

Published in final edited form as:

Hear Res. 2009 October ; 256(1-2): 64–74. doi:10.1016/j.heares.2009.06.014.

## Primate Auditory Recognition Memory Performance Varies With Sound Type

Ng Chi-Wing, Plakke Bethany, and Poremba Amy

Department of Psychology, Behavioral and Cognitive Neuroscience Program, University of Iowa, USA.

### Abstract

Neural correlates of auditory processing, including for species-specific vocalizations that convey biological and ethological significance (e.g. social status, kinship, environment), have been identified in a wide variety of areas including the temporal and frontal cortices. However, few studies elucidate how non-human primates interact with these vocalization signals when they are challenged by tasks requiring auditory discrimination, recognition, and/or memory. The present study employs a delayed matching-to-sample task with auditory stimuli to examine auditory memory performance of rhesus macaques (*Macaca mulatta*), wherein two sounds are determined to be the same or different. Rhesus macaques seem to have relatively poor short-term memory with auditory stimuli, and we examine if particular sound types are more favorable for memory performance. Experiment 1 suggests memory performance with vocalization sound types (particularly monkey), are significantly better than when using non-vocalization sound types, and male monkeys outperform female monkeys overall. Experiment 2, controlling for number of sound exemplars and presentation pairings across types, replicates Experiment 1, demonstrating better performance or decreased response latencies, depending on trial type, to species-specific monkey vocalizations. The findings cannot be explained by acoustic differences between monkey vocalizations and the other sound types, suggesting the biological, and/or ethological meaning of these sounds are more effective for auditory memory.

### Keywords

*Macaca mulatta*; delayed matching-to-sample; vocalizations; performance; same; different; rhesus; monkey

---

Monkeys have difficulty in learning a delayed matching-to-sample (DMTS) task requiring decisions about whether sounds match or not across memory delays (D'Amato and Colombo, 1985; Wright, 1998, 1999; Fritz et al., 2005). Rhesus monkeys generally learn the rule for visual and tactile versions of this trial-unique delayed matching- and nonmatching-to-sample at short delays, within a few hundred trials (Murray and Mishkin, 1998; Buffalo et al., 1999; Zola et al., 2000), while a similar task, using auditory stimuli, takes them on average 15,000 trials to learn the rule at 5-second memory delays (Fritz et al., 2005). Auditory memory performance seems rather poor compared to using visual and tactile stimuli in similar tasks.

---

© 2009 Elsevier B.V. All rights reserved.

**Corresponding author and postal address for all authors:** Amy Poremba, Ph.D., University of Iowa, E11 Seashore Hall, Department of Psychology, Iowa City, IA 52242, USA, amy-poremba@uiowa.edu, Office: 319-335-0372; Fax: 319-335-0191.

**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.

Monkeys show forgetting thresholds (i.e. scores falling to 75% accuracy) for visual and tactile stimuli at delays of 10 minutes or more, but thresholds for forgetting auditory stimuli are as short as 35 seconds. They require more training in discriminating auditory stimuli and are less efficient in maintaining auditory information for retention, compared to visual and tactile information, although it may be possible that experimenters have not yet devised the most robust way to test the auditory memory of non-human primates. A related finding similarly reports that human auditory recognition memory is relatively poor compared to visual recognition memory (Cohen et al., 2009). Here we investigate whether the auditory memory of monkeys is improved by, or if its expression is dependent on, particular sound types.

Species-specific vocalizations are salient stimuli to living organisms, for communication among individual members and about the surrounding environment (Fitch, 2000; Ghazanfar and Hauser, 2001). Imaging and neurophysiological studies identify “voice-sensitive” and “vocalization-sensitive” areas of secondary auditory regions, superior temporal gyri, temporal pole, insular cortex and prefrontal cortices in humans (Belin et al., 2000; Fecteau et al., 2004; Belin, 2006; Bélizaire et al., 2007) and non-human primates (Tian et al., 2001; Gil-da-Costa et al., 2004; Poremba et al., 2004; Romanski et al., 2005; Cohen et al., 2007, Petkov et al., 2008; Remedios et al., 2009). Similar neural correlates are also present in the second auditory cortical fields of birds (Theunissen and Shaevitz, 2006) and mice (Ehret, 1987; Geissler and Ehret, 2004). Like humans, non-human primates attend to distinctive acoustic cues of conspecific vocalizations for efficient auditory processing, compared to heterospecific vocalizations from non-rhesus monkeys or other animal species (Zoloth et al., 1979; Petersen et al., 1984; Hauser, 1998; Gifford et al., 2003; Rendall, 2003; Hienz et al., 2004; Fitch and Fritz, 2006). One possible explanation for differences in memory performance across stimulus types is that some sounds may be more readily processed or encoded by the brain. In humans, visual perception and memory performance are enhanced using faces, pictures, and words, which are more efficiently processed and categorized during human cognition (Seifert, 1997; Amrhein et al., 2002; Bulthoff and Newell, 2006). Species-specific vocalizations may then exert functional advantages in auditory learning and memory of monkeys over other sound types.

The present study aims to investigate if the memory performance of rhesus macaques varies across seven distinct sound types. Rhesus monkeys were tested with an auditory version of the delayed matching-to-sample (DMTS) task. They were trained to perform go/no-go responses for matching and nonmatching sounds respectively at fixed 5-second memory delays. In Experiment 1, a collection of approximately 900 auditory stimuli were used and classified based on acoustical, biological, and ethological characteristics. These sound groupings were then used for analyses of memory performance across match and nonmatch trials respectively. In Experiment 2, the total number of sound stimuli per sound type and the exact pairings of sound presentation were controlled and organized to achieve a trial-unique DMTS task to determine if particular types of sound stimuli would evoke better behavioral performance. The study hypothesized that monkey vocalizations, species-specific sounds to the animal subjects, would yield better memory performance than others in the task.

## Methods

### Experiment 1

**Subjects**—Six rhesus macaques (*Macaca mulatta*) were used, three males and three females between 11 and 12 years of age and weighing 6–11 kilograms. For approximately the first two years they were raised with other rhesus monkeys in a breeding facility in both indoor and outdoor corals. Since then, the monkeys have been in single housing or paired housing in animal colony rooms with a total room number of 7–23 other monkeys. During the testing included herein they were individually housed with a 12-hour light/dark cycle at the University of Iowa.

Food control was applied during behavioral training, in order to maintain them at 85% or more of their original weights. Monkey biscuits (Harlan Teklad, Madison, WI) were fed to animals daily, with fruits, vegetables, and treats scheduled throughout the week. All animals had access to water ad libitum. Treatment of the animals and experimental procedures were in accordance with the National Institutes of Health Guidelines and were approved by the University of Iowa Animal Care and Use Committee.

**Apparatus**—The auditory DMTS task took place inside a sound-attenuated chamber. Each animal was trained to sit in a primate chair and listened to a wide range of sound stimuli. The behavioral panel contained a speaker, an acrylic touch-sensitive button and a reward dish (Fig. 1). The speaker (3.5 inch × 3.5 inch) was 15 centimeters (cm) in front of the primate, at its eye level. The touch-sensitive button (2.8 inch × 2.8 inch) was 3 cm below the speaker to detect responses. The reward dish, 3 cm below the touch-sensitive button, released a food reinforcer from a pellet dispenser (Med Associates Inc, VT) for correct responses. A house light (a 40W light bulb) provided illumination throughout the training session. A library of 893 distinct sounds (containing significant spectral energy up to 10,000 Hertz) was presented at 70–75 decibels (dB) at sound pressure level (SPL), and each sound clip was truncated at 500 milliseconds (ms). LabView software (National Instruments, Austin, TX) controlled the lights, sound stimuli, pellet dispenser, and recorded button-pressing responses.

**General Procedures**—Training sessions were held five days a week and 50 trials were presented per session. The current setup employed go/no-go response rules for the auditory DMTS task (Fig. 1). The ratio of match to nonmatch trials was 1, randomly controlled by the LabView software. On match trials, the two sounds were the same and a correct go-response was made by touching the button resulting in the delivery of a small chocolate candy reward. On nonmatch trials, the two presented sounds were different and a correct response was sorted if the monkey avoided touching the button (i.e. a no-go response), which did not result in food delivery. Thus the current DMTS task, employing go/no-go rules, used an asymmetrical reinforcement contingency. In the two-alternative forced choice contingency, used in some other auditory primate studies (Wright, 1998,1999;Fritz et al., 2005), behavioral responses are always necessary for nonmatch trials. However, monkeys had difficulty in acquiring discrete button-pressing on match and nonmatch trials respectively. In order to learn the two-alternative forced choice contingency, responses for match and nonmatch trials needed to be spatially separated. Although the go/no-go setup does not require two separate behavioral responses, the monkeys learn this task faster, and the potential confound of spatial preference and/or processing was minimized while the goal was to elucidate auditory memory performance of monkeys in a non-spatial behavioral task.

In each trial, the memory delay between two sounds (i.e. inter-stimulus intervals) was five seconds long. The inter-trial interval (ITI) was set at 12 sec, and premature response during the ITIs reset the interval. The same response during 5-second memory delays reset that trial. There were no more than three consecutive trials of match or nonmatch trials in a row. Monkeys were trained to a criterion of 80% or better on match and nonmatch performance combined. All sounds (893 samples) were divided into 18 sound folders (50 unique sound stimuli each on average), and folder use was cycled across days. Two of the 18 sound folders were pre-selected for each session/monkey. The order and combination of 18 folders were randomized weekly, and thus a given stimulus was repeated once on average every 10 training days.

**Auditory stimuli**—Acoustic samples, 884 out of 893, were classified by two independent human researchers into *post-hoc* groupings that yielded seven sound types: animal vocalizations (Anivoc), human vocalizations (Hvoc), monkey vocalizations (Mvoc), music clips (Music), natural sounds (Nature), synthesized clips (Syn) and band-passed white noises (WhiteN). Animal vocalizations (Anivoc), 123 out of the 884 samples (13.9 %), included

vocalizations recorded from birds, domestic animals (e.g. cat, and dog etc.), and miscellaneous/wild animals (e.g. lion, elephant, and leopard etc.). Human vocalizations (Hvoc), 113 samples (12.8 %), included speech sounds (e.g. “girl”, “thank you” and “good morning” etc.) and non-speech sounds (e.g. laughing, crying and sneezing etc.) generated from unknown male and female speakers. Monkey vocalizations (Mvoc), 14 samples (1.6 %), included various vocalizations generated by unknown rhesus monkeys. Music clips (Music: 142 samples or 16.1 %) contained notes (e.g. harmonics), and sound clips (e.g. extracts of orchestra symphonies and melodies of TV commercials) generated from various musical instruments (e.g. violin, flute and trumpet etc.). Natural sounds (Nature; 28 samples or 3.2 %) contained recorded samples of natural phenomena such as fire burning, water ripple, flowing stream, wind breeze, hurricane, and thunder. Synthesized clips (Syn; 443 samples or 50.1 %) consisted of digitally generated sounds (e.g. pure tones and frequency-modulated sweeps), and recordings of man-made environmental sounds, such as engine noise, police siren, drilling, clock ticking, and sounds resulting from metallic bombardment. White noises (WhiteN; 21 samples or 2.3 %), were band-passed noises between 10 – 10000 Hz with different low/high-pass filters (e.g. 500, 1000, 2000 and 7500 Hz) and frequency bandwidths (min and max bandwidth between 390 to 9900 Hz). The remaining samples and data associated with them were discarded, as these sounds were not easily classified with mutually exclusive criteria. All stimuli were digitized and processed with a sampling frequency of 44100 Hz, and were 8-bit mono-recorded sound clips.

**Data Analysis**—Results are based on a *post-hoc* database analysis to determine if auditory memory performance of monkeys varied differentially across the seven sound types. From the available data, the study included sessions where the monkeys behavioral performance on both match and nonmatch trials was 60 percent correct or above. This behavioral criterion resulted in, on average, 70% of all behavioral sessions per monkey being included in the analysis. The criterion selection provided satisfactory performance from each monkey, while allowing enough response data for statistical analyses. Forty sessions (2000 trials) of data from each monkey were used (between February and June 2006), since the six monkeys received the original DMTS training at different times with differing numbers of total trials to criterion performance.

The current study employed go/no-go response rule for match and nonmatch trials. Performance data of match and nonmatch trials were analyzed separately. In match trials, both sounds presented as the first and second sound were the same and a button press response was required to release the food reward. Repeated-measures ANOVAs (SPSS 13.0; Chicago, IL) were conducted to examine auditory memory performance of match trials. For match trials, gender was a between-subject factor and sound type was a within-subject factor for conducting repeated-measures ANOVAs. In contrast, during nonmatch trials the two sounds presented were different and no button press response was to be made. Particular sound stimuli could either be presented as the sample stimulus (first position) or as the test stimulus (second position) on different trials. Thus for nonmatch trials, because there are two additional factors, the sound type of the first sound, and the sound type of the second sound, rather than an ANOVA, linear regression analysis (SPSS 13.0; Chicago, IL) was used to assess both factors. Here, percent correct of a given sound pairing was the dependent variable. Regressions were conducted hierarchically with gender entered on the first step to account for between-subject variability. Sound type presented as the first sound or the second sound was then entered to account for within-subject variability. Paired-sample *t*-tests were used for preplanned comparisons, and examined performance differences between monkey vocalizations and the other six sound types. Parallel analyses were used to examine differences of response latency when subjects gave correct go responses for two matching sounds (repeated-measure ANOVAs), and when subjects erroneously gave go responses for two nonmatching sounds (linear regression analysis).

## Experiment 2

After obtaining the results from Exp. 1 over a large number of behavioral testing sessions and analyzing them in a post-hoc manner, Exp. 2 was designed to exert more control over the comparison of sound exposures by using same numbers of sound stimuli across the seven sound types. In particular, presentations of sound stimuli during nonmatch trials were systematically organized, in order to reveal if particular sound types would improve auditory memory performance. The present design examines if the sound effects on memory performance derived from Exp. 1 could be replicated by Exp. 2.

**Subjects**—Exp. 2 used four monkeys, three males and one female, that participated in, and were housed as in Exp. 1.

**Auditory Stimuli**—For each of the seven sound types, 28 exemplars were chosen to represent each sound type used in Exp. 1 [animal vocalizations (Anivoc), human vocalizations (Hvoc), monkey vocalizations (Mvoc), music clips (Music), natural sounds (Nature), synthesized clips (Syn) and band-passed white noises (WhiteN)], for a total of 196 sounds. For monkey vocalizations, natural sounds and white noises, new stimuli were created in the same fashion as Exp. 1. New monkey vocalizations were recorded in a natural monkey reserve (South Carolina, USA; by the author A.P.). Calls representing coos, grunts, screams and harmonic arches were chosen from several hundred examples (frequency range: 100 – 10000 Hz, mean frequency: 1660 Hz).

**Procedures**—Exp. 2 was conducted approximately two years after Exp. 1, and monkeys had been receiving the auditory DMTS training with the 196 sound stimuli for Exp. 2 over the preceding six to eight weeks for a separate experiment. The same go/no-go response rule for the auditory DMTS task from Exp. 1 was used. The memory delay between two sounds (five seconds) and other training parameters were same as Exp. 1, with the exception that daily sessions consisted of 84 trials (42 match and 42 nonmatch trials; a ratio of 1 controlled by LabView software) instead of 50 to allow for controlled sound pairings on nonmatch trials.

**Pre-training**—Monkeys were first accustomed to 84 trials per session daily on the trial-unique DMTS task. All sounds were evenly distributed between seven control folders, containing four exemplars from each of the seven sound types. Two of the seven sound folders were pre-selected for each session/monkey pseudorandomly. The order and combination of the seven folders were randomized, and thus a given stimulus was repeated once on average every three training days. With pre-training monkeys normally took three to five days to reach the criterion of 80% or better before assessment of the auditory memory performance.

**Testing**—Everyday, the sound presentations were systematically organized, so that a given sound stimulus would appear either in match or nonmatch trials. A given stimulus was used once per daily session, and could be repeated on two successive days at most. On match trials, six sound exemplars from each sound type were used per day. On nonmatch trials, another 12 stimuli from each sound type were used per day. Moreover, positions of sound presentations on nonmatch trials (i.e. appeared as the sample or test sound stimulus) were completely counterbalanced among the seven sound types. There were no two sounds from the same sound type presented within a single nonmatch trial. Nonmatch trials in Exp. 2, hence, examined memory performance of monkeys when they discriminated one sound type against another type. The testing phase lasted for 10 to 15 daily sessions to achieve 10 sessions that met the performance criterion.

**Data Analysis**—Repeated-measure ANOVAs and linear regression analysis were used for memory performance of monkeys during match and nonmatch trials respectively, as in Exp.



1. Ten sessions for each monkey (approximately 85% of the behavioral sessions per monkey over 2–3 weeks) were used for data analysis where their memory performance for both match and nonmatch trials was correct on at least 60% of the trials for each trial type. As only one female monkey was included in this experiment, gender was not included as a between-subject factor. Based on the results of Exp. 1, effects of sound type were mainly due to performance associated with monkey vocalizations presented as the second sound. Preplanned comparisons were then focused on memory performance difference between this sound type and the other sound types (see methods, Experiment 1).

**Acoustic Analyses of Sounds**—To determine acoustic characteristics within each sound type, modulation spectra, adopted from Cohen et al. (2007), were created for the seven sound types and originally developed by Singh and Theunissen (2003). It is analogous to decomposing a sound waveform into a series of sine waves. A (log) spectrographic representation of each auditory stimulus could then be decomposed into a series of sinusoidal gratings that characterized the temporal modulation (in Hz) and the spectral modulations (in cycles per Hz or octave) of the stimulus. Modulation spectra of sound samples within a particular sound type were then averaged, and presented as the squared amplitude of the temporal and spectral modulation rates of that sound type.

The mathematic algorithm of the modulation spectrum first calculated the spectrographic representation for each sample of each sound type. It utilized a filter bank of Gaussian-shaped filters whose gain function had a bandwidth of 32 Hz. The 299 filters with center frequencies ranging from 32 Hz to 10 kHz, and the corresponding Gaussian-shaped windows in the time domain had a temporal bandwidth of 5 ms. These parameters defined the time-frequency scale of the spectrogram and the upper limits of the spectral and temporal modulation frequencies that could be characterized by the spectrogram: 16.25 cycles/kHz and 100.5 Hz respectively. The two-dimensional Fourier transform of each sound's log spectrogram was calculated for non-overlapping 1-second segments using a Hamming window. The modulation spectrum of each sound stimulus was calculated by averaging the power (amplitude squared) of the two-dimensional Fourier transform. The final modulation spectrum of a sound type was obtained by averaging individual modulation spectra from each sound stimulus within that sound type. All the spectral, temporal calculations and their visual presentations were created with MATLAB (The Math Works; Natick, MA).

The harmonics-to-noise ratio (HNR, expressed in dB), the degree of acoustic periodicity, was generated for each sound sample, using the freely available phonetic software, Praat (Boersma and Weenink, 2007; <http://www.fon.hum.uva.nl/praat/>). The HNR value served as an indicator of sound quality against noise, as how much acoustic energy of a signal was devoted to harmonics over time, relative to that of the remaining noise (i.e. representing nonharmonic, irregular, or chaotic acoustic energy). The HNR algorithm determined the degree of periodicity of a sound,  $x(t)$ , based on finding a maximum autocorrelation,  $r'_x(\tau_{\max})$ , of the signal at a time lag ( $\tau$ ) greater than zero.

$$\text{HNR (dB)} = 10 * \log_{10} \left[ \frac{\{r'_x(\tau_{\max})\}}{\{1 - r'_x(\tau_{\max})\}} \right].$$

## Results

### Experiment 1

**Match Memory Performance with Regard to the Seven Sound Types**—Males expressed significantly better auditory memory performance than females regardless of sound type, when subjects determined if the two sounds were the same (Fig. 2A). There was a main effect of gender (Repeated-measure ANOVAs,  $F_{(1,4)} = 10.48$ ;  $p < 0.05$ ), but no effect of sound

( $F_{(6,24)} = 0.13; p > 0.99$ ). There was also no interaction effect between gender and sound ( $F_{(6,24)} = 0.41; p = 0.87$ ). The study then examined effects of sound type on response latency during match trials. There was no main effect of gender ( $F_{(1,4)} = 0.01; p = 0.93$ ), or sound ( $F_{(6,24)} = 0.54; p = 0.77$ ), and no interaction effect ( $F_{(1,4)} = 1.51; p = 0.72$ ) on response latencies during the auditory DMTS task (results not shown). Auditory memory performance involved in two matching sounds is independent of response latencies for button-pressing.

**Nonmatch Memory Performance Between Male and Female Monkeys**—Linear regression analysis was used to examine memory performance on nonmatch trials. On the first step of the analysis, gender was entered and significantly accounted for 4 percent of the variance ( $R^2_{\text{change}} = 0.04$ ,  $F_{\text{change}(1,232)} = 8.57$ ,  $p < 0.005$ ). There was a main effect of gender in which males performed significantly better than females (Fig. 2B), parallel to the findings for match-trial conditions. On the second step of the analysis, sound type of the first sound was added to the regression model and did not account for any significant variance ( $R^2_{\text{change}} = 0.02$ ,  $F_{\text{change}(6,226)} = 0.61$ ,  $p = 0.72$ ). On the third step, sound type of the second sound was added to the regression model and there was a significant ( $R^2_{\text{change}} = 0.09$ ,  $F_{\text{change}(6,220)} = 3.89$ ,  $p < 0.005$ ) main effect of sound type when presented in the second sound position (Figure 3). This effect was further analyzed using paired-sample *t*-tests. When the second sound was a monkey vocalization, our animal subjects yielded significantly better memory performance than when the second sound was a human vocalization ( $t(5) = 4.13$ ,  $p < 0.05$ ), an animal vocalization ( $t(5) = 2.74$ ,  $p < 0.05$ ), a music clip ( $t(5) = 4.45$ ,  $p < 0.05$ ), a natural sound ( $t(5) = 5.45$ ,  $p < 0.05$ ), or a synthesized clip ( $t(5) = 3.19$ ,  $p < 0.05$ ). In addition, nonmatch trials associated with human or animal vocalizations also yielded significantly better memory performance than those using natural sounds (Hvoc versus Nature:  $t(5) = 5.98$ ,  $p < 0.05$ ; Anivoc versus Nature:  $t(5) = 7.57$ ,  $p < 0.05$ ). The study also evaluated if an interaction between gender and sound type presented as the second sound would contribute to the variance associated with nonmatch memory performance. This last factor was entered to the regression model, and did not account for any significant variance ( $R^2_{\text{change}} = 0.003$ ,  $F_{\text{change}(6, 199)} = 0.11$ ,  $p = 0.99$ ).

We further examined if monkeys would perform better using sounds with relatively simple acoustic structure (e.g. pure tone and frequency-modulated sweep). A grouping of simple sounds (10 samples) was culled from synthesized clips (Syn), and the corresponding memory performance for that group of 10 simple sounds was compared to the other seven sound types. Memory performance associated with simple sound type showed a similar level of accuracy to the other seven sound types. These findings suggest that a simple acoustic structure did not make it easier for the monkeys to hold information across a memory delay, and instead, factors beyond purely acoustic properties may be more important.

Analysis of response latency on nonmatch trials used the same regression analysis. On the first step, gender was added to the model and was not significant ( $R^2_{\text{change}} = 0.01$ ,  $F_{\text{change}(1, 182)} = 2.27$ ,  $p = 0.13$ ). On the second step, sound type of the first sound accounted for no additional variance ( $R^2_{\text{change}} = 0.02$ ,  $F_{\text{change}(6, 176)} = 0.50$ ,  $p = 0.81$ ). Lastly, sound type of the second sound was added to the model, and marginally accounted for 7% of the variance ( $R^2_{\text{change}} = 0.07$ ,  $F_{\text{change}(6, 170)} = 2.06$ ,  $p = 0.06$ ).

Males performed better than females on both match and nonmatch trials, regardless of the seven sound types. It is also important to inspect their individual data to assess memory performance range between genders. Table 1 illustrates average individual memory performance of the six monkeys across the seven sound types, separated by gender and trial type (match or nonmatch).

## Experiment 2

**Effects of Sound Type on Match Memory Performance**—Parallel to findings from Exp. 1, there was no main effect of sound ( $F_{(6,18)} = 0.95$ ;  $p = 0.48$ ). Auditory memory

performance for two matching sound stimuli was consistently good across the seven sound types (overall mean = 91.00, standard error =  $\pm 2.96$ ).

**Effects of Sound Type on Nonmatch Memory Performance**—Exp. 2 was a follow-up study to examine if monkey vocalizations served as better acoustic stimuli when monkeys discriminated them from other sound types during a memory task. As expected, there was a main effect of sound type presented as the second sound ( $R^2_{\text{change}} = 0.15$ ,  $F_{\text{change}}(6, 152) = 6.16$ ,  $p < 0.005$ ), but no main effect of sound type presented as the first sound ( $R^2_{\text{change}} = 0.02$ ,  $F_{\text{change}}(6, 158) = 0.80$ ,  $p = 0.57$ ), similar to the findings of experiment 1. When the second sound was a monkey vocalization, animal subjects yielded significantly better memory performance than those when it was an animal vocalization, a music clip, a synthesized clip, or a white noise (paired-sample  $t$ -tests,  $p < 0.05$ ; Figure 4). In contrast to Exp. 1, the study did not reveal significant performance differences between human or animal vocalizations and other non-vocalization sound types.

Sound pairings for nonmatch trials in Exp. 2 were systemically organized and counterbalanced so that each trial consisted of stimuli from two distinct sound types. Different pairings of sounds from distinct sound types may then influence auditory memory performance of monkeys during auditory discrimination and recognition. Memory improvement due to monkey vocalizations presented as the second sound may depend on which sound type was presented as the first sound. Thus, this factor was entered to the regression model:  $R^2_{\text{change}} = 0.04$ ,  $F_{\text{change}}(10, 142) = 1.03$ ,  $p = 0.42$ . The result showed no interaction between a given sound type and monkey vocalizations when considering the first and second sound position. This suggests that auditory memory performance was improved accordingly when a monkey vocalization test stimulus (second position) was compared against a sample stimulus of any sound type.

**Effects of Sound Type on Response Latencies for Match and Nonmatch Memory Performance**—A robust sound effect was shown in that monkey vocalizations generally provided advantages to our animals during auditory memory performance. Another behavioral measure, response latency of the button-press was assessed to determine if it would also indicate a similar relationship between sound type and memory performance. On match trials, there was a main effect of sound on response latency ( $F_{(6,18)} = 13.29$ ;  $p < 0.05$ ). Figure 5 illustrates average response latencies across the seven sound types during match-trial conditions, and indicates the effect of sound type mainly due to monkey vocalizations. Paired sample  $t$ -tests were used to reveal latency differences between monkey vocalizations and the other six sound types. Subjects showed significantly faster go-responses (correct) for monkey vocalizations than for any of the other six sound types ( $p < 0.05$ ).

Regression analysis of nonmatch-trial conditions for response latency showed neither effect of sound type presented as the first sound ( $R^2_{\text{change}} = 0.03$ ,  $F_{\text{change}}(6, 127) = 0.90$ ,  $p = 0.50$ ) or the second sound ( $R^2_{\text{change}} = 0.03$ ,  $F_{\text{change}}(6, 121) = 1.03$ ,  $p = 0.41$ ), when monkeys produced incorrect button presses.

**Acoustic Analyses of Sounds**—One possible explanation for the above findings which describe better memory performance associated with monkey vocalizations, is that differences in acoustic properties between monkey and non-monkey sound types may account for the observed difference in performance. Such acoustic differences may then facilitate auditory discrimination when monkeys determined two sounds to be different. To explore this possibility, we quantitatively compared the acoustic properties of the seven sound types (see methods). Figure 6 displays a series of modulation spectra for the seven sound types. For the three vocalization sound types, their modulation spectra have most of their acoustic energy at low to medium spectra and temporal frequencies, and their power levels decrease rapidly at high frequencies. This pattern is characteristically to animal vocalizations, including those



produced by birds, monkeys and humans (Singh and Theunissen, 2003;Cohen et al., 2007). In contrast, there is remarkable acoustic energy at medium to high spectral frequencies in music clips and synthesized clips. These results match with the expected acoustic energy profiles of these sounds in that music segments and man-made environmental sounds contain a wider range of frequencies and energy sources from higher spectral levels. For natural sounds and white noises, acoustic energy dominantly resides at very low spectral and temporal frequencies, consistent to monotonous features of these sound types.

The harmonic-to-noise ratio (HNR) indicates if certain sound types tend to carry more harmonic components over time relative to noise (Fig. 7). All three vocalization sound types and two of the non-vocalization sound types, music and synthesized clips, have positive HNR values, showing that they carry large, regular harmonic contents compared to noise. Natural sound and white noise have negative HNR values, reflecting the nonharmonic, irregular, or chaotic acoustic energy predominantly present in these types. The natural phenomena we recorded here mainly related to wind-, fire- and water-related events. These sounds resemble perceptual and acoustic features of the band-passed white noises used in the current study. We tested if an increased acoustic periodicity of a sound type (i.e. harmonic components against background noises, HNR) is associated with increased auditory memory performance in monkeys, especially on nonmatch trials during auditory discrimination. Correlational comparisons between each sound type and the corresponding nonmatch memory performance were conducted. For each sound stimulus, average performance (percentage correct) at nonmatch trials associated with that sound was calculated and averaged per session by subject. For each sound type, a Pearson's correlation coefficient (SPSS 13.0; Chicago, IL) was calculated between memory performance associated with each sound stimulus and its respective HNR value. These results show there is no significant relationship between acoustic quality and memory performance associated with a given sound type.

## Discussion

Using a delayed matching-to-sample task, the present findings suggest a measurable effect of sound type influencing auditory recognition memory of monkeys. In the first experiment, monkeys show better auditory recognition memory with vocalizations, strongest for species-specific monkey vocalizations, on nonmatch trials after a fixed 5-second memory delay. Additionally, male monkeys demonstrated better auditory recognition memory than female monkeys on both match and nonmatch trials, regardless of sound type. The findings of the second experiment, using a trial-unique design with balanced presentation of sound types, once again showed robust memory performance on nonmatch trials where monkey vocalizations as one of the sounds, and a decreased response latency to match trials using monkey vocalizations.

Evidence for increased memory performance comes primarily from the nonmatch trials, as behavioral performance on the match trials may have reached asymptote creating a "ceiling" effect. However, in addition to the higher performance level on nonmatch trials, the latency of correct responses is significantly faster on match trials using monkey vocalizations. This decreased latency to respond to monkey vocalizations is compatible with the increased number of correct responses on nonmatch trials with monkey vocalizations. The results suggest that the monkeys, both perceptually and behaviorally, distinguish their own species-specific sounds preferentially. The use of monkey vocalizations offers a performance advantage with behavioral specificity, not just over excitation. While the monkeys are responding faster for match trials using monkey vocalizations they are also better at withholding, or not responding, during nonmatch trials where the second sound is a monkey vocalization whereas they are responding erroneously more often to other sound types during those trials. The effects of monkey vocalizations on this short-term memory performance task suggest auditory recognition memory of rhesus monkeys may not be universally poor, in comparison to visual

recognition memory, as concluded by prior studies (D'Amato and Colombo, 1985; Wright, 1998, 1999; Fritz et al., 2005), which did not specifically address the use of monkey vocalizations. Future studies will need to ascertain the influence of species-specific monkey vocalizations at longer memory delays.

The acoustic differences of the different sound type groupings do not account for different levels of memory performance. For example, the three vocalization sound types, humans, monkeys, and other animals, share similar spectral and temporal modulations, and similar profiles of their acoustic energy spreads and densities. Despite their acoustic similarities, monkey vocalizations, relative to human and animal ones, provide an advantage during the recognition memory task and serve as better acoustic cues than non-vocalization sounds. Distinctiveness in sound structure, as shown by illustrations of modulation spectra and HNR values (Fig. 6 and Fig. 7), does not modulate memory performance. Neither acoustically simple (natural sounds and white noises) nor complex sound types (music and synthesized clips) make the memory task easier for the animal subjects. Overall, the findings of the acoustical sound analyses suggest monkeys do not simply rely on global spectrotemporal differences across sounds to assist auditory discrimination and recognition for memory use. The findings of both experiments reinforce the notion that better memory performance is selectively associated with monkey vocalizations suggesting that factors embedded in the acoustic properties, e.g., significance and/or familiarity, of monkey vocalizations make them preferable to monkeys during memory performance.

One reason monkey vocalizations may evoke better behavioral performance across memory delay intervals is that monkey vocalizations may be more familiar to our subjects than other sound types. Familiarity and experience with this particular sound type may contribute to their special status. Expertise in facial recognition, analogous to species-specific vocalizations, greatly influences discrimination performance in humans (Diamond and Carey, 1986), chimpanzees (Parr and Heintz, 2006), Japanese macaques (Tomonaga, 1994), and rhesus monkeys (Parr and Heintz, 2008). They are examples of face inversion effects, in which humans discriminate human faces easily when they are presented upright versus inverted. These nonhuman primates show an inversion effect to conspecific faces and even sometimes to human faces, but not unfamiliar faces and objects (e.g., heterospecific monkey faces and houses). Future studies could include heterospecific vocalizations from other primate species during an auditory memory task.

Another possibility is the converging evidence from the present study and other multi-disciplinary research proposes that biological, and/or ethological significance of monkey vocalizations, acoustically embedded inside these sounds, are more readily recognized by monkeys and this may mediate memory performance. Compatible with the current findings assessing auditory memory are studies involving auditory discrimination. Japanese macaques learn to discriminate conspecific coo calls faster than heterospecific coo calls (Petersen et al., 1984); and rhesus macaques responded to food-related species-specific vocalizations based on their functional referents (i.e. the quality of food) but not physical features (Hauser, 1998; Gifford et al., 2003). Species-specific vocalizations seem to be unique, as animal subjects not only attend to physical quality of sounds (e.g. timing and frequency bandwidth) but also the acoustic cues derived biological/ethological significance embedded inside (also called "acoustic signatures"; Fitch, 2000). Electrophysiological studies demonstrate higher-order auditory regions, for example ventrolateral prefrontal cortices, encode monkey vocalizations according to functional referents embedded inside the sounds, for instance low/high food quality and food/non-food differences (Gifford et al., 2005; Cohen et al., 2006; Russ et al., 2007). In the present task of 5-second delays, memory performance for two matching sounds appears to be asymptotic across sound types, while response latencies associated with monkey vocalization are the fastest. The authors speculate if memory performance using monkey

vocalizations would be well maintained, and better than other sound types if memory delays were sufficiently long. Therefore, future studies could focus on the influences of sound types when monkeys are challenged with long memory delays, in order to examine if monkeys' preferences on their own species-specific sounds would generalize to more demanding memory tests.

Species-specific vocalizations, analogous to faces, may provide essential cues for identity, sex, age, emotional status, and kinship for social interaction and survival (Ghazanfar and Hauser, 2001). Neural processing of faces in humans and monkeys is along the ventral visual information pathway and electrophysiological studies reveal neural correlates of face detection and recognition in the fusiform face area, occipital face area, and a region of superior temporal sulcus (fSTS) (Kanwisher and Yovel, 2006). The current behavioral results for monkey vocalizations imply that perhaps a network of auditory brain regions specialized in processing species-specific vocalizations is capable of influencing memory processing similar to visual processing of faces. Auditory discrimination utilizing species-specific vocalizations requires belt/parabelt regions and superior temporal gyri (STG) along the primate auditory system. Lesions to these areas, particularly the rostral regions of STG, abolish the functional advantages provided by monkey vocalizations in auditory discrimination learning (Kupfer et al., 1977). And impair monkeys' ability to hold auditory information across memory delays using the delayed matching-to-sample task (Colombo et al., 1990, 1996; Fritz et al., 2005). High-order auditory processing of complex sounds, including species-specific vocalizations, illustrates evidence from neuronal recording and imaging studies supporting a neural specialization for vocalization processing extending ventrally through the superior temporal gyrus and including prefrontal cortical regions (Tian et al., 2001, Cohen et al., 2004; Gil-da-Costa et al., 2004; Poremba et al., 2004, Romanski et al., 2005; Petkov et al., 2008; Remedios et al., 2009).

Auditory studies using the DMTS task (e.g. Wright, 1998, 1999; Fritz et al., 2005) do not separate memory performance into match and nonmatch trials, and instead combine them into measures of average memory performance. The present findings reveal the two trial types differentiate behaviorally during auditory memory performance, i.e., nonmatch trials are more influenced by different sound types. Critics may argue that the phenomenon is due to the nature of go/no-go response contingency, i.e., excitation versus inhibition of motor responses. However, implications from the current study lead the authors to reconsider the phenomenon of divergent behavior on match and nonmatch trials. Memory performance for two matching sounds is less susceptible to sound types, which significantly modulate memory performance during nonmatch trials containing two different sounds. It is perplexing that auditory recognition and discrimination are involved in both trial types, and yet there are expression differences for auditory memory. One possibility is that different levels of information processing are required for recognition and discrimination across match and nonmatch trials, e.g., simple versus complex tasks, and this difference may interact with memory delays. In vision, electrophysiological studies in monkeys differentiate neuronal profiles of inferior temporal cortices and prefrontal cortices when encoding perceptual information versus categorizing stimuli according to instructions of category-matching (Freedman et al., 2003; Muhammad et al., 2006). These studies propose a division of labor in the primate visual system for encoding, discrimination and recognition of task-relevant stimuli. Their findings may also support the suggestion of a network of multiple brain regions for different aspects of auditory processing. Future studies of the DMTS paradigm should be paired with functional imaging or neuronal recording to investigate if a similar division of labor for information processing is evoked by auditory behaviors, and how a series of brain regions accommodate such task challenges.

The results of Exp. 1 reveal an effect of gender on auditory memory performance of the current DMTS task for both match and nonmatch trial performance. Individual and group performance

data suggests that male monkeys show reliable, consistent performance accuracy at high levels in most sound type conditions, while female monkeys often show fluctuations of memory performance. Gender effects on perception, learning, and memory have been extensively studied in humans. Males generally excel in spatial tasks, such as mental rotation, maze learning, map reading, distance/location finding (Kimura, 1996; Postma et al., 1998; Rizk-Jackson et al., 2006). Females generally excel in nonspatial processing and nonspatial components of spatial tasks, such as verbal memory, face recognition, object/landscape recognition and memory (Kimura and Clarke, 2002; Levy et al., 2005; Voyer et al., 2007). There are also similar reports in rodents (Jonasson, 2005; Sutcliffe et al., 2007) and non-human primates (Lacreuse et al., 1999, 2005), that male excellence in spatial processing and female excellence in non-spatial processing. Several theories have been used for describing and explaining gender differences on performance of cognitive and behavioral tasks. The evolutionary history of humans, such as sexual selection for mate competition, task divisions between foraging, and nurturing young (Eals and Silverman, 1994; Ecuyer-Dab and Robert, 2004; Sutcliffe et al., 2007), are used to describe why males and females perform differently in spatial and non-spatial tasks respectively, and the evolutionary history of non-human primates may also relate to the gender differences observed here in auditory memory performance.

Interactions between hormonal actions in the brain and gender are suggested to affect cognitive performance in rodents (Warren and Juraska, 1997; Sutcliffe et al., 2007), and humans (Kimura and Hampson, 1994; Kimura, 1996), where females tend to perform better in spatial tasks at low-estrogen levels than at high-estrogen levels. Most of the findings concerning hormonal effects on gender differences in humans and non-human mammals are predominantly based on studies assessing spatial abilities, for instance, maze learning and space navigation. Interpretations about hormonal and physiological mechanisms on behaviors may not correlate with non-spatial domains of cognition and behavior on the same experimental subjects and would need to be investigated in auditory memory tasks.

To date, there is a lack of consistent evidence on how gender plays a role in nonspatial components of auditory perception, learning, and memory functions. Some field studies show gender differences in recognizing calls during mate selection and competition or producing food-associated calls. Female rhesus monkeys generally produce more monkey calls in food-associated contexts (e.g. coo, grunt, warble and harmonic arch) than males (Hauser and Marler, 1993). Female monkeys also show a greater responsiveness to copulation calls than males (Hauser, 2007). Overall, these field studies suggest that females may have a heightened capacity to perceive and recognize acoustic differences regarding call exemplars and caller identity, which are important for females to evaluate sexual fitness of males during male selection and reproduction. However on our auditory memory task, with a small sample size, the gender advantage was in the opposite direction, with males showing higher performance levels than females.

In auditory tasks specifically, there is some evidence that gender differences may rely on differences in auditory sensitivity. Human females are more sensitive than males to high frequencies ranging from 8000 to 16000 Hz when test stimuli are pure tones and frequency sweeps (Chung et al., 1983; Lopponen et al., 1991; Hallmo et al., 1994; Pearson et al., 1995; Dreisbach et al., 2007), though others suggest no gender difference at all (Osterhammel and Osterhammel, 1979; Frank, 1990; Betke, 1991), and it is unknown in rhesus macaques. Auditory sensitivity is not sufficient to explain gender differences on the auditory DMTS task. The present study uses a wide variety of sounds with different acoustic profiles, from simple pure tones, white noises to complex music clips, vocalizations, and man-made environmental sounds. Parallel to other primate studies (Cohen et al., 2006; Ghazanfar et al., 2007), monkeys do not seem to simply rely on acoustic differences among sound stimuli for auditory behaviors.

Limited and inconclusive evidence, which differs by species of subjects, experimental designs, and complexity of acoustic stimuli, neither agree nor contradict the present gender specific results. Although the current results are based on a very small sample size, wherein two of the three females clearly performed poorly compared to the males and the other female showed performance accuracy closer to the performance of the lowest male monkeys (Table 1), implications of the current study suggest follow-up investigations on gender differences in auditory memory performance of monkeys as the use of male monkeys has predominated in previous research.

As we have discussed, multi-disciplinary experimental approaches converge on the conclusion that species-specific sounds, usually bearing biological or ethological significance, are more readily processed, analyzed, and recognized by humans and monkeys. Monkey vocalizations may therefore be salient and potent conveyors of acoustic information increasing memory or recognition performance, and may be mediated by a network of specialized brain regions for processing species-specific sounds similar to face processing in monkeys and humans.

## Acknowledgements

We would like to thank Dr. Mortimer Mishkin for his invaluable support of our research, Dr. Yale Cohen for providing guidance and MATLAB scripts for acoustic analyses, and Dr. Robert McMurray for advising on methods of statistical analysis. This work was supported by funding awarded to Amy Poremba from University of Iowa Startup Funds and NIH, NIDCD, DC0007156.

## References

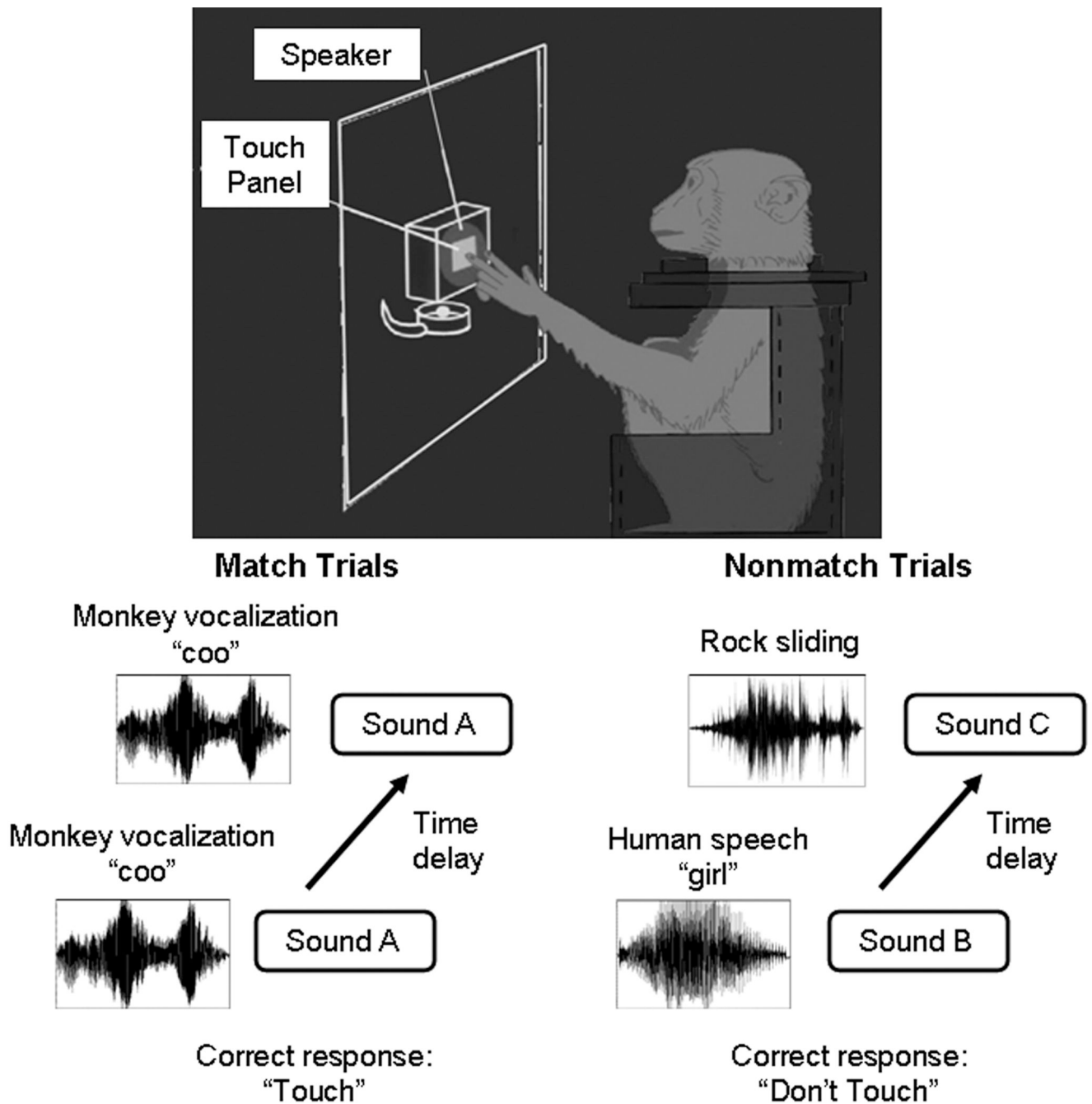
- Amrhein PC, McDaniel MA, Waddill P. Revisiting the picture-superiority effect in symbolic comparisons: do pictures provide privileged access? *J. Exp. Psychol. Learn. Mem. Cogn* 2002;28:843–857. [PubMed: 12219794]
- Belin P. Voice processing in human and non-human primates. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* 2006;361:2091–2107. [PubMed: 17118926]
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature* 2000;403:309–312. [PubMed: 10659849]
- Bélizaire G, Fillion-Bilodeau S, Chartrand JP, Bertrand-Gauvin C, Belin P. Cerebral response to 'voiceness': a functional magnetic resonance imaging study. *Neuroreport* 2007;18:29–33. [PubMed: 17259856]
- Betke K. New hearing threshold measurements for pure tones under free-field listening conditions. *J. Acoust. Soc. Am* 1991;89:2400–2403. [PubMed: 1860999]
- Boersma, P.; Weenink, D. Praat: doing phonetics by computer (Version 4.6.38) [Computer program]. 2007 [Retrieved November 19, 2007]. from <http://www.praat.org/>.
- Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM. Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn. Mem* 1999;6:572–599. [PubMed: 10641763]
- Bulthoff I, Newell FN. The role of familiarity in the recognition of static and dynamic objects. *Prog. Brain Res* 2006;154:315–325. [PubMed: 17010720]
- Chung DY, Mason K, Gannon RP, Willson GN. The ear effect as a function of age and hearing loss. *J. Acoust. Soc. Am* 1983;73:1277–1282. [PubMed: 6853839]
- Cohen MA, Horowitz TS, Wolfe JM. Auditory recognition memory is inferior to visual recognition memory. *Proc. Natl. Acad. Sci. U. S. A* 2009;106:6008–6010. [PubMed: 19307569]
- Cohen YE, Hauser MD, Russ BE. Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. *Biol. Lett* 2006;2:261–265. [PubMed: 17148378]
- Cohen YE, Russ BE, Gifford GW 3rd, Kiringoda R, MacLean KA. Selectivity for the spatial and nonspatial attributes of auditory stimuli in the ventrolateral prefrontal cortex. *J. Neurosci* 2004;24:11307–11316. [PubMed: 15601937]



- Cohen YE, Theunissen F, Russ BE, Gill P. Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. *J. Neurophysiol* 2007;97:1470–1484. [PubMed: 17135477]
- Colombo M, D'Amato MR, Rodman HR, Gross CG. Auditory association cortex lesions impair auditory short-term memory in monkeys. *Science* 1990;247:336–338. [PubMed: 2296723]
- Colombo M, Rodman HR, Gross CG. The effects of superior temporal cortex lesions on the processing and retention of auditory information in monkeys (*Cebus apella*). *J. Neurosci* 1996;16:4501–4517. [PubMed: 8699260]
- D'Amato MR, Colombo M. Auditory matching-to-sample in monkeys (*Cebus apella*). *Anim. Learn. Behav* 1985;13:375–382.
- Diamond R, Carey S. Why faces are and are not special: an effect of expertise. *J. Exp. Psychol. Gen* 1986;115:107–117. [PubMed: 2940312]
- Dreisbach LE, Kramer SJ, Cobos S, Cowart K. Racial and gender effects on pure-tone thresholds and distortion-product otoacoustic emissions (DPOAEs) in normal-hearing young adults. *Int. J. Audiol* 2007;46:419–426. [PubMed: 17654083]
- Eals M, Silverman I. The hunter-gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. *Ethol. Sociobiol* 1994;15:95–105.
- Ecuyer-Dab I, Robert M. Have sex differences in spatial ability evolved from male competition for mating and female concern for survival? *Cognition* 2004;91:221–257. [PubMed: 15168896]
- Ehret G. Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. *Nature* 1987;325:249–251. [PubMed: 3808021]
- Fecteau S, Armony JL, Joanette Y, Belin P. An fMRI study. Is voice processing species-specific in human auditory cortex? *Neuroimage* 2004;23:840–848. [PubMed: 15528084]
- Fitch WT. The evolution of speech: a comparative review. *Trends Cogn. Sci* 2000;4:258–267. [PubMed: 10859570]
- Fitch WT, Fritz JB. Rhesus macaques spontaneously perceive formants in conspecific vocalizations. *J. Acoust. Soc. Am* 2006;120:2132–2141. [PubMed: 17069311]
- Frank T. High-frequency hearing thresholds in young adults using a commercially available audiometer. *Ear. Hear* 1990;11:450–454. [PubMed: 2073979]
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J. Neurosci* 2003;23:5235–5246. [PubMed: 12832548]
- Fritz J, Mishkin M, Saunders RC. In search of an auditory engram. *Proc. Natl. Acad. Sci. U. S. A* 2005;102:9359–9364. [PubMed: 15967995]
- Geissler DB, Ehret G. Auditory perception vs. recognition: representation of complex communication sounds in the mouse auditory cortical fields. *Eur. J. Neurosci* 2004;19:1027–1040. [PubMed: 15009150]
- Ghazanfar AA, Hauser MD. The auditory behaviour of primates: a neuroethological perspective. *Curr. Opin. Neurobiol* 2001;11:712–720. [PubMed: 11741023]
- Ghazanfar AA, Turesson HK, Maier JX, Dinther R, Patterson RD, Logothetis NK. Vocal-tract resonances as indexical cues in rhesus monkeys. *Curr. Biol* 2007;17:425–430. [PubMed: 17320389]
- Gifford GW 3rd, Hauser MD, Cohen YE. Discrimination of functionally referential calls by laboratory-housed rhesus macaques: implications for neuroethological studies. *Brain Behav. Evol* 2003;61:213–224. [PubMed: 12784058]
- Gifford GW 3rd, MacLean KA, Hauser MD, Cohen YE. The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. *J. Cogn. Neurosci* 2005;17:1471–1482. [PubMed: 16197700]
- Gil-da-Costa R, Braun A, Lopes M, Hauser MD, Carson RE, Herscovitch P, Martin A. Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. *Proc. Natl. Acad. Sci. U. S. A* 2004;101:17516–17521. [PubMed: 15583132]
- Hallmo P, Sundby A, Mair WS. Extended high-frequency audiometry: Air- and bone-conduction thresholds, age, and gender variations. *Scand. Audiol* 1994;23:165–170. [PubMed: 7997833]

- Hauser MD. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim. Behav* 1998;55:1647–1658. [PubMed: 9642008]
- Hauser MD. When males call, females listen: sex differences in responsiveness to rhesus monkey, *Macaca mulatta*, copulation calls. *Anim. Behav* 2007;73:1059–1065.
- Hauser MD, Marler P. Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behav. Ecol* 1993;4:194–205.
- Hienz RD, Jones AM, Weerts EM. The discrimination of baboon grunt calls and human vowel sounds by baboons. *J. Acoust. Soc. Am* 2004;116:1692–1697. [PubMed: 15478436]
- Jonasson Z. Meta-analysis of sex differences in rodent models of learning and memory: a review of behavioral and biological data. *Neurosci. Biobehav. Rev* 2005;28:811–825. [PubMed: 15642623]
- Kanwisher N, Yovel G. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B. Biol. Sci* 2006;361:2109–2128. [PubMed: 17118927]
- Kimura D. Sex, sexual orientation and sex hormones influence human cognitive function. *Curr. Opin. Neurobiol* 1996;6:259–263. [PubMed: 8725969]
- Kimura D, Clarke PG. Women's advantage on verbal memory is not restricted to concrete words. *Psychol. Rep* 2002;91:1137–1142. [PubMed: 12585526]
- Kimura D, Hampson E. Cognitive pattern in men and women is influenced by fluctuations in sex hormones. *Curr. Dir. Psychol. Sci* 1994;3:57–61.
- Kupfer K, Jurgens U, Ploog D. The effect of superior temporal lesions on the recognition of species-specific calls in the squirrel monkey. *Exp. Brain Res* 1977;30:75–87. [PubMed: 563342]
- Lacreuse A, Herndon JG, Killiany RJ, Rosene DL, Moss MB. Spatial cognition in rhesus monkeys: male superiority declines with age. *Horm. Behav* 1999;36:70–76. [PubMed: 10433888]
- Lacreuse A, Kim CB, Rosene DL, Killiany RJ, Moss MB, Moore TL, Chennareddi L, Herndon JG. Sex, age, and training modulate spatial memory in the rhesus monkey (*Macaca mulatta*). *Behav. Neurosci* 2005;119:118–126. [PubMed: 15727518]
- Levy LJ, Astur RS, Frick KM. Men and women differ in object memory but not performance of a virtual radial maze. *Behav. Neurosci* 2005;119:853–862. [PubMed: 16187814]
- Löppönen H, Sorri MRB. High-frequency air-conduction and electric bone-conduction audiometry. Age and sex variations. *Scand. Audiol* 1991;20:181–189. [PubMed: 1842287]
- Muhammad R, Wallis JD, Miller EK. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J. Cogn. Neurosci* 2006;18:974–989. [PubMed: 16839304]
- Murray EA, Mishkin M. Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. *J. Neurosci* 1998;18:6568–6582. [PubMed: 9698344]
- Osterhammel D, Osterhammel P. High-frequency audiometry. Age and sex variations. *Scand. Audiol* 1979;8:73–81. [PubMed: 515692]
- Parr LA, Heintz M. The perception of unfamiliar faces and houses by chimpanzees: influence of rotation angle. *Perception* 2006;35:1473–1483. [PubMed: 17286118]
- Parr LA, Heintz M. Discrimination of faces and houses by rhesus monkeys: the role of stimulus expertise and rotation angle. *Anim. Cogn* 2008;11:467–474. [PubMed: 18256863]
- Pearson JD, Morrell CH, Gordon-Salant S, Brant LJ, Metter EJ, Klein LL, Fozard JL. Gender differences in a longitudinal study of age-associated hearing loss. *J. Acoust. Soc. Am* 1995;97:1196–1205. [PubMed: 7876442]
- Petersen MR, Beecher MD, Zoloth SR, Green S, Marler PR, Moody DB, Stebbins WC. Neural lateralization of vocalizations by Japanese macaques: communicative significance is more important than acoustic structure. *Behav. Neurosci* 1984;98:779–790. [PubMed: 6487413]
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK. A voice region in the monkey brain. *Nat. Neurosci* 2008;11:367–374. [PubMed: 18264095]
- Poremba A, Malloy M, Saunders RC, Carson RE, Herscovitch P, Mishkin M. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature* 2004;427:448–451. [PubMed: 14749833]
- Postma A, Izendoorn R, De Haan EH. Sex differences in object location memory. *Brain Cogn* 1998;36:334–345. [PubMed: 9647682]

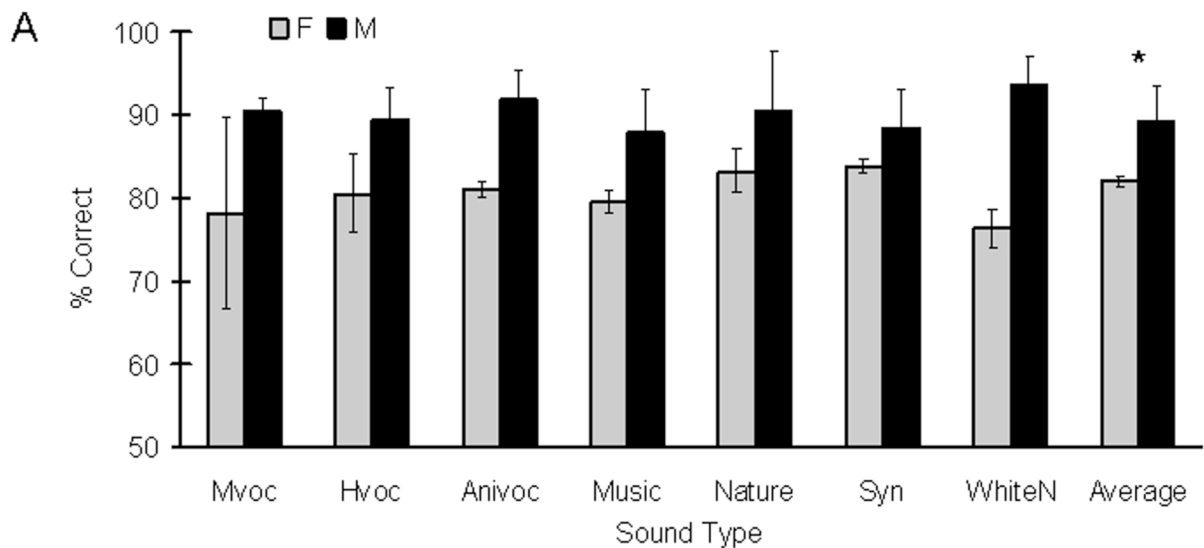
- Remedios R, Logothetis NK, Kayser C. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J Neurosci* 2009;29:1034–1045. [PubMed: 19176812]
- Rendall D. Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J. Acoust. Soc. Am* 2003;113:3390–3402. [PubMed: 12822809]
- Rizk-Jackson AM, Acevedo SF, Inman D, Howieson D, Benice TS, Raber J. Effects of sex on object recognition and spatial navigation in humans. *Behav. Brain Res* 2006;173:181–190. [PubMed: 16887201]
- Romanski LM, Averbach BB, Diltz M. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J. Neurophysiol* 2005;93:734–747. [PubMed: 15371495]
- Russ BE, Lee YS, Cohen YE. Neural and behavioral correlates of auditory categorization. *Hear. Res* 2007;229:204–212. [PubMed: 17208397]
- Seifert LS. Activating representations in permanent memory: different benefits for pictures and words. *J. Exp. Psychol. Learn. Mem. Cogn* 1997;23:1106–1121. [PubMed: 9293624]
- Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J. Acoust. Soc. Am* 2003;114:3394–3411. [PubMed: 14714819]
- Sutcliffe JS, Marshall KM, Neill JC. Influence of gender on working and spatial memory in the novel object recognition task in the rat. *Behav. Brain Res* 2007;177:117–125. [PubMed: 17123641]
- Theunissen FE, Shaevitz SS. Auditory processing of vocal sounds in birds. *Curr. Opin. Neurobiol* 2006;16:400–407. [PubMed: 16842993]
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. *Science* 2001;292:290–293. [PubMed: 11303104]
- Tomonaga M. How laboratory-raised Japanese monkeys (*Macaca fuscata*) perceive rotated photographs of monkeys: Evidence for an inversion effect in face perception. *Primates* 1994;35:155–165.
- Voyer D, Postma A, Brake B, Imperato-McGinley J. Gender differences in object location memory: a meta-analysis. *Psychon. Bull. Rev* 2007;14:23–38. [PubMed: 17546728]
- Warren SG, Juraska JM. Spatial and nonspatial learning across the rat estrous cycle. *Behav. Neurosci* 1997;111:259–266. [PubMed: 9106666]
- Wright AA. Auditory list memory in rhesus monkeys. *Psychol. Sci* 1998;9:91–98.
- Wright AA. Auditory list memory and interference processes in monkeys. *J. Exp. Psychol. Anim. Behav. Process* 1999;25:284–296. [PubMed: 10423854]
- Zola SM, Squire LR, Teng E, Stefanacci L, Buffalo EA, Clark RE. Impaired recognition memory in monkeys after damage limited to the hippocampal region. *J. Neurosci* 2000;20:451–463. [PubMed: 10627621]
- Zoloth SR, Petersen MR, Beecher MD, Green S, Marler P, Moody DB, Stebbins W. Species-specific perceptual processing of vocal sounds by monkeys. *Science* 1979;204:870–873. [PubMed: 108805]



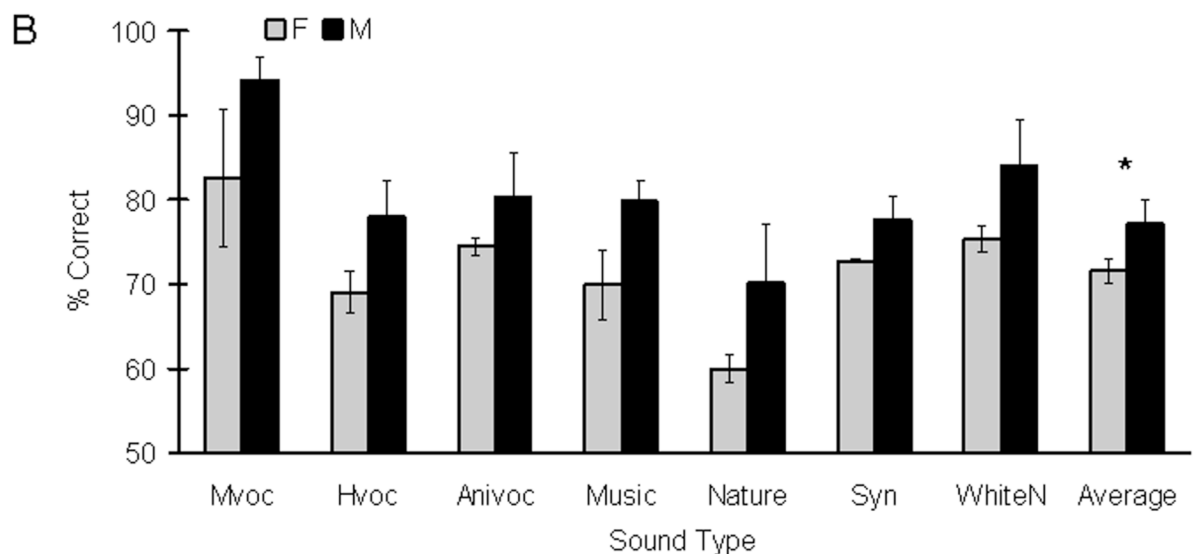
**Figure 1.**

Schematic diagram depicting the auditory delayed matching-to-sample task. Daily sessions contain equal numbers of match and nonmatch trials. During match trials, the first sound, followed by a 5-second delay, was same as the second sound. The correct response was a touch (go-response) and the animal was then rewarded. During nonmatch trials the two sounds were different and also separated by a 5-second delay, and the correct response was to not touch. The correct no-go response was not rewarded, and thus the study utilized asymmetric reinforcement contingency. An erroneous touch response during nonmatch trials resulted in an extended inter-trial interval before the next trial started.

## Match Memory Performance in Exp. 1



## Nonmatch Memory Performance in Exp. 1

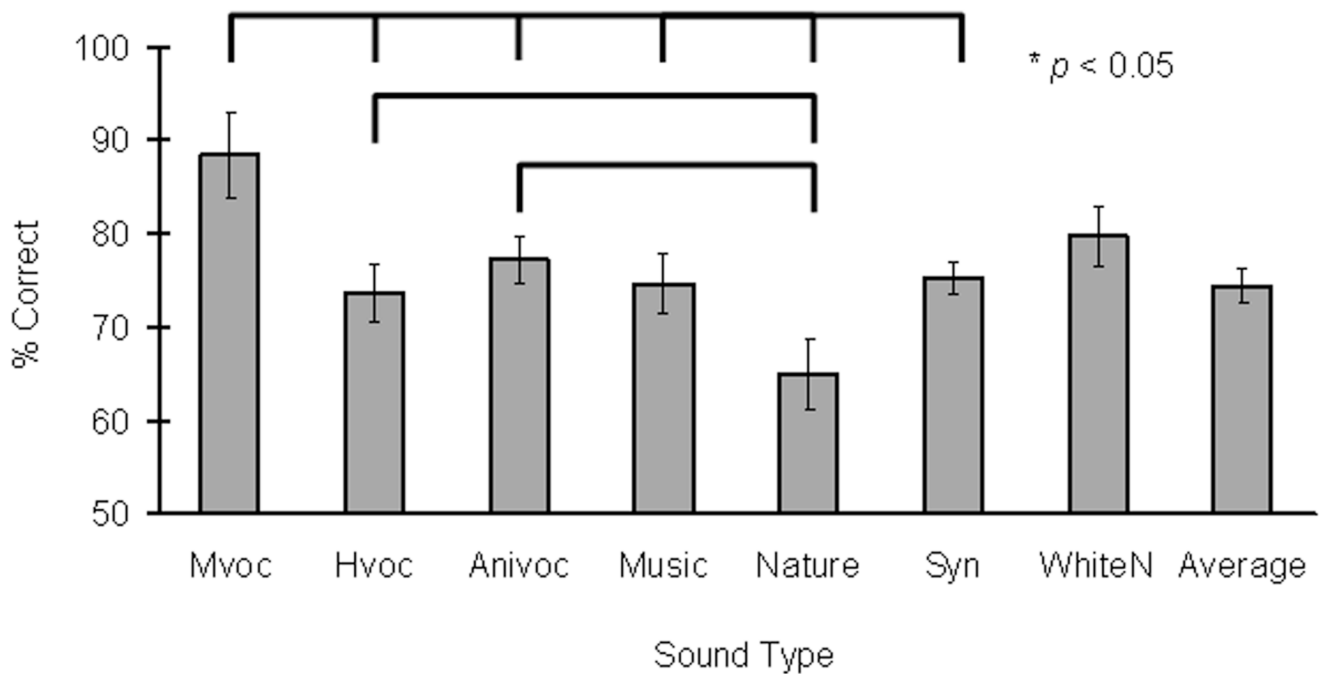
**Figure 2.**

Experiment 1: Gender influence on auditory memory performance of rhesus monkeys.

Auditory memory performance for match (A) and nonmatch trials (B) differs by gender. The graphs show average memory performance across seven sound types at fixed 5-second delays during match and nonmatch trials. Grey and black bars represent memory performance of female and male monkeys respectively. The asterisk indicated a significant performance difference between genders (A: repeated-measures ANOVAs,  $P < 0.05$ ; B: linear regression analysis,  $P < 0.005$ ). Males were always better than females in general. *Abbreviations:* monkey vocalizations (Mvoc), human vocalizations (Hvoc), animal vocalizations (Anivoc), music clips

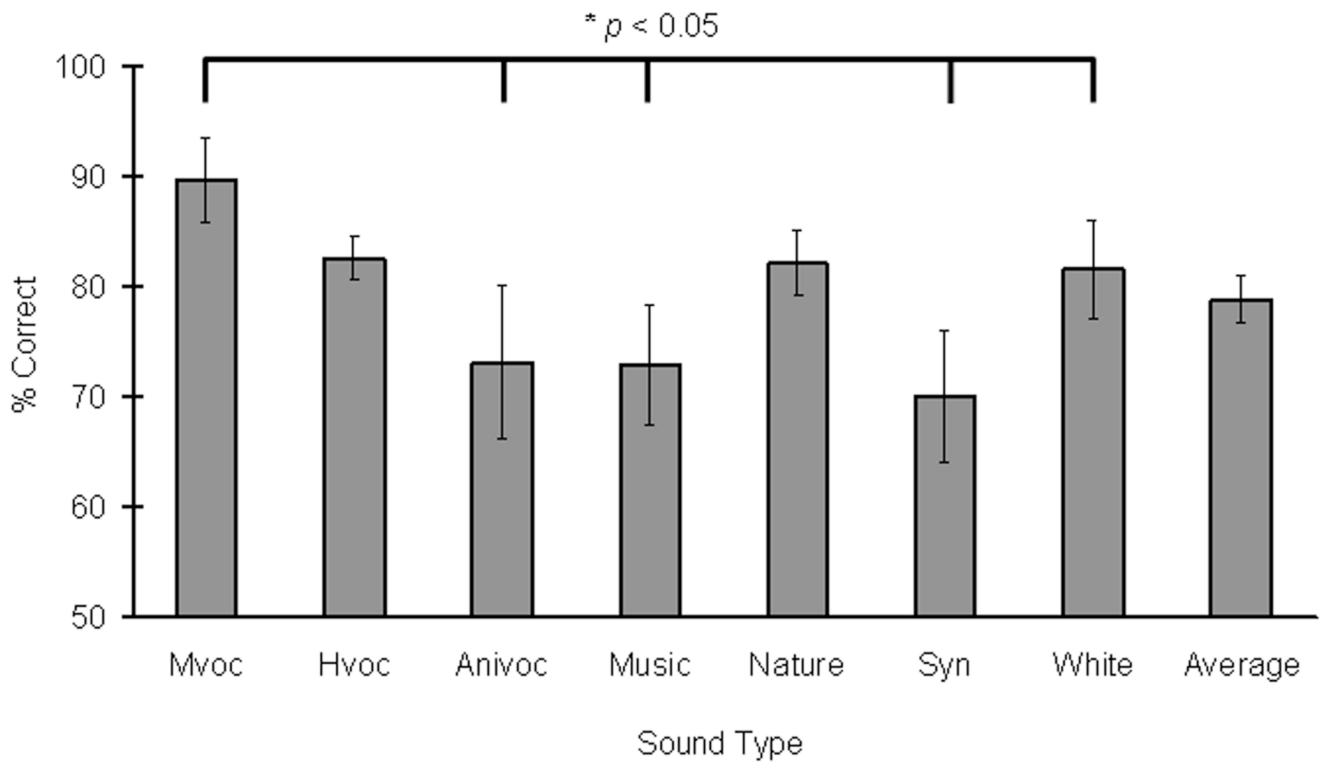


(Music), sounds of natural phenomena (Nature), synthesized clips (Syn) and band-passed white noises (WhiteN).



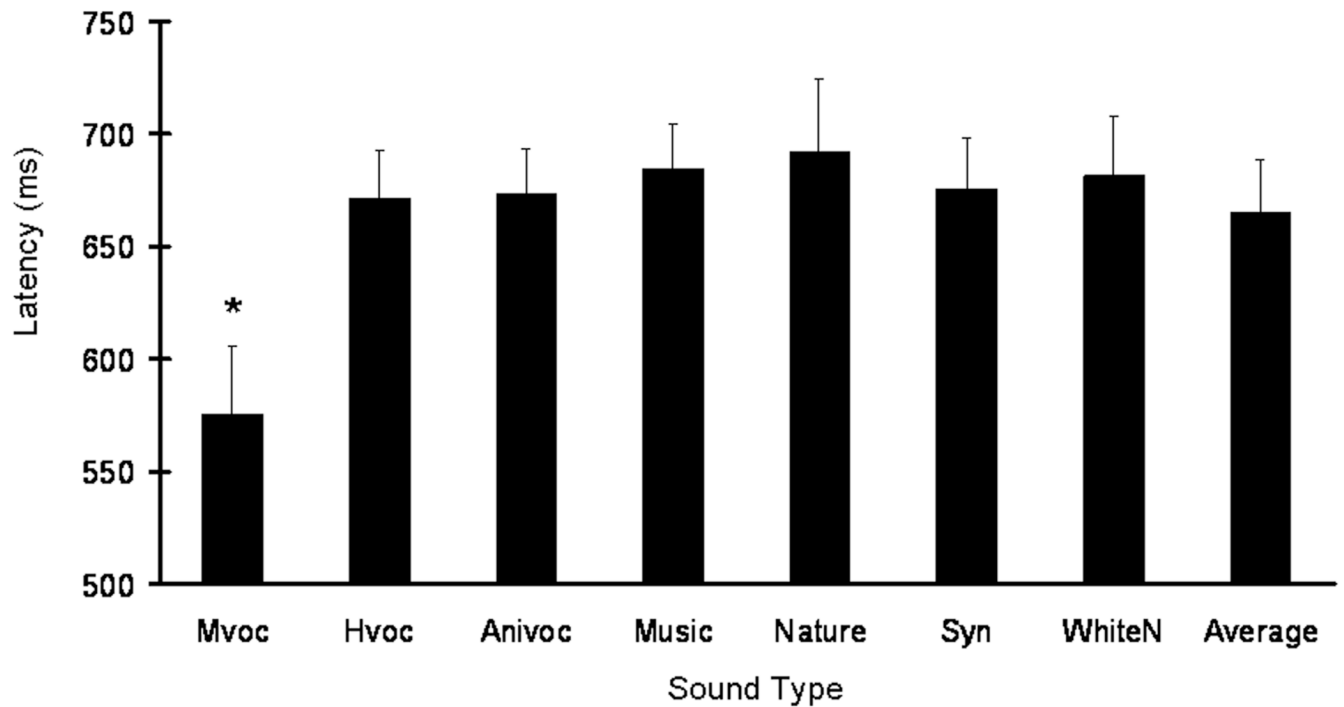
**Figure 3.**

Auditory memory performance for distinguishing two different sounds depends on sound types. The graph shows average performance at fixed 5-second delays during nonmatch trials when specified sound type was presented as the second sound. Asterisks and brackets indicated significant performance difference between two sound types (paired-sample *t*-tests, \*  $p < 0.05$ ). Monkey vocalizations (Mvoc) yielded better memory performance than those associated with five out of the six sound types (Hvoc, Anivoc, Music, Nature, and Syn). Human and animal vocalizations (Hvoc and Anivoc) also yielded better memory performance than those associated with natural sounds. Note that fewer samples in monkey vocalizations (Mvoc,  $n = 14$ ), natural sounds (Nature,  $n = 28$ ) and white noises (WhiteN,  $n = 21$ ) were not always associated with better recognition memory in general and number of stimuli per type did not account for the current findings.



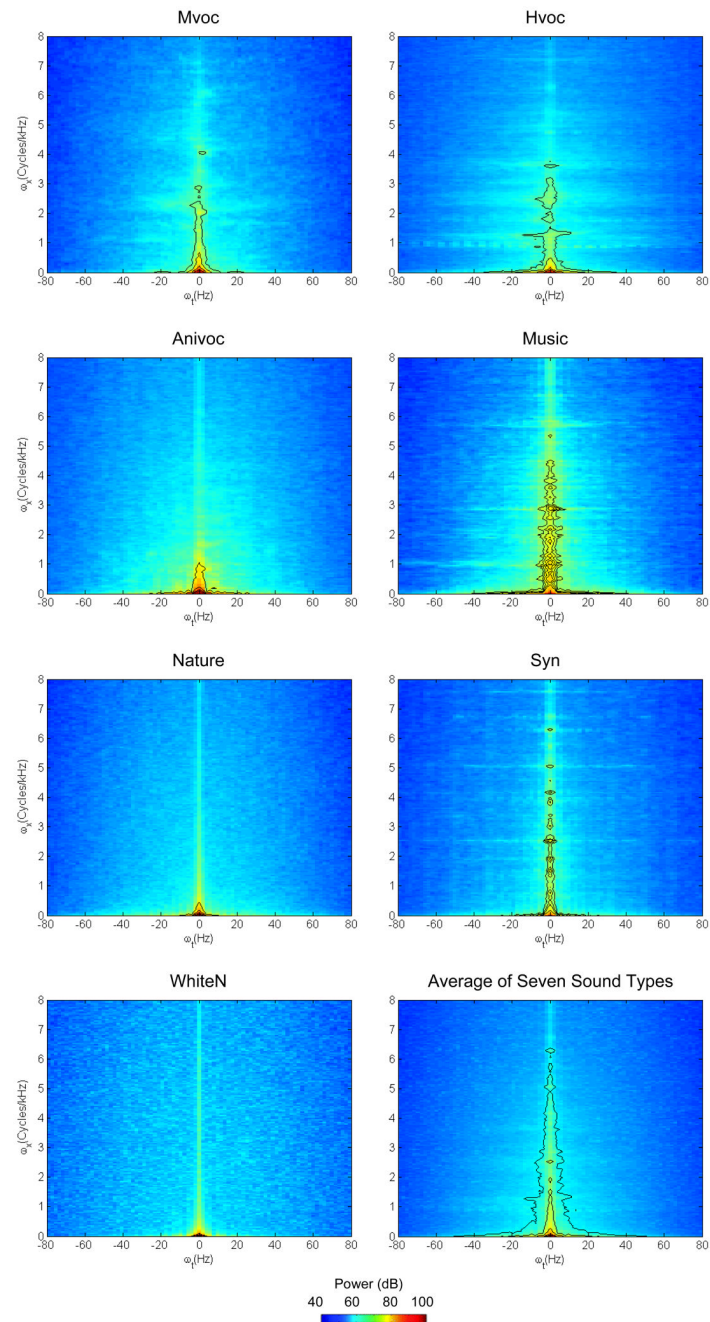
**Figure 4.**

Experiment 2: Auditory memory performance in a trial-unique delayed matching-to-sample task. Auditory memory performance for distinguishing two different sounds based on sound types. The graph shows average nonmatch performance at fixed 5-second delays when sound type was presented as the second sound. Asterisks and brackets indicated a significant performance difference between two sound types (paired-sample *t*-tests, \*  $p < 0.05$ ). Monkey vocalizations (Mvoc) yielded better memory performance than those associated with animal vocalizations (Anivoc), music clips (Music), synthesized clips (Syn), or white noises (WhiteN). Auditory memory performance of monkeys was improved accordingly when a given sound type as the first sound (sample) was compared against monkey vocalizations as the second sound (test).



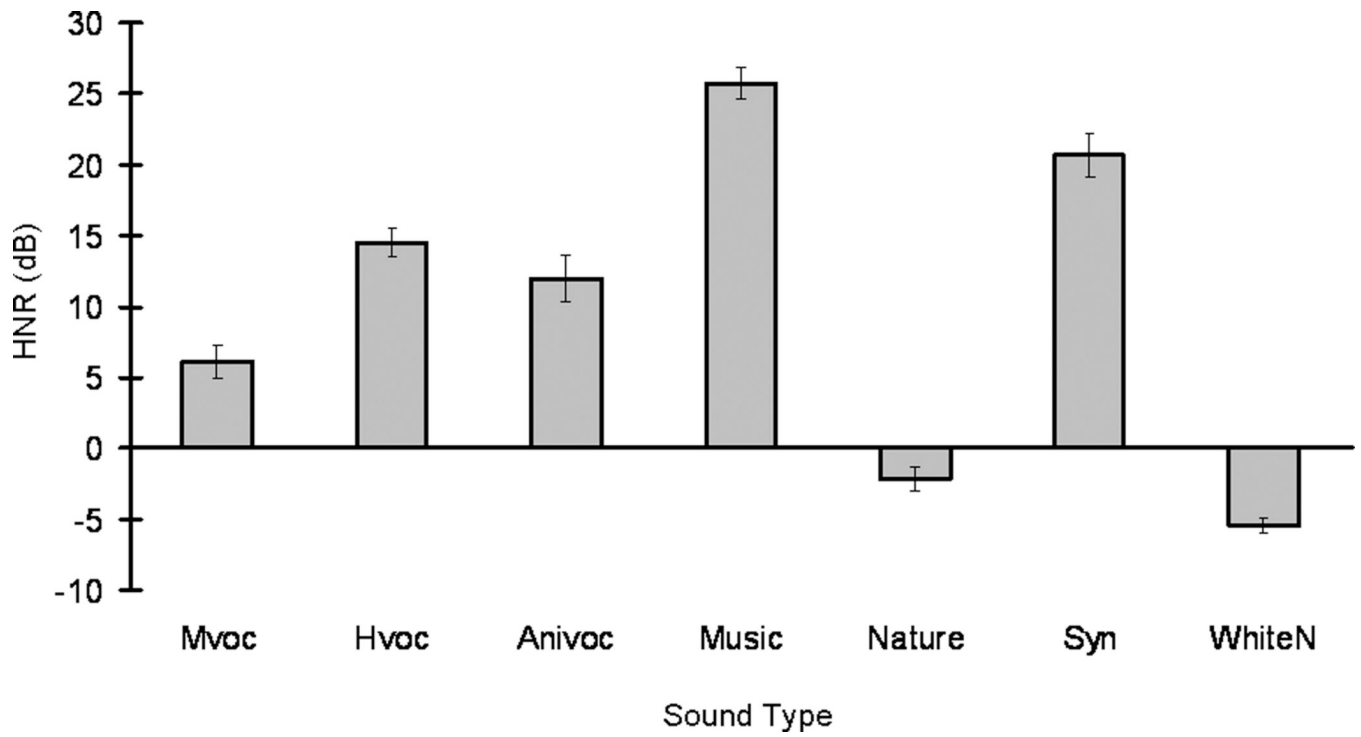
**Figure 5.**

Effects of sound type on response latencies during auditory memory performance of rhesus monkeys during Experiment 2. The graph illustrates average response latencies when monkeys showed correct go-responses for match trials. An asterisk indicated that monkeys responded faster to monkey vocalizations than any other sound type during match trials (paired-sample *t*-tests, \*  $p < 0.05$ ). Response latencies were similar across the seven sound types during nonmatch trials. The results suggest that monkeys' preference on their own species-specific sounds concomitantly influences both auditory memory performance and its respective motor expression.



**Figure 6.** Modulation spectra of seven sound types in Experiment 2. Power density is indicated by color using a decibel (dB) scale, with red showing the spectrotemporal modulations with the most energy. The x-axis and y-axis represent the frequency of the temporal modulations ( $\omega_t$  or cycles/Hz) and spectral modulations ( $\omega_f$  or Hz) respectively. Black lines in each spectrum are contour lines showing 50–90% of the power.





**Figure 7.** Harmonic-to-noise ratios of seven sound types in Experiment 2. Values of harmonic-to-noise ratio (HNR) indicate degree of acoustic periodicity for each sound type. A high and positive ratio value for a sound type indicates more acoustic energy for that sound type devoted to harmonics over time.

Individual memory performance on the auditory delayed matching-to-sample task in Experiment 1. Percentage correct of memory performance for males ( $n = 3$ ) and females ( $n = 3$ ) was shown (mean  $\pm$  standard error) on match and nonmatch trials respectively. The results were based on 40 sessions of data per monkey subject.

**Table 1**

Gender	Subject	Percentage Correct			
		Match Trials		Nonmatch Trials	
		Mean	S.E.	Mean	S.E.
Female	1	83.40	2.20	73.72	2.39
	2	81.50	1.71	71.97	1.96
	3	81.50	1.79	68.90	4.00
Male	1	85.31	1.59	80.16	3.11
	2	97.71	0.50	71.46	3.52
	3	84.61	1.63	79.71	3.32