers. The distance between each loudspeaker and the gerbil's head is approximately 55 cm. A video camera was installed directly above the setup to observe the gerbils' performance outside the chamber. Four halogen light spots provided setup lighting.

Stimuli

The stimuli used in this study consisted of six maskers and a signal. The maskers and signal were low-pass filtered white noise with a cut-off frequency of 1 kHz (slope, 24 dB/octave). The maskers were played continuously from all six loudspeakers during the whole training session, and the signal was added to the masker at a randomly chosen loudspeaker. The maskers at the six speakers were presented under two conditions, either the maskers at the six different speakers were correlated (that is, six identical noises) or they were uncorrelated (that is six independently generated noises). The sound pressure level of the maskers at the gerbils' starting position was approximately 60 dB SPL for the uncorrelated condition and 68 dB SPL for the correlated condition, respectively. The SNR ranged from 24 to -6 dB in 3 dB steps. The stimuli were played back through a Delta 410 PCI Audio card (M-Audio, Germany) at a sampling rate of 22.05 kHz and amplified (Rotel, RB-976 MKII) before being presented to the gerbils via the loudspeakers. Maskers were generated with a duration of 10 s and played in a loop to assure continuous playback throughout the whole training session. The signal was generated independently from the masker noise and had a duration of 300 ms including 10 ms raised-cosine ramps.

Experimental procedure

Starting at an age of approximately P40, gerbils were trained to localize a signal in a six-alternative forced-choice (6-AFC) paradigm by means of operant conditioning using food pellets as positive reinforcement. Gerbils were trained to jump on the platform and position their nose within the ring, ensuring a defined head position. This also prevents any head movements during the presentation of the signal. The outer diameter of the ring was 3 cm to ensure unrestricted sound transmission to the gerbils' ears. Breaking a light barrier in the ring started a trial by adding the signal to one of the six continuous maskers. Gerbils learned to move towards the signal speaker and to indicate their decision by pressing the foot-switch in front of that speaker. Gerbils had to respond within 15 s to complete a trial. Correct decisions within this time were rewarded with a food pellet; incorrect decisions resulted in no reward. Activation of a new trial within the 15-s time period without finishing the preceding trial (by pressing a foot-switch) was not possible. There was no timeout between completed trials, regardless of the correctness of the trial. An experimental session started with six trials at the highest SNR, three trials for each of the two masker conditions. The SNR was then reduced by 3 dB, and another six trials were obtained. This procedure was repeated until the lowest SNR was reached, and then the SNR was reset to its maximum. After starting the experimental session, the procedure ran automatically, controlled by custom software (MatLab, The Mathworks, Natick, MA, USA), until the experimenter stopped the training session. Experimental sessions took place

once a day between 1 and 5 pm and lasted on average between 30 and 60 min for each gerbil. Within this time, about 70–130 trials could be obtained per gerbil per day.

Data analysis

Data acquisition was completed after recording at least 60 trials per SNR and masker condition per gerbil. To determine a localization threshold (i.e., the lowest SNR at which the signal can still be correctly localized), we fitted a sigmoidal function to the recorded psychometric functions. Localization thresholds were determined as the point where this function crossed the significance threshold of 25%. This threshold was determined using a two-tailed binomial test and gives the percentage at which a localization performance differed significantly ($p < 0.05$) from chance level of 16.67% in a 6-AFC when the performance is based on 60 trials per SNR.

Human psychoacoustics

Subjects

Four human listeners (three male and one female, aged 26, 27, 31, and 32) participated in this study. These included the first author and three other listeners. All subjects had experience with psychoacoustic tasks. Listeners participated voluntarily in this study and showed normal hearing at audiometric frequencies between 250 and 8,000 Hz.

Apparatus

Ensuring comparable results between human and gerbil psychophysics, the loudspeakers of the human psychophysical setup were arranged similarly to gerbil experimental ones. The subjects were seated in a double-walled sound attenuated chamber (Industrial Acoustics Company GmbH, Niederkrüchten, Germany) surrounded by a semi-circular loudspeaker array. Six loudspeakers (Canton XS.2, Weilrod, Germany) were positioned at $\pm 20^\circ$, $\pm 60^\circ$, and $\pm 100^\circ$ off midline. This resulted in a separation of 40° for neighboring speakers and a separation of 200° for the two outer speakers. Similar to the gerbils, a calibration routine was used to assure a flat frequency and phase response from approximately 200 Hz to 10 kHz for each loudspeaker. Listeners were seated exactly in the middle of the loudspeaker array with their head fixed to prevent any head movements. The distance between each loudspeaker and the humans' head was approximately 97 cm. Walls, ceiling, and floor were covered with 20 cm foam wedges. Figure 1B depicts the human experimental setup.

Stimuli

Two versions of the human psychophysics were carried out in this study. In the first version, we presented the same stimuli as used for the gerbil psychophysics. As with the gerbils, sound pressure level was set to 60 and 68 dB SPL for the uncorrelated and correlated masker condition, respectively. The SNRs ranged from 9 to -15 dB in 3 dB steps. The second version of the experiment was identical to the first version, except that the low-pass cut-off frequency for both maskers and signal was reduced to 500 Hz. The lowered cut-off frequency was used to account for the difference in the head size/wavelength ratio of gerbils and humans in the first set of experiments.

Consider a fixed wavelength but two subjects with different head sizes, i.e., gerbils with a small head and humans with a big head. To create a head size/wavelength ratio for humans that is more similar to that of gerbils, one has to expand the wavelength, thus reducing the cut-off frequency. The masker sound pressure levels for the reduced cut-off frequency were the same as for the higher cut-off frequency. Stimuli were generated in Matlab and digital analog converted with an Motu PCI424 board and Motu HD192 converters (Cambridge, MA, USA). Stimuli were then amplified (Rotel CI9120, Halle, Germany) and presented via the loudspeakers.

Experimental procedure

The experimental procedure was identical to that of the gerbils apart from the following differences. While the continuous maskers were already active, listeners triggered each trial and the presentation of a signal by pressing a button on a graphical user interface displayed on a touch screen. Listeners were required to indicate from which of the six speakers the signal was presented by pressing one of six buttons arranged in a semi-circle on the touch screen. Visual feedback was provided after every trial. As in the gerbil experiments, psychometric functions for the two masker conditions were obtained using a non-adaptive one-interval, 6-AFC procedure: An experimental run was started by presenting the signal with the highest SNR (9 dB). At this SNR, six trials were obtained for the uncorrelated-masker condition followed by six trials for the correlated-masker condition. The SNR was then reduced by 3 dB and another 12 trials were obtained. This sequence was repeated until a minimum SNR of -15 dB was reached, which finished one run. Listeners were free to decide how many runs they performed in an experimental session. A minimum of ten runs were acquired for each listener, yielding at least 60 trials per point on the psychometric functions

for each of the two masker conditions. Data analysis was identical to that of the gerbils.

Simulation

To develop an understanding for the differences in masked thresholds between gerbils and humans, we simulated the behavioral experiments in a numerical model of binaural processing. The model is divided into three stages: I, the peripheral processing; II, the binaural processing; and III, the decision device. A block diagram of the different stages of the model is shown in Figure 2.

Peripheral processing

For each ear of the subject, the sounds from the loudspeakers were added after the corresponding interaural time differences (ITDs) were applied. These ITDs depended on both the azimuthal position of each speaker and the subject's head size. The head diameter was set to 3.2 cm for the gerbils and 18 cm for the

humans. The signals were then passed through a gammatone filterbank with five center frequencies equally spaced on a log frequency axis between 250 and 1,000 Hz, simulating the apical (low-frequency) part of the basilar membrane. The model assumed different auditory-filter bandwidths for gerbils and humans: Both auditory nerve and auditory-brainstem recordings from the gerbil have shown that low-frequency channels are broadly tuned (Faulstich and Kossel 1999; Siveke et al. 2008; Versteegh et al. 2011)). For the humans, the equivalent rectangular bandwidth (ERB) of the gammatone filter was set according to Glasberg and Moore (1990) as $ERB = 24.7 + 0.108 \times \text{center frequency}$, whereas ERB and center frequency are both given in Hertz. For the gerbils, the factor of 0.108 was replaced by 0.8, which results in ERBs of the gerbil being between three and six times the human ERBs. As a last step of the peripheral processing, the signal transduction at the inner hair cells was modeled as a half-wave rectifier, exponential compression with an exponent of 0.4 (Oxenham and Moore 1994) and a second order low-pass filter with a cut-off frequency of 1 kHz.

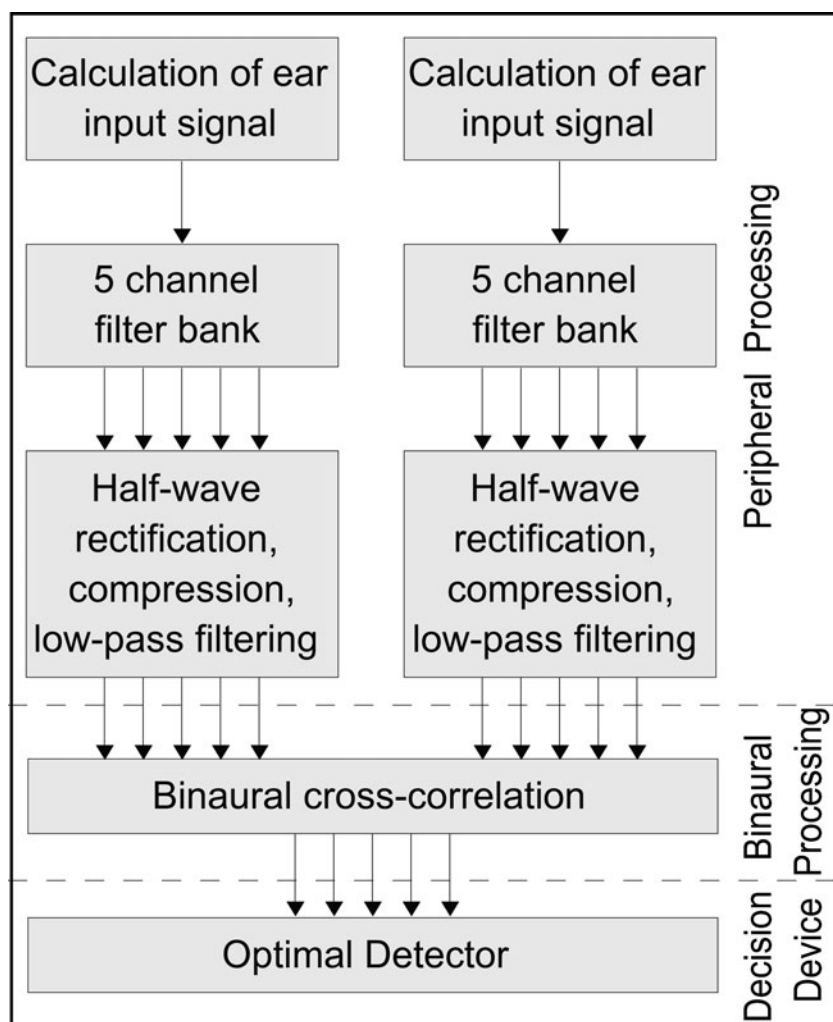


FIG. 2. Block diagram of the model. Auditory periphery is modeled by calculating the input signal for each ear, sending the signal through a five channel filterbank and performing a half-wave rectification, compression, and a low-pass filtering. Within the binaural processing, a cross-correlation of both monaural signals is performed. The resulting signal is then compared to internally stored templates within the decision device.

Binaural processing

The second step of the model simulates binaural interactions, by performing a cross-correlation between the left and right ear signal. In contrast to the model published in Bernstein and Trahiotis (1996), the cross-correlation in this model was not normalized. No internal noise was added after the binaural cross-correlation. The output of the binaural processor was excitation as a function of frequency and correlation lag axis (further interpreted as ITD); it is called the binaural display.

Decision device

A decision device was modeled using a so-called optimal detector as introduced by Dau et al. (1996): The decision device requires a masker and a signal template. The masker template is the binaural display (i.e., the unnormalized cross-correlation between the left and right ear signal) in response to either the uncorrelated or the correlated maskers. The signal template is the change in the binaural display affected by adding an above-threshold signal to the maskers at one of the six speakers. Following the simulation approach of Dau et al., the exact SNR for the template generation is not critical; only it must be well above threshold. As the perceptual thresholds for the gerbils were much higher than those for the humans, the SNRs for the template generation were set to 36 and 12 dB for the gerbils and humans, respectively. For each experimental condition and each subject, the masker references and each of the six signal templates were averaged across 60 presentations of the stimuli. The masker template and the six different signal templates for each masker condition and each experimental subject are shown in Figure 6.

As in the psychophysical experiment, psychometric functions were generated by the model by presenting 180 repetitions of each stimulus with randomized signal

position for each subject (gerbil or human), for each masker condition, and each SNR. The range of SNRs was 36 to -6 dB for the gerbils and 24 to -18 dB for the humans. The SNR step size was 3 dB. The model calculates the binaural display for the signal plus masker and subtracts the masker template. Finally, it calculates the cross-correlation of the resulting representation with each of the six signal templates. The decision device chooses the signal position corresponding to the signal template where the cross-correlation is maximal. Data analysis and presentation is identical to that of the experimental data.

RESULTS

Sample psychometric functions for the signal localization are shown for one control and one noise-reared gerbil in Figure 3A. With the uncorrelated masker condition, the localization threshold for the control gerbil is at a SNR of 0.9 dB and for the noise-reared gerbil at a SNR of 0.1 dB. With the correlated masker condition, the thresholds increase to an SNR of 9.2 dB (control gerbil) and 9 dB (noise-reared gerbil). These data indicate that, while localization thresholds differ strongly between masker conditions, they are independent of whether the gerbil was reared in a noisy environment or not. A two-way ANOVA of the parameters “masker condition” and “noise rearing” reveals a significant effect of the masker condition ($p < 0.001$; F value, 41.87; and degrees of freedom, 1) but no effect of noise rearing ($p > 0.8$; F value, 0.07; and degrees of freedom, 1). Consequently, noise-reared and control gerbils are grouped together for the subsequent analyses.

A sample psychometric function for one human listener is shown in Figure 3B. The localization threshold for this listener is at an SNR of -3.8 and

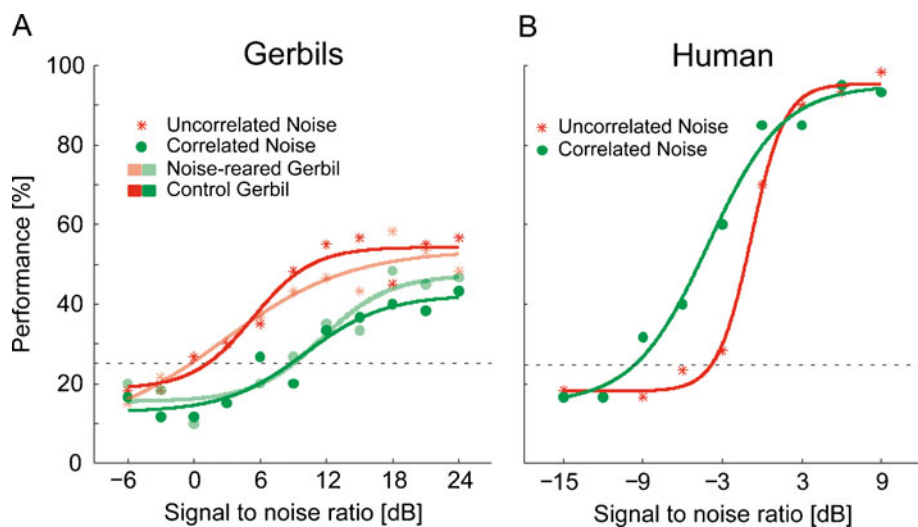


FIG. 3. Sample psychometric functions. **A** Sample psychometric function for one control gerbil (dark symbols) and one noise-reared (pale symbols). **B** Sample psychometric function for one human listener. Symbols indicate the localization performance in percent correct at each SNR for each subject. Red stars represent the localization performance in presence of the uncorrelated masker condition; green filled circles represent the localization performance in presence of the correlated masker condition. Green and red solid lines indicate the fitted sigmoidal function the corresponding data. The black dashed line depicts the significance level of 25% ($p < 0.05$, two-tailed binomial test in a 6-AFC with 60 trials per point).

−9.5 dB for the uncorrelated and correlated masker condition, respectively.

The mean localization thresholds for gerbils and humans are shown in Figure 4A. For the gerbils, an SNR of 0.9 and 8.5 dB was necessary to localize a signal in uncorrelated and correlated masker condition, respectively. For the humans, an SNR of −4.7 and −8.7 dB was necessary to localize a signal in uncorrelated and correlated masker condition, respectively. Thus, humans showed an overall localization performance that is markedly better than that of gerbils. Furthermore, assuming the uncorrelated masker condition as reference, gerbils showed a masking release of −7.6 dB, whereas humans showed a masking release of +4 dB (see Fig. 4B).

To understand the differences in the release from masking between gerbils and humans, we created a numerical model of binaural processing, simulating the behavioral experiments (see methods, simulations). The comparisons of the localization thresholds measured in the experiments and those derived from the simulation are shown in Figure 5A. Overall, the simulation results follow the experimental data in that for the gerbils, localization thresholds are generally higher than in humans. The effect of the masker condition, i.e., the release from masking, on the localization thresholds is qualitatively predicted by the model for both gerbils and humans. A direct comparison of the release from masking is shown in Figure 5B.

To clarify the features of the model that guided its detection sensitivity, the inspection of the masker and signal templates is informative: The plots of the

masker templates (highest row in Fig. 6) show that for the gerbils, the switch from correlated to uncorrelated maskers decreases the excitation in the binaural display. The shape of the excitation pattern, however, is relatively unchanged. For the humans, the switch from correlated to uncorrelated maskers changes not only the excitation level but also the shape of the excitation pattern: at higher center frequencies, side ridges occur (upper right panel in Fig. 6B).

The signal templates show how the addition of the signal at one of the speakers changes the excitation in the binaural display. Signal templates are shown in rows 2–7 of Figure 6. Due to the small head size, the relatively simple excitation pattern in the gerbil is only slightly shifted along the ITD axis when the signal speaker is moved from -87.5° to $+87.5^\circ$.

The larger head size in humans leads to a much stronger signal-induced changes in the binaural display: The more complex pattern of ridges and troughs is shifted by larger amounts along the ITD axis when the signal speaker is moved from -100° to $+100^\circ$. Thus, these signal-induced shifts in the binaural display are easier to detect, and consequently, the simulated human localization thresholds are lower than the simulated gerbil localization thresholds.

This analysis suggests that differences in the overall performance and the effect of masker condition on localization thresholds are due to differences in the input characteristics (head size, auditory filter bandwidth) of gerbils and humans.

As a control experiment to verify this hypothesis, humans were tested with a second set of stimuli using

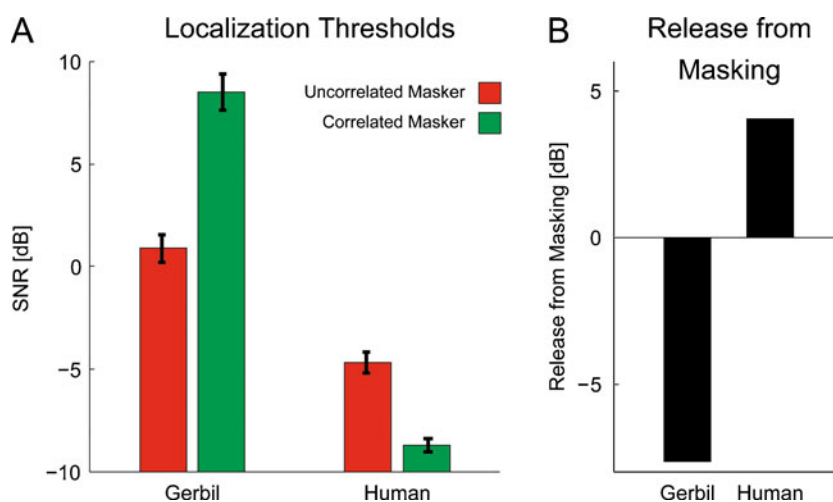


FIG. 4. Localization thresholds and the resultant release from masking. The mean localization thresholds for gerbils and humans are shown in **A**. The *red bars* represent mean localization thresholds in presence of the uncorrelated masker condition, whereas the *green bars* represent the mean localization thresholds in presence of the correlated masker condition. Concerning gerbils, SNRs of 0.9 and 8.5 dB were required to localize a signal in uncorrelated and

correlated masker condition, respectively. Concerning humans, SNRs of −4.7 and −8.7 dB were required to localize a signal in uncorrelated and correlated masker condition, respectively. *Error bars* depict the standard deviation. **B** Resultant release from masking for both gerbils and humans, assuming the uncorrelated masker condition as a reference. This result in a release from masking of −7.6 dB for gerbils and a release from masking of +4 dB for humans.

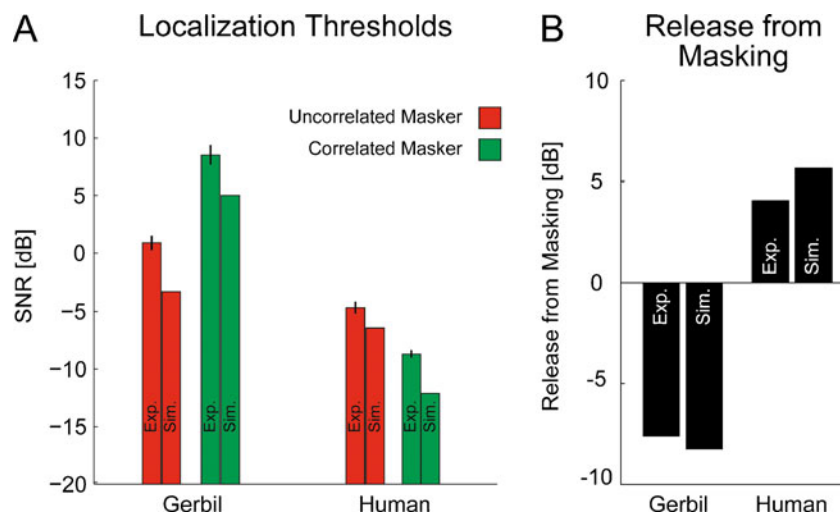


FIG. 5. Localization thresholds and resultant release from masking as a comparison to the model. Localization thresholds for the experimental (*Exp.*) and simulated data (*Sim.*) for gerbils and humans are shown in **A**. The *red bars* represent the mean localization thresholds in the presence of the uncorrelated masker condition, whereas the *green bars* represent the mean localization thresholds in presence of the correlated masker condition. **B** Resultant release

from masking calculated from the experimental and simulated data for both gerbils and humans, assuming the uncorrelated masker condition as a reference. Localization thresholds for both gerbils and humans are at least qualitatively correct predicted by the model. Thus, the model correctly predicts a negative release from masking for the gerbils and a positive release from masking for the humans (**B**).

a reduced low-pass cut-off frequency of 500 Hz to test whether the ratio of head size and wavelength is critical in the current experiments. The effect of lowering the cut-off frequency and a direct comparison of the simulated release from masking for both cut-off frequencies is shown in Figure 7. Again assuming the uncorrelated masker condition as a reference, the experimentally determined release from masking decreased from about 4 dB at a cut-off frequency of 1,000 Hz to about 2 dB at a cut-off frequency of 500 Hz. Similar to the experimental data, the model also followed this decrease in the release from masking as an effect of lowered cut-off frequency.

DISCUSSION

The localization thresholds of a low-pass filtered noise signal in the presence of correlated and uncorrelated spatially distributed maskers were examined in gerbils and humans with the same free-field paradigm. The experimental data show that, overall, gerbils need a much higher SNR to localize the low-frequency signal than humans. When the masker is switched from uncorrelated noises across the six speakers to correlated noises, the localization thresholds change for gerbils and humans but in opposite directions: While the gerbils show more masking (SNR at localization threshold increases by 7.6 dB), the humans show a 4-dB release from masking.

These data are qualitatively predicted by a numerical model of binaural processing. The inspection of

the model, particularly the inspection of the binaural display of the masker and signal templates, shows that differences in the psychophysical performance between gerbils and humans can be explained qualitatively with the differences of the inputs to the binaural processor: The smaller head size and the broader auditory filters of the gerbil lead to deterioration of the salience of signal-related features in the binaural display.

Further runs of the simulation (data not shown) indicated that head size and auditory-filter bandwidth contributed in different ways to the observed pattern of localization thresholds. Specifically, head size is mainly responsible for the observed difference in the overall performance, whereas both head size and auditory filter bandwidth account for the differential effect of the masker conditions. This is in line with the inspection of the binaural display for the signal template: It shows that the larger head of humans' results in much stronger signal-induced changes in the binaural display making the localization of a signal already at high SNRs easier. Moreover, the complex pattern of ridges, produced by the auditory filters, facilitates the localization ability of humans even at low SNRs.

The benefit of binaural hearing for signal detection has been assessed both with psychophysical studies under headphones (BMLDs) and with studies in the free sound field (spatial unmasking). In the following, the current results are discussed with respect to these two approaches performed in animal models of human hearing.

BMLDs measured under headphones in animals are, due to the difficult experimental approach, rare.

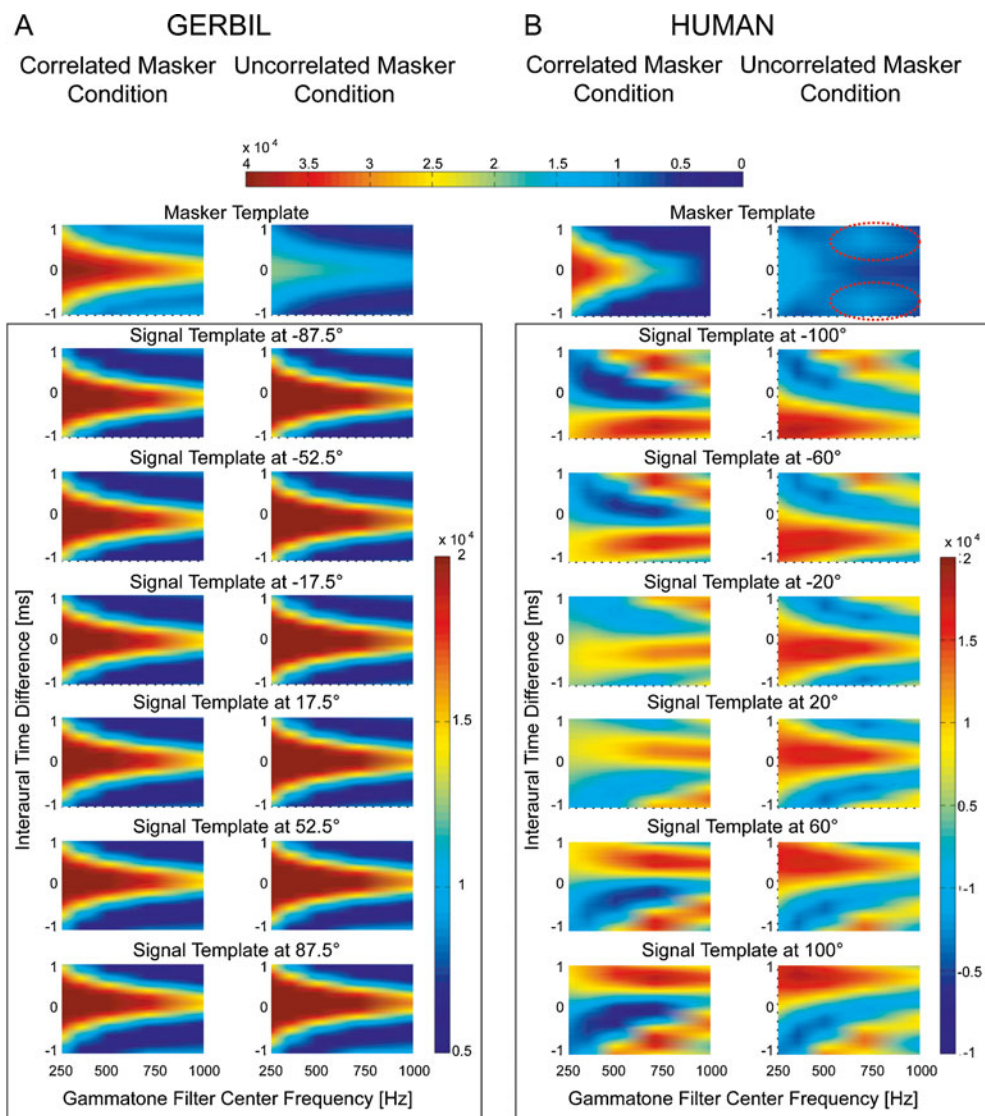


FIG. 6. Simulated masker and signal templates. This figure shows the binaural excitation pattern of the simulated masker and signal templates as function of gammatone filter center frequency and interaural time difference for both masker conditions for gerbils (A) and humans (B). Upper row of each panel depicts the masker template; dotted circles indicate side ridges that occur for humans at higher frequencies (for more information, see text). Rows 2–7 of each

panel depict the signal template for each corresponding loudspeaker separation. The upper color bar depicts the excitation for the masker template. The two lower color bars depict the signal induced changes of excitation for the simulated gerbil (left) and human (right). Both excitation and change of excitation are given in arbitrary model units. For further explanations of the differences in the binaural excitation pattern, see text.

Cats and rabbits exhibit BMLDs of approximately 8 dB for a 500-Hz signal (Wakeford and Robinson 1974; Early et al. 2001), similar to those of humans. Even though these data were recorded with low-frequency stimuli and the subjects had rather small heads as compared to humans, the occurrence of pronounced BMLDs is not surprising; under headphones, differences in level and arrival time at the two ears (ILD and ITD cues, respectively) can be independently manipulated. Inverting the phase of a 500-Hz signal corresponds to a change of ITD by 1 ms, which is much more than most animal models would ever encounter in the free sound field.

In contrast to head-phone experiments, any manipulation of the stimulus presented in the free-field leads to changes of the overall sound field. Spatial unmasking in the free sound field has been investigated in various animal models: Ferrets show markedly lower thresholds for a 500-Hz tone signal presented with a signal-masker separation of 180° (bilateral) compared to co-located signal and masker (Hine et al. 1994). Compared to the current results, the observation of low-frequency spatial unmasking in the ferret is surprising at first sight: The head size of ferrets is not much larger than that of gerbils, and thus, neither is the physiological range of ITDs. However,

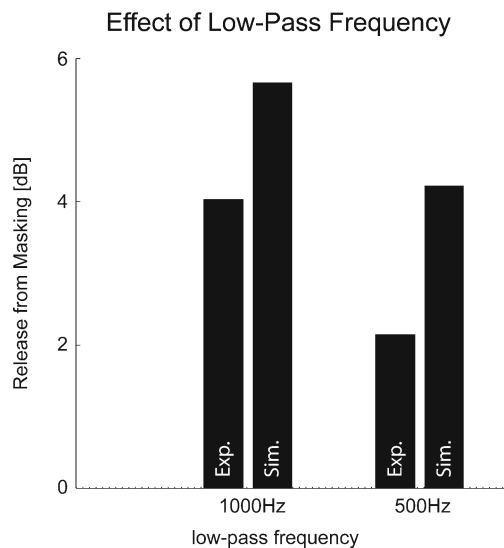


FIG. 7. Effect of low-pass frequency. The comparison of the release from masking determined from the experimental (*Exp.*) and simulated data (*Sim.*) for humans is shown as a function of the two different low-pass cut-off frequencies. A lowering of the cut-off frequency resulted in a smaller release from masking, both for experimental and simulated data.

a more detailed inspection of the two studies reveals an important procedural difference: In the Hine et al. paradigm, the animals had to detect the target that was always emitted from the same position in space. Thus, the paradigm of Hine et al. is a signal-detection paradigm. In the current study, the subjects had to localize the signal emitted from one of six different speaker positions in space. Thus, the principal difference of detection vs. localization may cause the different results. Another possible reason for the discrepancies of the two studies may lie in the occurrence of low-frequency ILD effects (see below).

In budgerigars, Dent et al. (1997) showed that the detection threshold of low-frequency sounds decreased when masker and signal were separated in space. Based on the small head of budgerigars (comparable to that of gerbils), these results are also surprising. However, in line with current arguments, one explanation for the deviating results could again be the procedural difference (detection vs. localization) of the two experiments. Another possibility that can account for the results of Dent et al. is an anatomical difference between birds and mammals. In contrast to mammals, most birds exhibit middle ears that are acoustically connected. Such a middle ear connection is shown to most likely be responsible for an enlargement of ITDs (Larsen et al. 2006) and thus could explain the low-frequency spatial unmasking in budgerigars despite the fact that they have small heads.

Even in two species of water living mammals, harbor seals and California sea lions, a benefit from

spatial separation of signal and masker was shown (Holt and Schusterman 2007). Changing the position of aurally presented pure tones (1, 8, or 16 kHz) from being co-located with the masker at 0° to either 45° or 90° induced a detection threshold decrease similar to that found in humans (Blauert 1997). The head radii of both pinnipeds and humans are virtually identical (Holt et al. 2004). Thus, not only the ITD cues at 1 kHz but also ILD cues occurring at the higher pure-tone frequencies are qualitatively similar in pinnipeds and humans, and they support the hypothesis that pinnipeds process binaural information as effectively humans.

The current paradigm differs in important points from both BMLD experiments and those on spatial unmasking: One reason for using the current paradigm has been that BMLDs are measured under headphones, which is very difficult to do with gerbils. An advantage of the current paradigm over the spatial-unmasking paradigms cited above is that there are no apparent “better ear” effects in terms of ILDs (see also below) because the spatial arrangement of maskers is symmetric around the head and the signal can occur at any azimuth within the range covered by the maskers.

However, there are several challenges to be considered in the interpretation of free-field studies: In the current experiment, the change from the uncorrelated- to the correlated-masker condition leads to a sound level increase from 60 to 68 dB SPL at the position of the subject’s head. This is the case because with uncorrelated noises, a doubling of the number of speakers emitting the noises leads to a 3-dB increase in sound level, whereas with correlated masker condition, the increase is 6 dB per doubling of the number of speakers. Thus, irrespective of the correlation degree of the noise at the six speakers, lower localization thresholds are expected for the uncorrelated masker condition. Indeed, this was found for the gerbils.

As opposed to BMLD experiments, the current free-field paradigm generates only physiologically plausible ITDs. However, ILDs can occur, at least for the human listeners: Several studies have shown that also at frequencies below 1 kHz, ILDs of up to 5 dB are generated at sound-source distances around 1 m (Brungart and Rabinowitz 1999; Kuwada et al. 2010). Even though both cues were shown to be present in low-frequency sounds, a recent study of Macpherson and Middlebrooks (2002) showed that listeners judged ITDs as the prominent cue to determine the azimuthal position of a low-frequency signal. Therefore, we did not implement ILDs in our simulation. Similar to the Macpherson and Middlebrooks study, the success of our model in explaining both the human and gerbil results indicates that ITD cues

alone are sufficient to explain the psychophysical performance.

For the localization with low-frequency spectral cues, the magnitude and phase response of all six speakers in our setups is identical (referenced to an omnidirectional microphone at the subject's head position). At the subject's ear drum, however, spectral localization cues can occur with the correlated-noise maskers. For humans, the masker from a far-left speaker takes about 0.5 ms longer to the right ear than the masker from the far-right speaker. This will result in a comb filter effect with an f_0 of 2 kHz, i.e., the first spectral notch would occur at 1 kHz. If the signal (which is not correlated to any of the masker noises) is added to a lateral positioned loudspeaker, it reduces the correlation of the waveforms at the two ears. Thus, presentation of the signal will not only introduce ITD cues but also reduce the correlation and decrease the depth of the spectral notch at 1 kHz, i.e., ILD cues are potentially introduced. Note, however, that these effects would be fully captured by the simulations and the decision device, which looks at the (unnormalized) interaural correlation as a function of ITD and frequency. Thus, the decision device could pick up these spectral cues possibly associated with the detection of a lateral signal masked by correlated noises in humans.

In summary, the current paradigm cannot isolate ITD cues from low-frequency ILD cues as well as for the classical BMLD paradigm, implemented under headphones. Nevertheless, the current simulations indicate that low-frequency ILDs are not required to quantitatively predict the performance of humans and gerbils.

Overall, the current empirical data from gerbils and humans in an identical free-field unmasking paradigm show that gerbils needed higher SNRs to localize the signal than humans. Moreover, when the maskers were switched from being uncorrelated across the six masking speakers to being correlated, the gerbils' performance decreased further while the humans' performance improved. The simulations show that these dramatic performance differences can be fully accounted for by the fact that, due to the smaller head and wider auditory filters in gerbils, the inputs to the binaural processor are less salient for the gerbils than for the humans. This indicates that the binaural processor itself is equally sensitive in gerbils and humans. However, the physical limitations imposed by the small head prevent the gerbil from performing equally well in the current paradigm. Thus, the current data, together with earlier experiments on the precedence effect (Wolf et al. 2010) argue against the notion that the gerbil may not be a good animal model of human binaural processing due to its relatively large minimum resolvable angle

(Heffner and Heffner 1988; Carney et al. 2011). The current data and simulations highlight the need to evaluate in detail differences in the experimental paradigms and threshold criteria to judge the quality of an animal model.

ACKNOWLEDGEMENTS

This work was funded by a grant from the "Deutsche Forschungsgemeinschaft" (GR 1205/14-1 8419087) to B.G. We thank Teresa Kindermann for the initial training of the animals.

REFERENCES

- BERNSTEIN LR, TRAHOTIS C (1996) On the use of the normalized correlation as an index of interaural envelope correlation. *J Acoust Soc Am* 100:1754–1763
- BLAUERT J (1997) Spatial hearing: the psychophysics of human sound localization. MIT Press, Cambridge
- BRAASCH J, HARTUNG K (2002) Localization in presence of a distractor and reverberation in the frontal horizontal plane. *I Psychoacoustical Data Acustica* 88:942–955
- BRAND A, BEHREND O, MARQUARDT T, McALPINE D, GROTHE B (2002) Precise inhibition is essential for microsecond interaural time difference coding. *Nature* 417:543–547
- BRONKHORST AW (2000) The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acustica* 86:117–128
- BRUNGART DS, RABINOWITZ WM (1999) Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am* 106:1465–1479
- CARNEY LH, SARKAR S, ABRAMS KS, IDROBO F (2011) Sound-localization ability of the Mongolian gerbil (*Meriones unguiculatus*) in a task with a simplified response map. *Hear Res* 275:89–95
- CHERRY EC (1953) Some experiments on the recognition of speech, with one and with two ears. *J Acoust Soc Am* 25:975–979
- DAU T, PUSCHEL D, KOHLRAUSCH A (1996) A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am* 99:3615–3622
- DENT ML, LARSEN ON, DOOLING RJ (1997) Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). *Behav Neurosci* 111:590–598
- EARLY SJ, MASON CR, ZHENG L, EVILSIZER M, IDROBO F, HARRISON JM, CARNEY LH (2001) Studies of binaural detection in the rabbit (*Oryctolagus cuniculus*) with Pavlovian conditioning. *Behav Neurosci* 115:650–660
- FAULSTICH M, KOSSL M (1999) Neuronal response to cochlear distortion products in the anteroventral cochlear nucleus of the gerbil. *J Acoust Soc Am* 105:491–502
- GLASBERG BR, MOORE BC (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138
- GOOD MD, GILKEY RH (1996) Sound localization in noise: the effect of signal-to-noise ratio. *J Acoust Soc Am* 99:1108–1117
- HEFFNER RS, HEFFNER HE (1988) Sound localization and use of binaural cues by the gerbil (*Meriones unguiculatus*). *Behav Neurosci* 102:422–428
- HINE JE, MARTIN RL, MOORE DR (1994) Free-field binaural unmasking in ferrets. *Behav Neurosci* 108:196–205
- HIRSH IJ (1948) Binaural summation and interaural inhibition as a function of the level of masking noise. *Am J Psychol* 61:205–213

- HOLT MM, SCHUSTERMAN R (2007) Spatial release from masking of aerial tones in pinnipeds. *J Acoust Soc Am* 121:1219–1225
- HOLT MM, SCHUSTERMAN RJ, SOUTHALL BL, KASTAK D (2004) Localization of aerial broadband noise by pinnipeds. *J Acoust Soc Am* 115:2339–2345
- INOUE J (2001) Effects of stimulus intensity on sound localization in the horizontal and upper-hemispheric median plane. *J UOEH* 23:127–138
- ISON JR, AGRAWAL P (1998) The effect of spatial separation of signal and noise on masking in the free field as a function of signal frequency and age in the mouse. *J Acoust Soc Am* 104:1689–1695
- KAPFER C, SEIDL AH, SCHWEIZER H, GROTHE B (2002) Experience-dependent refinement of inhibitory inputs to auditory coincidence-detector neurons. *Nat Neurosci* 5:247–253
- KUWADA CA, BISHOP B, KUWADA S, KIM DO (2010) Acoustic recordings in human ear canals to sounds at different locations. *Otolaryngol Head Neck Surg* 142:615–617
- LARSEN ON, DOOLING RJ, MICHELSEN A (2006) The role of pressure difference reception in the directional hearing of budgerigars (*Melopsittacus undulatus*). *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 192:1063–1072
- LICKLIDER JCR (1948) The influence of interaural phase relations upon masking of speech by white noise. *J Acoust Soc Am* 20:150–159
- LORENZI C, GATEHOUSE S, LEVER C (1999) Sound localization in noise in normal-hearing listeners. *J Acoust Soc Am* 105:1810–1820
- MACPHERSON EA, MIDDLEBROOKS JC (2002) Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am* 111:2219–2236
- OXENHAM AJ, MOORE BC (1994) Modeling the additivity of non-simultaneous masking. *Hear Res* 80:105–118
- RYAN A (1976) Hearing sensitivity of the mongolian gerbil, *Meriones unguiculatus*. *J Acoust Soc Am* 59:1222–1226
- SIVEKE I, EWERT SD, GROTHE B, WIEGREBE L (2008) Psychophysical and physiological evidence for fast binaural processing. *J Neurosci* 28:2043–2052
- VERSTEEGH CP, MEENDERINK SW, VAN DER HEIDEN M (2011) Response characteristics in the apex of the gerbil cochlea studied through auditory nerve recordings. *J Assoc Res Otolaryngol* 12:301–316
- WAKEFORD OS, ROBINSON D (1974) Detection of binaurally masked tones by the cat. *J Acoust Soc Am* 56:952–956
- WOLF M, SCHUCHMANN M, WIEGREBE L (2010) Localization dominance and the effect of frequency in the Mongolian Gerbil, *Meriones unguiculatus*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 196:463–470