Spring 2011

Behavioral and neural correlates of auditory encoding and memory functions in Rhesus Macaques

Chi-Wing Ng
University of Iowa

Recommended Citation

Ng, Chi-Wing. "Behavioral and neural correlates of auditory encoding and memory functions in Rhesus Macaques." PhD (Doctor of Philosophy) thesis, University of Iowa, 2011.
http://ir.uiowa.edu/etd/1041.

BEHAVIORAL AND NEURAL CORRELATES OF AUDITORY ENCODING AND MEMORY FUNCTIONS IN RHESUS MACAQUES

by

Chi-Wing Ng

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Associate Professor Amy Poremba

ABSTRACT

Auditory recognition memory in non-human primates is not well understood. Monkeys have difficulty acquiring auditory memory tasks, and limited capability maintaining auditory information over memory delays, relative to studies of visual memory. Neural substrates of auditory discrimination and recognition memory depend on superior temporal gyrus (STG), instead of rhinal cortex necessary for visual memory (Fritz et al., 2005). The current project assessed behavioral and neural correlates of auditory processing and memory function in monkeys, particularly focusing on the dorsal temporal pole (dTP), the rostral portion of STG. Chapter 2 examined recognition memory of monkeys under influences of various sound types. In a delayed matching-to-sample (DMTS) task, rhesus monkeys were trained to determine if two sounds, separated by a 5-second delay, were same (match trials) or different (nonmatch trials). Results demonstrated monkey vocalizations served as better cues than other sound types for auditory memory performance. Memory improvements may be due to familiarity and biological significance of con-specific sounds, analogous to using facial stimuli during visual tasks. Chapter 3 examined neuronal activity of dTP, when two monkeys performed an auditory DTMS task and listened to sound stimuli. Population encoding of sample stimuli in dTP was closely associated with memory accuracy. Moreover, a suppression effect on identical sounds was present, similar to processing in the ventral visual processing stream, inferior temporal cortex (ITC) and ventral temporal pole (vTP). Delay-related activity of dTP was weak, limited and short-lived, in contrast to visual studies reporting sustained activity over memory delays in ITC, vTP and prefrontal cortex. The findings provide preliminary evidence on why monkeys show limited memory capability, compared to visual memory, for auditory information. Neurons of dTP were sound-selective, and mainly evoked by one to four discrete stimuli only. Sound types and simple acoustic properties of sound stimuli cannot completely account for response profiles of dTP neurons. The findings suggest dTP is a higher order auditory area, and receives information from various auditory areas along STG. Dorsal temporal pole fits into proposals of neural networks for auditory processing, in which a hierarchical organization of information flow exists within the primate auditory nervous system.

Abstract Approved: ____________________________________
Thesis Supervisor

____
Title and Department

____
Date

BEHAVIORAL AND NEURAL CORRELATES OF AUDITORY ENCODING AND MEMORY FUNCTIONS IN RHESUS MACAQUES

by

Chi-Wing Ng

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Associate Professor Amy Poremba

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Chi-Wing Ng

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the May 2011 graduation.

Thesis Committee: Amy Poremba, Thesis Supervisor

John H. Freeman

Daniel T. Tranel

Bob McMurray

Paul J. Abbas

Dedicated to my mother and father. The sacrifice of my mother actualizes my journey to higher knowledge, as well as the current dissertation. For you.

ACKNOWLEDGMENTS

I would like to thank my mentor and teacher, Dr. Amy Poremba, for leading me into behavioral neurosciences and opening my horizon for auditory neurophysiology in non-human primates. In the past six and a half years, her guidance has shown me the ways of scientific thinking, design of experiments, academic writing and speech. I thank her because of her tolerance and patience during my difficult moments in science. I would also like to thank my long-time friend and colleague, Bethany, as we established the beginning of a laboratory and worked together to accomplish an abundance of experimental projects. I also thank Ryan and James for providing support and assistance in my work, especially Ryan for making my life experiences in America one to remember. Amy Opheim helped me develop proficiency in English writing by helping me write in a more precise and reader-friendly manner. I thank Dr. Yasuo Nagasaka for giving me guidance on developing MATLAB scripts in data processing and analysis. I also thank all of the research assistants, past and present, for making daily progress in the laboratory. I would like to thank Dr. Mortimer Mishkin, Richard Saunders and Megan Malloy for their early work of monkey training and feedback on current projects.

I especially thank the committee members of Dr. John Freeman, Dr. Bob McMurray, Dr. Daniel Tranel, and Dr. Paul Abbas, for their suggestion and guidance to my dissertation.

Lastly, I am in debt to my father’s support, patience and tolerance as I am no longer available to him after my mother passed away eleven years ago.

ABSTRACT

Auditory recognition memory in non-human primates is not well understood. Monkeys have difficulty acquiring auditory memory tasks, and limited capability maintaining auditory information over memory delays, relative to studies of visual memory. Neural substrates of auditory discrimination and recognition memory depend on superior temporal gyrus (STG), instead of rhinal cortex necessary for visual memory (Fritz et al., 2005). The current project assessed behavioral and neural correlates of auditory processing and memory function in monkeys, particularly focusing on the dorsal temporal pole (dTP), the rostral portion of STG. Chapter 2 examined recognition memory of monkeys under influences of various sound types. In a delayed matching-to-sample (DMTS) task, rhesus monkeys were trained to determine if two sounds, separated by a 5-second delay, were same (match trials) or different (nonmatch trials). Results demonstrated monkey vocalizations served as better cues than other sound types for auditory memory performance. Memory improvements may be due to familiarity and biological significance of con-specific sounds, analogous to using facial stimuli during visual tasks. Chapter 3 examined neuronal activity of dTP, when two monkeys performed an auditory DTMS task and listened to sound stimuli. Population encoding of sample stimuli in dTP was closely associated with memory accuracy. Moreover, a suppression effect on identical sounds was present, similar to processing in the ventral visual processing stream, inferior temporal cortex (ITC) and ventral temporal pole (vTP). Delay-related activity of dTP was weak, limited and short-lived, in contrast to visual studies reporting sustained activity over memory delays in ITC, vTP and prefrontal cortex. The findings provide preliminary evidence on why monkeys show limited memory capability, compared to visual memory, for auditory information. Neurons of dTP were sound-selective, and mainly evoked by one to four discrete stimuli only. Sound types and simple acoustic properties of sound stimuli cannot completely account for response profiles of dTP neurons. The findings suggest dTP is a higher order auditory area, and receives information from various auditory areas along STG. Dorsal temporal pole fits into proposals of neural networks for auditory processing, in which a hierarchical organization of information flow exists within the primate auditory nervous system.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF ABBREVIATIONS

Anivoc	Animal vocalization
dB	Decibel
DMTS	Delayed matching to-sample
dTP	Dorsal temporal pole
fSTS	Face area of superior temporal sulcus
FR	Firing rate
HNR	Harmonic-to-noise ratio
Hvoc	Human vocalization
ITC	Inferior temporal cortex
ISI	Inter-stimulus interval
ITI	Inter-trial interval
lPFC	Lateral prefrontal cortex
MC	Match correct
MI	Match incorrect
MRI	Magnetic resonance imaging
Mvoc	Monkey vocalization
NC	Nonmatch correct
NI	Nonmatch incorrect
PFC	Prefrontal cortex
A1	Primary auditory core
V1	Primary visual cortex
R1, R2, R3	Response periods 1, 2, or 3
rSTP	Rostral superior temporal plane
rSTG	Rostral superior temporal gyrus
RT	Rostral portion of A1
SPL	Sound pressure level
STG	Superior temporal gyrus
Syn	Synthesized clip
TP	Temporal pole
vTP	Ventral temporal pole
WhiteN	White noise

CHAPTER 1. INTRODUCTION

Animal communication requires a series of multi-level information processing procedures that pertain to a signal receiver and its surrounding environment. Sensory detection alerts the receiver about the presence of external stimuli, and encoding and processing stimuli help to distinguish relevant information from irrelevant background noises. Analyses of these external stimuli identify the nature of external cues and facilitate the receiver to derive significance or meaning (e.g., identity, gender, body size, and kinship) embedded inside these particular cues. This process in animal communication may facilitate the execution of appropriate decision-making processes and behavioral actions in regard to a particular environmental situation, e.g., courtship and mating between two sexes and fight or flight responses induced by a potential predator/competitor. Multi-level signal processing varies by stimulus quality, salience, and complexity of information processing and can occur successively and simultaneously within the nervous system.

Beyond signal communication, auditory stimulus recognition is essential for animals to manipulate cues for advanced usages such as attention, learning, memory, and goal-directed behavior. Auditory learning and memory processes have been examined extensively in rodents, birds, and humans, from experiments such as classic Pavlovian conditioning with tone-shock associations (Davis, 1989; LeDoux, 1993; Weinberger, 1998) and the development of song learning in birds (Konishi, 1985; Nottebohm et al., 1990; Marler, 2004) to a more complex and linguistic use of speech/non-speech sounds with a capability for storage and retrieval (Belin and Zatorre, 2000; Hickok and Poeppel, 2007). Auditory sensitivity, pattern discrimination and memory performance for comparative, psychoacoustic studies are often examined in non-human primates. It is fascinating that non-human primates show poor acquisition and retention of auditory information in contrast to their excellence with visual information. There is no doubt that a more in-depth investigation is required to further our understanding of how non-human primates use and comprehend auditory information and the neural encoding underlying auditory-linked behaviors.

Complexity in Auditory Recognition Memory of Non-Human Primates

Interactions between Sound Type and Memory Expression

One example of poor auditory learning and memory in monkeys is illustrated in the delayed matching-to-sample (DMTS) task. This DMTS task requires subjects to discriminate and retain a representation of auditory information across time delays. Monkeys are trained to distinguish whether two presented sounds, separated by a time delay, are the same (match trials) or different (nonmatch trials). Then, they are trained to make discrete responses to match and nonmatch trials for rewards respectively, such as left/right or go/no-go response contingencies. Rhesus macaques (*Macaca mulatta*) generally demonstrate a slow acquisition in learning the DMTS task for auditory stimuli. Wright and colleagues (Wright et al., 1990; Wright, 1998, 1999) concluded that monkeys with prior DMTS training of visual stimuli fail to transfer their learning to a similar task that uses auditory stimuli. The use of a series of training parameters to help monkeys attend to the task took researchers years of training for two monkeys (Wright, 1998, 1999; Wright et al., 2000) to become successfully trained and tested. A more refined approach in training monkeys in this auditory memory task usually takes 15,000 trials on average to reach performance criteria at 5-second memory delays (Fritz et al., 2005), while a similar one (delayed matching- and nonmatching-to-sample task), which uses visual or tactile stimuli, and takes them a few hundred trials to learn at short delays such as 10 minutes (Murray and Mishkin 1998; Buffalo et al. 1999; Zola et al. 2000).

Monkeys also sometimes exhibit significant discrepancies on memory performance when using auditory, visual or tactile information. Major findings assessing short-term visual recognition memory of monkeys employ training methods of fixed, incremental memory delay intervals. The temporal duration of delay intervals is fixed and then lengthened once memory performance of monkeys reached criterion. For instance, the delay interval for testing sessions starts with 10 seconds (s) fixed, then 30 s, 60 s, 120 s and so on (e.g. Murray and Mishkin 1998). Instead, variable delay training (i.e., a mix of short and long delay intervals) is used to encourage monkeys to higher performance levels on the auditory DMTS task at longer delay intervals (e.g., Fritz et al, 2005). Differential setup of memory delays during testing implies dissimilar expressions of recognition memory between sensory modalities. Monkeys poorly maintain auditory information for recognition at time delays longer than 35 s, but they can remarkably hold visual or tactile information at time delays beyond 10 minutes. Do certain sound types serve as better acoustic cues to monkeys for improving memory performance across time delays?

Researchers have examined some of the possible reasons why non-human primates show poor auditory performance in learning and memory while humans exhibit remarkable, efficient use of acoustic stimuli. Early studies examined the extent to which humans and animals, particularly monkeys, share similar psychoacoustic capacities in auditory processing in order to establish animal models of speech-like, communicative signal perception (Kuhl, 1979; Sinnott et al., 1985). These experiments utilized a variant of the *AX* technique, which is now referred to as the same/different task for auditory discrimination. In the same/different task, an animal is presented with a standard stimulus repeatedly and is always reinforced for behavioral responses in the presence of a stimulus change. Humans are superior in frequency discrimination of pure tones at frequencies below 4000 Hertz (Hz), as compared to birds (Sinnott et al., 1980) and monkeys (Sinnott et al., 1985, 1987; Prosen et al., 1990). However, monkeys express similar performance to humans in frequency discrimination when using speech-like stimuli, such as the formant frequency of vowels (Sinnott and Kreiter, 1991; Sommers et al., 1992; LePrell et al., 2001). Additionally, chinchillas (Kuhl and Miller, 1978) and monkeys (Waters and Wilson, 1976; Sinnott and Adams, 1987; Sinnott and Kreiter, 1991) are able to categorize phonemes that differ in voice-onset times similarly to humans, such as the phoneme continuums of ba/pa or da/ta. These studies imply that sounds of a dynamic nature serve as better cues than pure tones during auditory discrimination of monkeys. Converging evidence potentially suggests that a general, common auditory system in processing complex, dynamic acoustic stimuli is present for both animals and humans.

However, humans and animals differ in their attention to the acoustic features of sounds. One of the common features examined extensively in human and animal studies is frequency modulations (FM). FM are important cues in human speech and are also valuable for signal communication in various species like song birds (Marler, 2004), bats (Ulanovsky and Moss, 2008), and monkeys (Moody et al., 1986; May et al., 1989).

Perception of frequency pattern or direction of frequency change is often used to examine how humans and animals rely on characteristics of FM to solve discrimination problems. Humans attend preferably to frequency contours of sound in recognizing and discriminating auditory patterns, e.g., directions of frequency change (ascending and descending FM) and tonal/melodic sequences of sounds (Dowling, 1978, 1991; Trehub et al., 1987; Schouten and Pols, 1989; Hauser and McDermott, 2003). Animals predominantly rely on absolute pitch differences or local cues of sound sequences to discriminate sounds (monkeys and rats: D'Amato and Salmon, 1982; D'Amato and Colombo, 1988; ferrets: Yin et al., 2010; birds: Hulse et al., 1984; Cynx et al., 1986; MacDougall-Shackleton and Hulse, 1996; Mongolian gerbils: Wetzel et al., 1998; dolphins: Ralston and Herman, 1995). Animals often show evidence of using frequency contours for discrimination only during species-specific sounds of animal subjects (monkeys: Zoloth et al., 1979; Hauser, 1998; Ghazanfar et al., 2007; birds: Hurly et al., 1992; Weisman et al., 1994; MacDougall-Shackleton and Hulse, 1996). In general, the use of absolute sound cues (e.g., pitch) rather than relative features of sounds are strongly shown in animals with regard to humans. Birds (e.g., pigeons and zebra finches) show superior frequency-range discrimination of pure tones to humans, while the latter species preferably attends to melodic patterns of sounds without noticing differences in pitch intervals within the sound segments (Friedrich et al., 2007). Interpretations of the FM discrimination performance of monkeys may suggest possible reasons why non-human primates generally express poor auditory recognition memory compared to humans. Prior studies (e.g. Wright, 1998, 1999; Fritz et al., 2005) primarily presented non-biological sounds (e.g., pure tones and environmental sounds, etc.), but they also presented some heterospecific vocalizations (e.g., human speech words and dog barks) and a small amount of monkey vocalizations as well. However, acoustic stimuli used by these studies may not accurately assess auditory memory performance of monkeys.

There is strong evidence to propose that species-specific sounds generally lead to faster acquisition and sometimes better discrimination performance for birds and non-human primates. Japanese macaques learn faster than control monkeys of a different species when discriminating conspecific calls of Japanese macaques, even though both groups achieve similar levels of accurate performance (Zoloth et al., 1979; Petersen et al., 1984). Zebra finches similarly learn faster to discriminate their own songs from other conspecific songs of the same species, or songs from different avian species (Cynx and Nottebohm, 1992; Cynx, 1993). Discrimination performance is sometimes better when these animals distinguish conspecific sounds from pure tones or white noises. Species-specific vocalizations are salient stimuli for communication among individual members and the surrounding environment for living organisms (Fitch, 2000; Ghazanfar and Hauser, 2001). Like humans, non-human primates attend to distinctive acoustic cues of conspecific vocalizations for efficient processing when compared to heterospecific sounds from other animal species (Hauser, 1998; Gifford et al., 2003; Rendall, 2003; Hienz et al., 2004; Fitch and Fritz, 2006). These psychoacoustic findings are also compatible with results obtained from imaging and neurophysiological studies. “Voice-sensitive” and “vocalization-sensitive” areas are shown in secondary auditory regions, the 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; Andics et al., 2010) and non-human primates (Gil-da-Costa et al., 2004; Poremba et al., 2004; Petkov et al., 2008; Remedios et al., 2009: Kikuchi et al., 2010). 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). One possible explanation for differences in memory performance across sound types is that some sounds may be more readily processed or encoded by the brain, perhaps due to familiarity. In humans, visual perception and memory performance are enhanced using faces, pictures, and words that are more efficiently processed and categorized in human cognition (Seifert, 1997; Amrhein et al., 2002; Bulthoff and Newell, 2006). This suggests that species-specific sounds may exert functional advantages in auditory learning and memory of monkeys over other sound types. In order to properly assess short-term auditory recognition memory of monkeys, a broad range of sound types as well as species-specific vocalizations are necessary to ascertain the capabilities of monkeys for processing and retaining auditory information across memory delays.

Problems Associated with Studying Auditory Processing of Higher-Order Auditory Brain Regions

The investigation of audition and its neural mechanisms has benefited from experimental techniques and methodologies employed during studies of the visual system. Numerous studies have provided information on how our visual system detects and processes incoming visual stimuli, e.g., retinotopic mapping of the primary and secondary visual cortices (Hubel and Wiesel, 1968; Desimone and Gross, 1979) and analyses of stimulus feature and object identity (Desimone et al., 1984; Tanaka, 1993). The proposed neural network of visual processing also influences attention (Luck and

Vogel, 1997), learning, and memory (Fuster and Jervey, 1982; Miller et al., 1991). Psychophysically, there are particularly salient stimuli that are sent to the signal receiver for efficient processing in monkeys (faces: Hauser, 1993; Kobatake and Tanaka, 1994) and humans (words, pictures and faces: Seifert 1997; Amrhein et al. 2002). The auditory system of mammals also shares some neural and functional characteristics of the visual system, such as cochleotopic mapping of sound frequencies and analysis of spectrotemporal features of sounds in the primary and secondary auditory cortices (Merzenich and Brugge, 1973; Rauschecker et al., 1995; Recanzone, 2000; Tian et al., 2001). Empirical studies of auditory behavior and neurophysiology often employ experimental metaphors used in visual studies to address questions about auditory processing of monkeys and humans. However, there are still crucial differences between auditory and visual information which may result in significant distinctions in neuronal and behavioral manifestations of auditory processing.

Compared to visual stimuli, auditory information is generally more dynamic and involves fluctuation along several physical dimensions simultaneously. Visual objects are typically static pictures or figures presented on a screen consisting of different colors, shapes, and spatial arrangements of individual units on a visual presentation. Contrarily, sounds are a dynamic series of varying frequencies, phases, and intensities across time. Thus, dynamic visual stimuli (e.g., moving objects) serve as a better analogy to sound dynamics. Sounds may or may not include spatial components regarding their locations or movements. Distinct segments of a given sound may provide significance to the signal receiver; for example, humans are more attentive to consonances, vowels, and musical segments and animals are capable of making favorable distinctions between hetero/conspecific vocalizations. Thus, temporal relationship among sound frequencies is especially important for auditory processing. Increasing or decreasing the temporal duration of a sound remarkably influences neuronal profiles in the primary auditory cortex compared to the unaltered version of the same sound (Wang et al., 1995; Nagarajan et al., 2002; Wang, 2007). In order to understand principles of auditory processing for unique, dynamic structures of auditory information, a wide array of auditory stimuli should be utilized, ranging from simple acoustic cues (e.g., pure tones) to complex, perceptually meaningful sound segments (e.g., species-specific vocalizations).

Experimental design and behavioral setup directly affect how subjects, as well as the corresponding nervous system, process and manipulate auditory information according to task demands. Historically, the majority of findings about the principles of auditory processing are predominantly based on anesthetized animals. The main concern with this technique is that neurons are more difficult to evoke from anesthetized animals and event-related spike activities can become affected by anesthesia (Cheung et al., 2001). Past recording studies primarily revealed short-lived, rapid onset responses of auditory neurons within primary auditory regions in anesthetized monkeys, rather than continuously discharging responses to long-duration sounds (deCharms and Merzenich, 1996). In awake monkeys, long-lived and sustained responses to simple and complex sounds are shown within the primary auditory core region (Wang et al., 2005). A large proportion of these sustained-firing neurons often illustrate greater spectral and temporal complexity to their preferred stimuli than pure tones or broadband noises. Top-down influence, accompanied by stimulus-driven effects, mediates response characteristics of auditory cortical areas when processing acoustic information concerning its relationship to the signal receivers. In addition, task demand, subjects' attention, and experience with sounds significantly alter response patterns of auditory-responsive brain regions (Wang, 2000; Cohen et al, 2004; Romanski et al., 2005; Blake et al., 2006). In human imaging studies, frontal regions are more likely to be activated in conditions in which the task required subjects to actively attend to sound stimuli than in those conditions where subjects passively listened to the stimuli (Cohen et al., 2004). Similarly, owl monkeys exhibit enhanced auditory response strength and stimulus selectivity to auditory stimuli after discrimination learning with the same stimulus set (Blake et al., 2006). The converging implication is that a behavioral paradigm requiring active attention to task-relevant sounds favorably activates some cell populations, which may be involved in encoding and analyzing identity and/or behavioral relevance of sounds with respect to the experimental testing.

Recent imaging and electrophysiological studies tend to measure auditory spike activity in awake animals and produce fruitful evidence about auditory processing and its computational mechanisms inside the auditory cortical areas. Most of these studies, however, employ a passive auditory environment with minimal attention on sounds presented throughout sessions (e.g., Gil-da-Costa et al., 2004; Rauschecker and Tian, 2004; Poremba et al., 2004; Romanski et al., 2005; Wang et al., 2005; Cohen et al, 2007; Petkov et al., 2008; Sadagopan and Wang, 2008). Auditory stimuli that are presented to subjects do not necessarily carry operational relevance to the experiment itself. The activity of auditory neurons may be merely stimulus-driven and can manifest based on quality and complexity of sounds. At times researchers optimize subjects' attention to sounds by instructing monkeys to fixate on a reference point (e.g., a light spot), and to make saccade responses to a target on a video screen (e.g., Romanski et al., 2005; Cohen et al, 2007). This procedure may inadvertently confound results that demonstrate considerable activity from visual and spatial processing in the monkey auditory system. It undoubtedly complicates the neural computational mechanisms of auditory cortical regions during sound analysis and identification. In order to understand how the brain generally encodes and processes sounds, it is necessary to employ a behavioral task coupled with task-relevant auditory information which requires subjects to accurately assess stimuli for memory performance. The auditory delayed matching-to-sample task serves as a compatible paradigm for examining neural mechanisms of learning and memory in monkeys, and it also allows for empirical comparisons to those in visual memory. It is particularly important to use complex acoustic stimuli in this paradigm rather than pure tones. These acoustic stimuli are necessary to evoke substantial spike activity within higher-order auditory brain regions of monkeys such as secondary auditory areas, superior temporal gyrus and temporal pole.

Temporal Pole: A Possible Area for Higher-Order, Non-Spatial Auditory Processing of Sounds

Dual streams of spatial and non-spatial processing have been revealed in the visual system of non-human primates and humans. The “where” and “what” pathways both originate in the primary visual cortex (V1) and diverge into dorsal and ventral streams (Mishkin et al., 1983). For the non-spatial aspects of stimulus attributes, projections from V1 to V4, inferior temporal cortex, and temporal pole signify the ventral “what” stream of visual processing. The occipitotemporal pathway is critically involved in identification of visual objects. Neurons in V4 and area TE selectively respond to the color and size of visual objects. In particular, neurons of ITC and temporal pole fire selectively to highly specific complex objects such as face and hand features (Desimone et al., 1984; Nakamura and Kubota, 1996). The dual streams of information eventually converge to the frontal cortex. The dorsal pathway terminates at the dorsolateral prefrontal cortex, while the ventral pathway terminates at the ventrolateral prefrontal cortex (Wilson et al., 1993; Romanski, 2004). Anatomical, electrophysiological, and metabolic mapping studies recently suggested that auditory information may be transmitted and divided into “what” and “where” pathways, which are analogous to the visual system (Figure 1; Rauschecker and Tian, 2000; Wang, 2000; Poremba et al., 2003; Rauschecker and Scott, 2009). Dual streams of information processing for auditory cues for spatial (“where”) and nonspatial (“what”) processing exist among regions of the prefrontal cortex, the primary auditory cortex and auditory association areas in cats (Lomber and Malhotra, 2008), monkeys (Romanski et al., 1999b; Rauschecker and Tian, 2000; Tian et al., 2001; Poremba et al., 2003) and humans (Alain et al., 2001; Maeder et al., 2001; Rämä et al., 2004; Barrett and Hall; 2006; Renier et al., 2009). Temporal pole, as shown to be a component of the anteroventral network along superior temporal gyrus (STG) for visual processing (Nakamura et al., 1994; Nakamura and Kubota, 1995, 1996) is similarly proposed to be important for auditory, non-spatial processing.

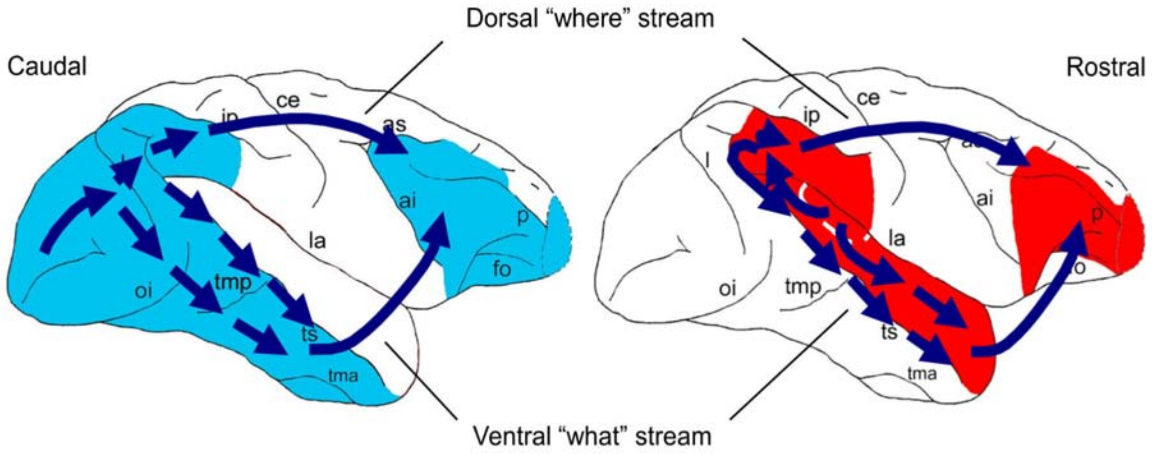


Figure 1. Schematic diagrams of dorsal and ventral streams of information processing in primate visual and auditory nervous systems. Blue and red colors represented visual and auditory responsive areas respectively, adapted from Poremba et al. (2003). Sulcus abbreviations: inferior occipital (io); superior temporal (ts); anterior and posterior of middle temporal (tma, tmp); lateral (la), intraparietal (ip); central (ce); superior arcuate (as); inferior arcuate (ai); principal (p); orbitofrontal (fo).

The temporal pole (TP) is situated on the tip of the temporal lobe, or the rostral end of STG (Figure 2). It has been suggested as a transition zone where information converges from the frontal lobe, temporal lobe, and limbic system. The dorsal temporal pole (dTP), including granular and dysgranular areas on the dorsal part of TP, is found to have extensive connections with surrounding auditory regions of STG, the frontal lobe, and the limbic system. Retrograde and anterograde tracing studies, which use horseradish peroxidase and fluorescent tracers, reveal that dTP predominantly receives auditory afferent inputs from the rostral part of STG and rostral parabelt region of the auditory association cortex (Moran et al., 1987; Hackett et al., 1998; Saleem et al., 2008; Reser et al., 2009). Retrograde-labeled cells present in medial pulvinar nucleus and the nucleus limitans and medial geniculate nucleus indicate that dTP receives direct thalamic inputs from the early stages of the auditory transmission pathway (Markowitsch et al., 1985; Moran et al., 1987; Yeterian and Pandya, 1989).

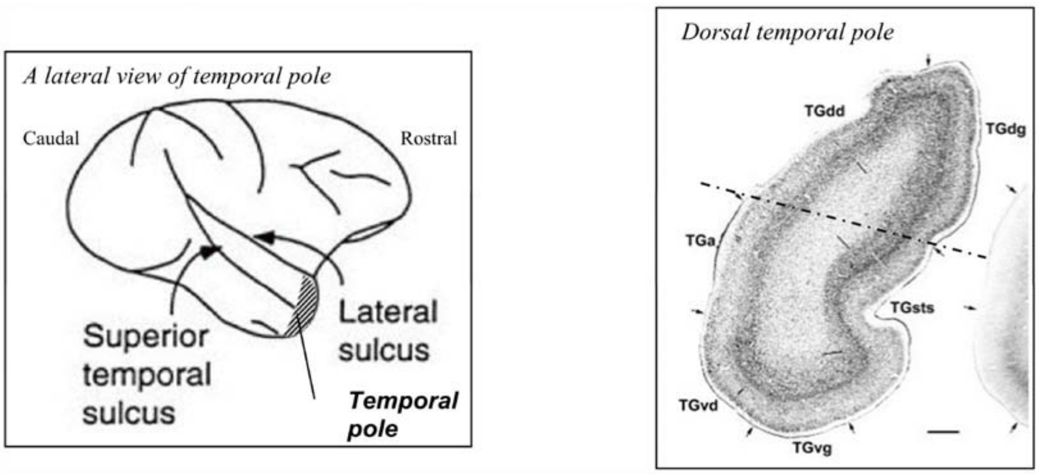


Figure 2. The lateral and coronal views of temporal pole. The area is divided into agranular (TGa), dysgranular (TGdd) and granular (TGdg) components, with dorsal and ventral regions separated by the STS (Kondo et al., 2003).

The dorsal temporal pole also has bidirectional connections with the frontal lobe, such as lateral, medial and orbital networks of the prefrontal cortex, including frontal pole (area 10m), inferior convexity (areas 12o, 12vl, 25), lateral orbital area 13 and anterior cingulate cortex (Barbas et al., 1999; Romanski et al., 1999a; Kondo et al., 2003). It receives numerous projections from the medial temporal lobe, including the entorhinal cortex (area 28), parahippocampal cortex (area TH and TF), anterior hippocampus and amygdala (Markowitsch et al., 1985, Moran et al., 1987; Suzuki and Amaral, 1994; Carmichael and Price, 1995; Kondo et al., 2003, 2005; Saleem et al., 2008). The most distinct anatomical and architectonic feature of TP is that the neuronal network of connection for the dorsal and ventral regions of the temporal pole (vTP) seems not to overlap. The auditory association areas connect to dTP while visual associative areas of the inferior temporal cortex connect to vTP (Seltzer and Pandya, 1978; Baylis et al., 1987; Moran et al., 1987). A similar trend of separate frontal connections to the dorsal and ventral parts of TP is also revealed. A strong bidirectional connection exists between the dTP, medial, and lateral orbital networks of the prefrontal cortex while heavy interconnection is present between the vTP and orbital network of the prefrontal cortex (Carmichael and Price, 1995; Kondo et al., 2003, 2005). Using metabolic mapping by the 2-deoxyglucose method, the respective functional segregation between visual and auditory processing at TP is further confirmed (Poremba et al., 2003). A distinct, non-overlapping network within dTP implies its unimodal specificity in auditory information processing. In short, the massive network of information flow within dTP may be able to support analysis and integration of different auditory information regarding attention, memory experiences and motivational significance to guide goal-directed behavior in non-human primates (Figure 3). Furthermore, the fact that dTP is anatomically located within the ventral "what" pathway provides hypotheses for its role in auditory stimulus identification and corresponding functions in learning and memory.

The area of dTP along with STG has been implicated in the auditory processing system of non-human primates. Early studies show that ablations of the auditory cortex along the superior temporal plane normally do not destroy basic ability of auditory perception. Monkeys with STG ablation are able to discriminate two tones in terms of frequency difference (Weiskrantz and Mishkin, 1958; Iversen and Mishkin, 1973).

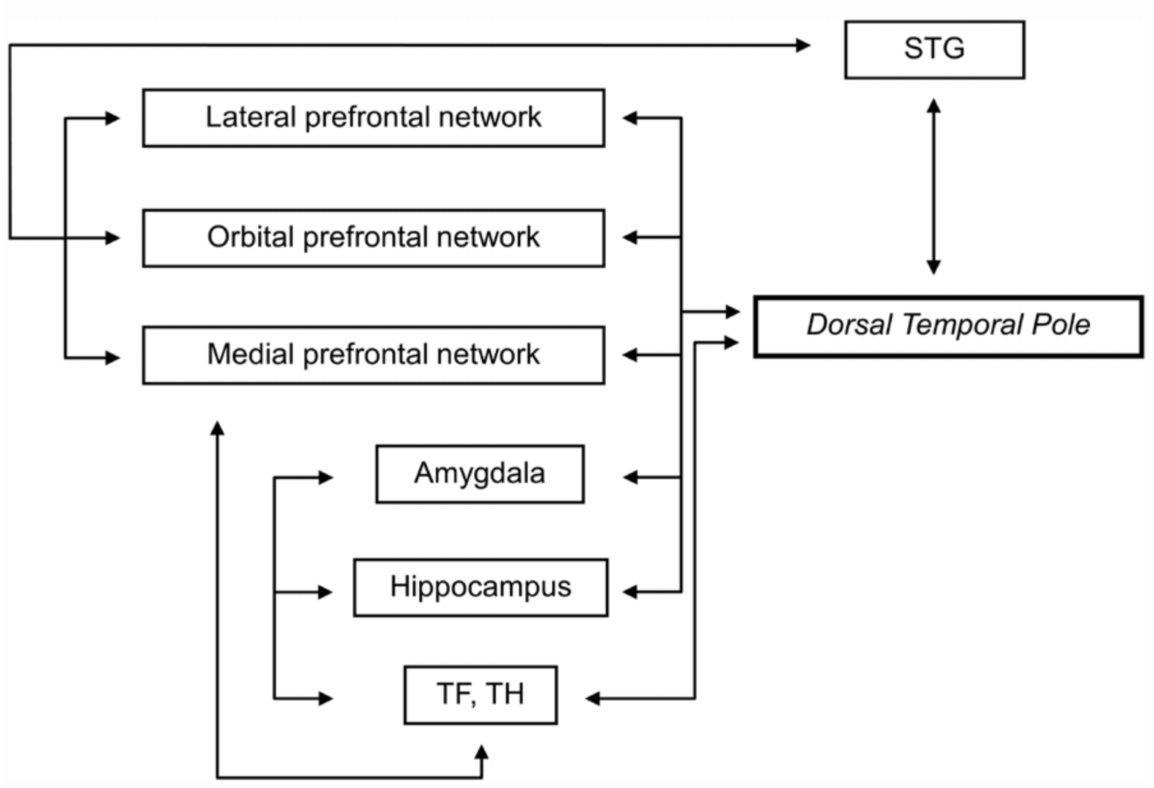


Figure 3. A schematic diagram depicting the neural network of dorsal temporal pole (dTP) and its connections with major cortical/subcortical brain regions. Superior temporal gyrus (STG) and the dorsal bank of superior temporal sulcus (STSd) provide major auditory inputs to dTP. Emotional and memory processing are linked to functional roles of dTP, where bidirectional connections are innervated with the medial and orbital prefrontal networks, and also the medial temporal lobe, including amygdala, hippocampus and parahippocampal cortices (TF and TH).

The network diagram is based on anatomical studies of monkeys: Suzuki and Amaral, 1994; Stefanacci et al., 1996; Hackett et al., 1999; Kondo et al., 2003, 2005; Höistad and Barbas, 2008; Saleem et al., 2008.

However, the corresponding damage impairs the ability to process complex sound stimuli, such as auditory discrimination based on pattern differences (Dewson et al., 1969, 1970). Bilateral lesions, including the association cortex and the rostral part of STG and small portions of dTP, severely impair the ability of monkeys to discriminate monkey vocalizations from non-monkey vocalizations or white noises. These monkeys can only respond to conspecific vocalizations as “go” stimuli for social reinforcement as if they are performing a basic auditory detection task. Similar lesions including dTP impair primates’ ability in discriminating subtypes of monkey vocalizations (Heffner and Heffner, 1984, 1986). Neural correlates of dTP for processing complex sounds, including species-specific vocalizations, are also established. Metabolic mapping by the 2-deoxyglucose method are utilized to examine activations in different brain areas when monkeys were passively listening to a variety of acoustic stimuli, including white noises, pure tones, FM sweeps, environmental sounds, musical phrases, human voices, monkey vocalizations and other animal sounds (Poremba et al., 2003). Dorsal temporal pole showed significantly greater local cerebral glucose utilization in the intact (hearing) hemisphere than in the acoustically isolated hemisphere. Similar results are also present in those brain regions with interconnections to dTP, for example, the rostral parabelt region, frontal pole, inferior convexity, orbitofrontal cortex, medial geniculate body, amygdala and parahippocampal cortex. In addition, the left dTP is associated with high brain activation evoked by species-specific monkey vocalizations or arrays of sounds containing those conspecific sounds, compared to simple and complex non-vocal sounds, background noise, human speech and phase-scrambled monkey vocalizations (Poremba et al., 2004). “Voice-sensitive” regions of the rostral STG and dTP revealed in human imaging studies suggest a common auditory mechanism for humans and non-human primates to process species-specific sounds accordingly (Belin et al., 2000, 2002; Petkov et al., 2009; Andics et al., 2010). Cumulative findings imply that the dorsal temporal pole is critical to encode and analyze complex sounds, particularly species-typical sounds relevant to animals. Its resultant contribution may support auditory learning and memory of monkeys when accurate discrimination and recognition of task-relevant sounds is required.

Empirical Questions and Outline for the Design of Experiments

The current experiments propose to investigate how sound types modulate auditory working and recognition memory of rhesus macaques, and how dTP encodes different sounds during passive listening and an auditory memory task. In Chapter 2, Experiment 1 aims to address whether a particular sound type would serve as a better acoustic cue to monkeys during discrimination and recognition in an auditory delayed matching-to-sample task. It evaluates whether auditory memory of monkeys is modulated by particular sound types. It is hypothesized that monkey vocalizations, species-typical sounds to the animals, would improve retention of auditory information compared to other sound types. In Chapter 3, Experiment 2 will examine the neural mechanisms of auditory processing in dorsal temporal pole. The recording study is coupled with passive listening sessions to evaluate how the spiking activity of dTP might represent a wide range of sound stimuli, and also to investigate an auditory DMTS task in which some of the previously heard sounds become task-relevant and the recognition of which becomes crucial to the attainment of short-term memory performance in monkeys. Spike activity of dTP is hypothesized to be associated with auditory-relevant events during the memory task, and may also reflect a variety of task-relevant conditions, for example, delay interval, behavioral outcome and reward delivery. It is expected that dTP would be more difficult to be evoked by sound stimuli, compared to the primary or secondary auditory brain regions. Neurons of dTP may thus be more stimulus-selective, as suggested by models for a hierarchical organization of information flows, in which higher-order auditory sensory areas receive large amounts of information from primary and secondary sensory areas (Poremba and Mishkin, 2007; Rauschecker and Scott, 2009; Hackett, 2010).

CHAPTER 2. PRIMATE AUDITORY RECOGNITION MEMORY PERFORMANCE VARIES WITH SOUND TYPE

Abstract

Neural correlates of auditory processing including 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 1A 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 1B, controlling for number of sound exemplars and presentation pairings across types, replicates Experiment 1A, 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.

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. 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 second. 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 whether 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-s memory delays. In Experiment 1A, 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 1B, 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.

Experiment 1A: Methods

Subjects

Six rhesus macaques (*Macaca mulatta*) were used, three males and three females between 11 and 12 years of age and weighing 6 – 11 kg. For approximately the first two years they were raised with other rhesus monkeys in a breeding facility in both indoor and outdoor corrals. 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-h 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 listen to a wide range of sound stimuli. The behavioral panel contained a speaker, an acrylic touch-sensitive button and a reward dish (Figure 4).

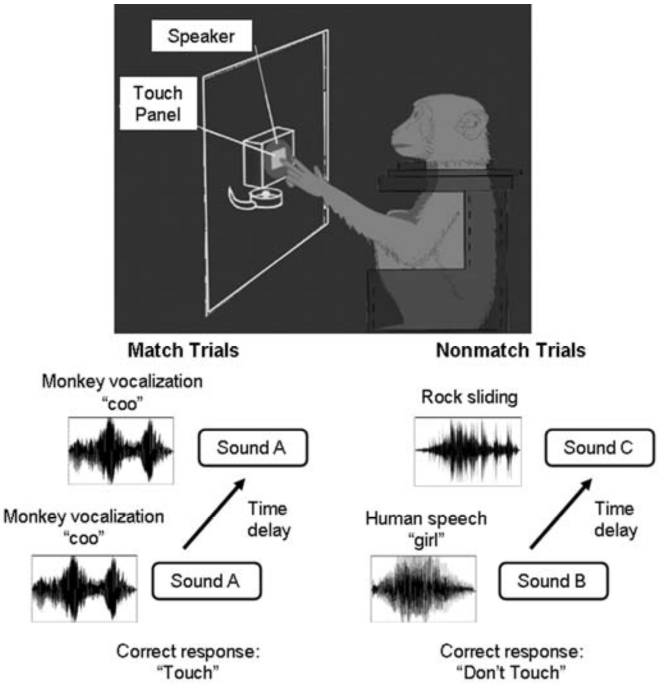


Figure 4. Schematic diagram depicting the auditory delayed matching-to-sample task. Daily sessions contain equal numbers of match and nonmatch trials. During match trials, the first sound, followed by a 5-second delay, was same as the second sound. The correct response was a touch (go-response) and the 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.

The speaker (3.5 inch x 3.5 inch) was 15 centimeters (cm) in front of the primate, at its eye level. The touch-sensitive button (2.8 inch x 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. The touch-sensitive button lit up to indicate the possible 500-ms response window on both match and nonmatch trials. 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 5 days a week and 50 trials were presented per session. The current setup employed go/no-go response rules for the auditory DMTS task (Figure 4). 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 5 s long. The inter-trial interval (ITI) was set at 12 s, and premature response during the ITIs reset the interval. The same response during 5-s 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, dog, etc.), and miscellaneous/wild animals (e.g., lion, elephant, leopard, etc.). Human vocalizations (Hvoc), 113 samples (12.8 %), included speech sounds (e.g., “girl”, “thank you”, “good morning”, etc.) and non-speech sounds (e.g., laughing, crying, 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, 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 44,100 Hz and were 8-bit mono-recorded sound clips.

Data Analysis

Results are based on a *post-hoc* database analysis to determine whether 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% 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 the 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, sex 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 sex 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 1B: Methods

After obtaining the results from Experiment 1A over a large number of behavioral testing sessions and analyzing them in a *post hoc* manner, Experiment 1B 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 systemically organized in order to reveal whether particular sound types would improve auditory memory performance. The present design examines whether the sound effects on memory performance derived from Experiment 1A could be replicated by Experiment 1B.

Subjects

Experiment 1B used four monkeys, three males and one female, that participated in, and were housed as in Experiment 1A.

Auditory Stimuli

For each of the seven sound types, 28 exemplars were chosen to represent each sound type used in Experiment 1A [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 Experiment 1A. New monkey vocalizations were recorded in a natural monkey reserve (South Carolina, USA; by the thesis supervisor 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

Experiment 1B was conducted approximately two years after Experiment 1A, and monkeys had been receiving the auditory DMTS training with the 196 sound stimuli for Experiment 1B over the preceding six to eight weeks for a separate experiment. The same go/no-go response rule for the auditory DMTS task from Experiment 1A was used. The memory delay between two sounds (5 s) and other training parameters were the same as Experiment 1A, 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

Every day, 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 Experiment 1B, 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-measures ANOVAs and linear regression analysis were used for memory performance of monkeys during match and nonmatch trials, respectively, as in Experiment 1A. 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, sex was not included as a between-subject factor. Based on the results of Experiment 1A, 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 Section 2, Experiment 1A).

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-s 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 of 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 (τ) greater than zero.

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

Results: Experiment 1A

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 (Figure 5A). There was a main effect of sex (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 sex 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 sex ($F_{(1,4)} = 0.01$; $p = 0.93$), 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.

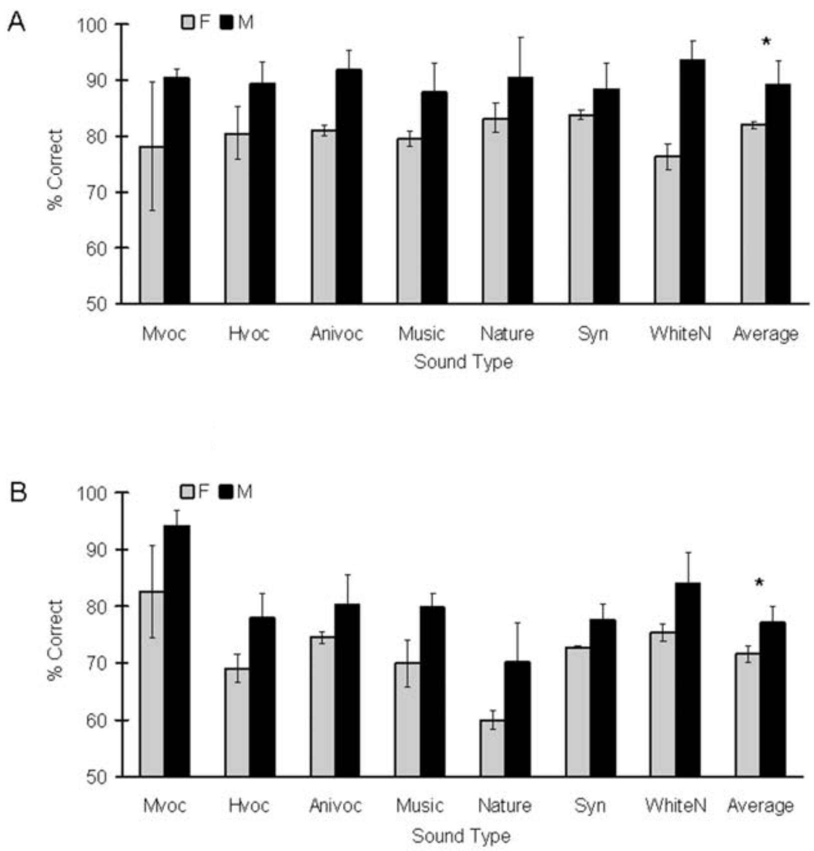


Figure 5. Experiment 1A: Sexual influence on auditory memory performance of rhesus monkeys. Auditory memory performance for match (A) and nonmatch trials (B) differs by sex. The graphs show average memory performance across seven sound types at fixed 5-second delays during match and nonmatch trials. Grey and black bars represent memory performance of female and male monkeys respectively. The asterisk indicated a significant performance difference between sexes (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).

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, sex was entered and significantly accounted for 4% 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 sex in which males performed significantly better than females (Figure 5B), 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 main effect of sound type ($R^2_{\text{change}} = 0.09$, $F_{\text{change}\ (6,220)} = 3.89$, $p < 0.005$) when presented in the second sound position (Figure 6). 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 whether an interaction between sex 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 ($R^2_{\text{change}} = 0.003$, $F_{\text{change}\,(6,\,199)} = 0.11$, $p = 0$).

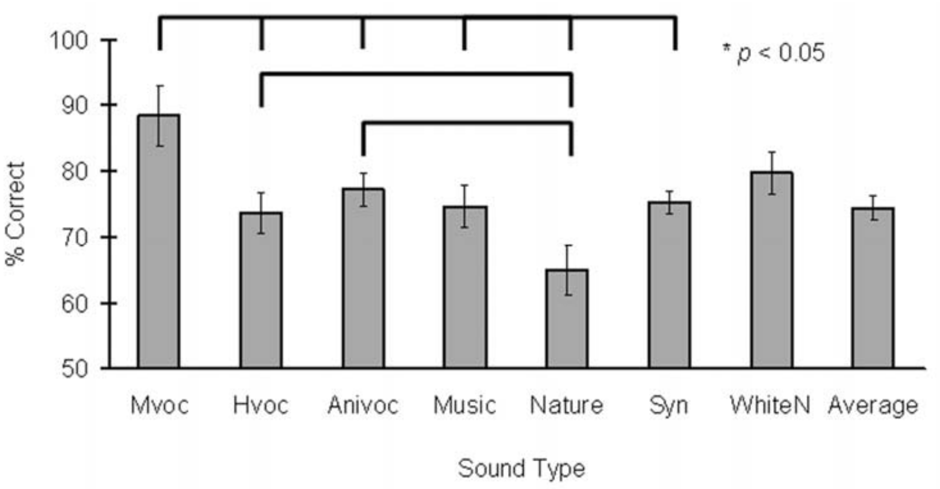


Figure 6. Auditory memory performance for distinguishing two different sounds depends on sound types. The graph shows average performance at fixed 5-second delays during nonmatch trials when specified sound type was presented as the second sound. Asterisks and brackets indicated significant performance 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.

We further examined whether 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 types showed a similar level of accuracy to the other seven sound types. These findings suggest that a simple acoustic structure did not necessarily 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, sex 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 sexes. Table 1 illustrates average individual memory performance of the six monkeys across the seven sound types, separated by sex and trial type (match or nonmatch).

Results: Experiment 1B

Effects of Sound Type on Match Memory Performance

Parallel to findings from Experiment 1A, 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 = ± 2.96).

Table 1. Individual memory performance on the auditory delayed matching-to-sample task in Experiment 1A.

		Percentage Corrrect					
Sex	Subjects	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

Effects of Sound Type on Nonmatch Memory Performance

Experiment 1B was a follow-up study to examine whether 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 1A. 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 7). In contrast to Experiment 1A, the study did not reveal significant performance differences between human or animal vocalizations and other non-vocalization sound types.

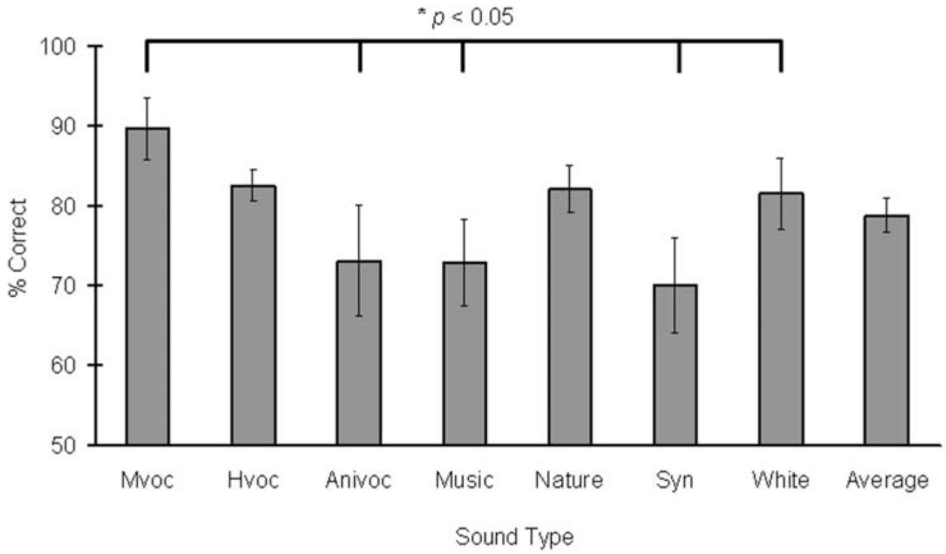


Figure 7. Auditory memory performance in a trial-unique delayed matching-to-sample task. Auditory memory performance for distinguishing two different sounds based on sound types. The graph shows average nonmatch performance at fixed 5-second delays when sound type was presented as the second sound. Asterisks 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).

Sound pairings for nonmatch trials in Experiment 1B 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 8 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 ($p < 0.05$).

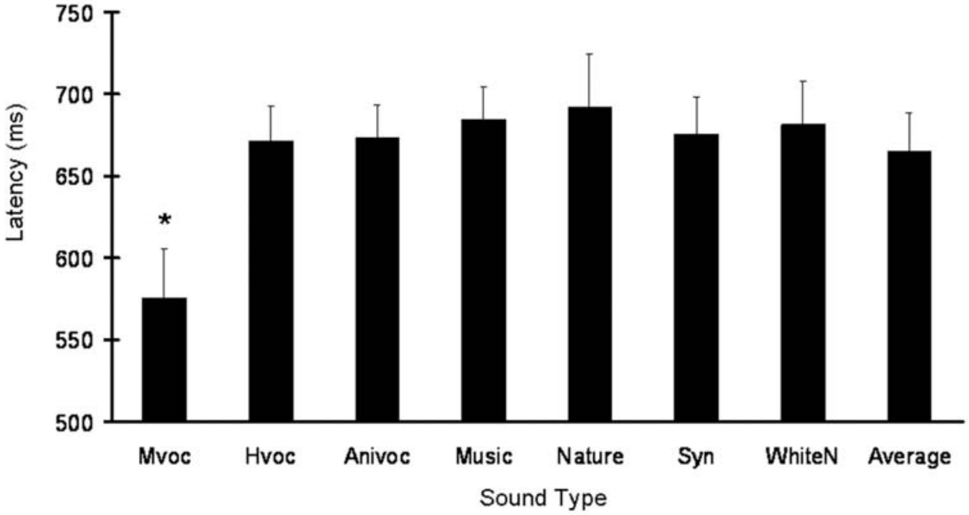


Figure 8. Effects of sound type on response latencies during auditory memory performance of rhesus monkeys during Experiment 1B. The graph illustrates average response latencies when monkeys showed correct go-responses for match trials. An asterisk indicated that monkeys responded faster to monkey vocalizations 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.

Regression analysis of nonmatch-trial conditions for response latency showed neither effect of sound type presented as the first sound ($R^2_{\text{change}} = 0.03$, $F_{\text{change }(6,\,127)} = 0.90$, $p = 0.50$) nor 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 9 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 spectral and temporal frequencies, and their power levels decrease rapidly at high frequencies. This pattern is characteristic 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.

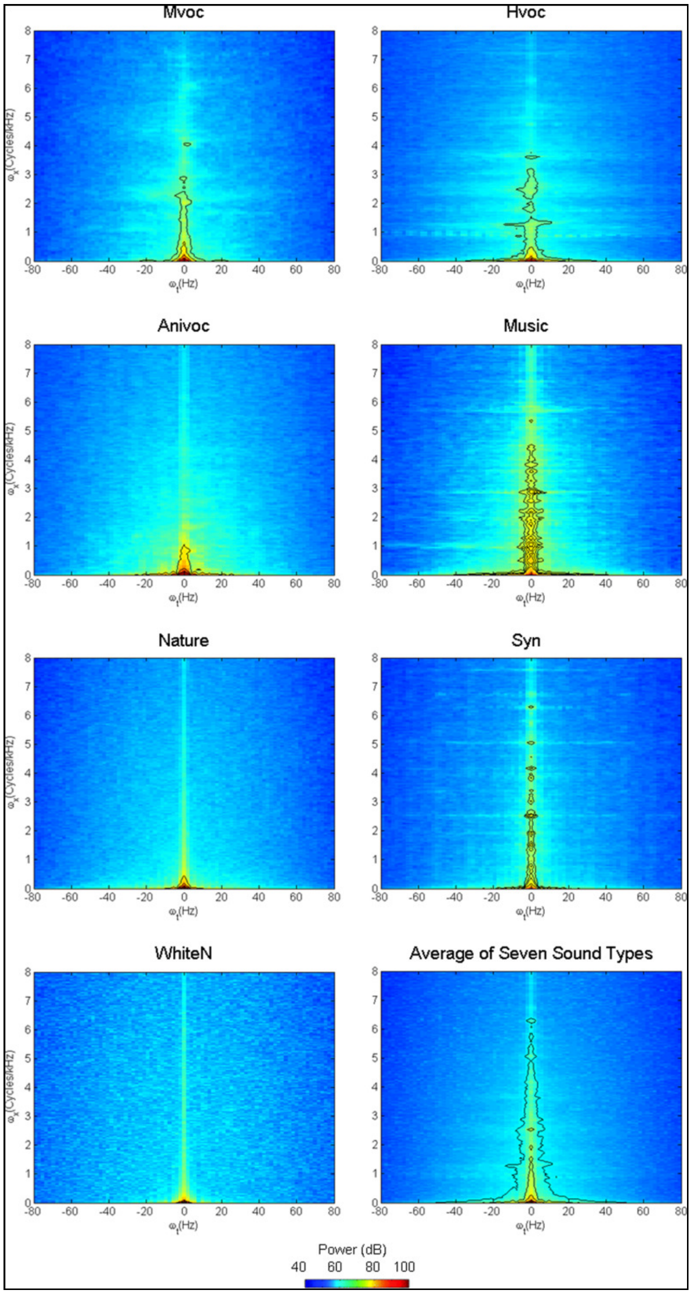


Figure 9. Modulation spectra of seven sound types in Experiment 1B. Power density is indicated by color using a decibel (dB) scale, with red showing the spectrotemporal modulations with the most energy. The x-axis and y-axis represent the frequency of the temporal modulations (ω_t or cycles/Hz) and spectral modulations (ω_f or Hz) respectively. Black lines in each spectrum are contour lines showing 50-90% of the power.

The harmonic-to-noise ratio (HNR) indicates if certain sound types tend to carry more harmonic components over time relative to noise (Figure 10). 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.

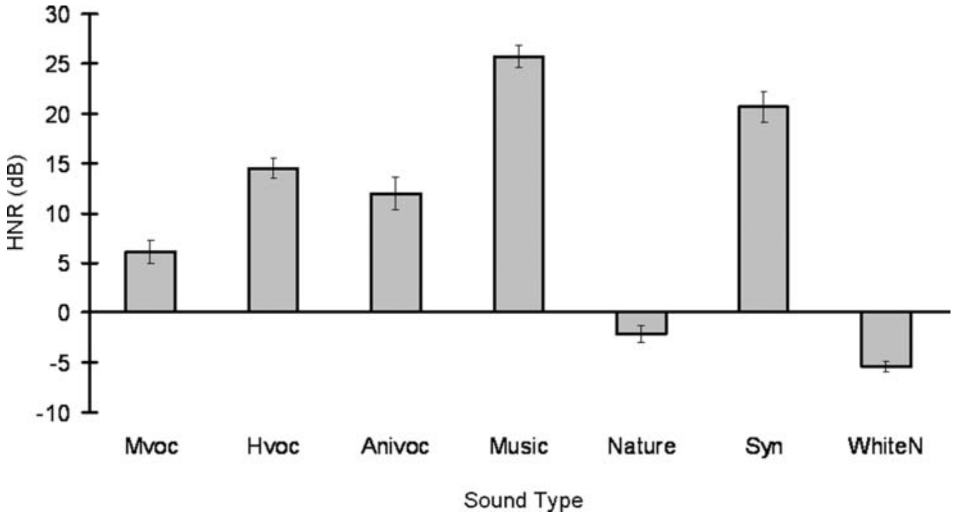


Figure 10. Harmonic-to-noise ratios of seven sound types in Experiment 1B. Values of harmonic-to-noise ratio (HNR) indicate degree of acoustic periodicity for each sound type. A high and positive ratio value for a sound type indicates more acoustic energy for that sound type devoted to harmonics over time.

We tested whether 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 Experiment 1A, monkeys show better auditory recognition memory with vocalizations, strongest for species-specific monkey vocalizations, on nonmatch trials after a fixed 5-s 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 Experiment 1B, using a trial-unique design with balanced presentation of sound types, once again showed robust memory performance on nonmatch trials with 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 vocalizations, 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 (Figures 9 and 10), 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, which 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-s delays, memory performance for two matching sounds appears to be asymptotic across sound types, while response latencies associated with monkey vocalizations are the fastest. The authors speculate whether 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 whether 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). Higher-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, but instead combine them into measures of average memory performance. The present findings reveal that 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 modulates 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-relevant behaviors, and how a series of brain regions could accommodate such task challenges.

The results of Experiment 1A reveal an effect of sex on auditory memory performance of the current DMTS task for both match and nonmatch trial performance. Individual and group performance data suggest 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. Sexual effects on perception, learning, and memory have been extensively studies 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) of male excellence in spatial processing and female excellence in non-spatial processing. Several theories have been used for describing and explaining sexual 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 sexual differences observed here in auditory memory performance.

Interactions between hormonal actions in the brain and sex 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 sexual 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 sex plays a role in nonspatial components of auditory perception, learning, and memory functions. Some field studies show sexual 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, with a small sample size on our auditory memory task, the sexual advantage was in the opposite direction with males showing higher performance levels than females.

In auditory tasks specifically there is some evidence that sexual 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; Löppönen et al., 1991; Hallmo et al., 1994; Pearson et al., 1995; Dreisbach et al., 2007), though others suggest no sexual difference at all (Osterhammel and Osterhammel, 1979; Frank, 1990; Betke, 1991), and it is unknown in rhesus macaques. The current male and female monkeys were pre-tested for basic hearing abilities before acquisitions of the auditory DMTS task, but they may not necessarily be matched on auditory sensitivity and general intelligence. The present study uses a wide variety of sounds with different acoustic profiles, from simple pure tones 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 design, and complexity of acoustic stimuli, neither agree nor contradict the present results specific to the sex of the animal. Although the currents 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 sexual 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.

CHAPTER 3. SINGLE-UNIT RECORDINGS IN DORSAL TEMPORAL POLE DURING PASSIVE LISTENING AND AN AUDITORY DELAYED MATCHING-TO-SAMPLE TASK IN MONKEYS

Abstract

Dorsal temporal pole (dTP), a cortical region at the rostral portion of the superior temporal gyrus, has been associated with auditory perception for complex sound stimuli (Poremba and Mishkin, 2007; Olson et al., 2007), and it is particularly sensitive to species-specific vocalizations in humans and monkeys (Belin et al., 2002; Poremba et al., 2004). Evidence related to neural mechanisms of auditory processing and memory function is less well-known for dTP compared to other auditory association areas. The present study employed single-unit recording techniques to examine neuronal activity of dTP when two monkeys performed an auditory version of a delayed matching-to-sample (DMTS) task. Results show that dTP encodes various task-related events during the DMTS task, cue presentations, memory delays, behavioral responses, and food rewards. Population response to sample stimuli is crucial for correct memory performance, regardless of match and nonmatch rules. Importantly, dTP showed effects of match suppression to identical sound stimuli when monkeys correctly identified two matching sounds. Delay-related activity revealed in dTP was short-lived and irregular. During passive listening, neurons of dTP were sound-selective, and mostly responsive to up to four sound stimuli although no strong relationship was identified between unit responsiveness to sounds and acoustic properties of sounds estimated from spectral and amplitude envelopes. Current findings suggest that dTP is a higher-order region for auditory processing, and it is compatible to a hierarchical organization of information

flow in the auditory nervous system of non-human primates. Dorsal temporal pole demonstrates match suppression in the current recognition memory task, similar to the activity observed in the visual object identification pathway located more ventrally than dTP, inferior temporal cortex and ventral temporal pole. However, sustained delay activity, one of the important hallmarks for models of visual working/short-term memory, is very limited and transient in dTP. It may therefore provide preliminary evidence for why monkeys have less robust auditory memory and short thresholds of forgetting, compared to their visual memory functions.

Discriminating and recognizing a variety of auditory stimuli is vital to survival in nature, from detecting potential predator-related sounds to identifying diverse communicative signals regarding biological statuses and social relationships within and between species. Working memory facilitates real-time decision-making and goal-directed behaviors for holding important acoustic information for a short time. In non-human primates, the paradigm of the delayed matching-to-sample task (DMTS) is used to test auditory memory performance. Monkeys are trained to determine whether two sounds, separated by a memory delay, are same or different. Concluding evidence suggests monkeys have low thresholds of forgetting for auditory stimuli, which is a significant drop in overall performance to 75% correct at memory delays of 35 seconds (Fritz et al., 2005). This is in contrast to the relatively high thresholds of forgetting for visual stimuli at memory delays of 10 minutes or longer (Murray and Mishkin, 1998; Buffalo et al., 1999; Zola et al., 2000). Bilateral damage to higher-order auditory regions along the superior temporal gyrus severely impairs short-term memory with auditory stimuli (Colombo et al., 1990, 1996; Fritz et al., 2005). The respective neural basis of auditory memory and possible differences with visual memory remain unclear.

Inquiries on neural mechanisms of auditory working and recognition memory performance may benefit from understanding the well-established network of visual working memory in monkeys. Along the ventral “what” stream of the information processing pathway, inferior temporal cortex (ITC) and the prefrontal cortex (PFC) are necessary for working memory of visual objects (Passingham, 1975; Goldman et al., 1971; Bachevalier and Mishkin, 1994; Ungerleider et al., 1998; Buffalo et al., 1999). In visual DMTS tasks, many visual neurons of ITC and PFC show stimulus-selective activity and sustained firing during memory delay intervals (Fuster and Jervey, 1981, 1982; Miller et al., 1991, 1993; Colombo and Gross, 1994; Tanaka, 1996; Woloszyn and Sheinberg, 2009). A functional difference between these two cortical regions is that ITC primarily encodes details of visual information and contributes to visual analysis of viewed objects. Compared to PFC, the majority of ITC neurons are more stimulus-selective and more phasic to stimulus onset (Miller et al., 1996; Freedman et al., 2003). Their delay activities are prone to change with stimulus identity or category when two presented visual objects are different during nonmatch trials (Muhammad et al., 2006; Meyers et al., 2008). Auditory association areas along the anteroventral axis of the temporal lobe, analogous to ITC, may then become potential candidates to support auditory analysis of sound details for recognition and working memory of monkeys. The current focus for neural mechanisms of auditory encoding and memory functions is the dorsal temporal pole (dTP), located at the rostral tip of the temporal lobe.

Functions of temporal pole (TP) have been associated with sensory perception, memory, and emotional processing based on human functional imaging (Blair et al., 1999; Kim et al., 2005; Jimura et al., 2009) and clinical studies of patients suffering from partial or complete TP damage (Olson et al., 2007). Lesions of superior temporal gyrus (STG), including the dorsal portion of TP, spare hearing but impair monkeys' ability to perceive auditory pattern of sounds (Weiskrantz and Mishkin, 1958; Dewson et al., 1969, 1970; Iversen and Mishkin, 1973). The auditory deficits weaken their ability to discriminate between conspecific monkey vocalizations or distinguish monkey vocalizations from complex acoustic stimuli (e.g., dog barks, human voices, FM sweeps and white noises; Kupfer et al., 1977; Heffner and Heffner 1984, 1986). Functional imaging studies hint at the roles of TP for auditory processing of species-specific vocalizations in rhesus monkeys (Poremba et al., 2004) and voices in humans (Belin et al., 2000; 2002, Nakamura et al., 2001; Andics et al., 2010). The temporal pole receives auditory, visual, olfactory, and polysensory information from different cortical regions. Anatomical studies suggest these various streams of sensory information reach different parts of TP and remain separate (Nakamura and Kubota, 1996; Kondo et al., 2003, 2005). Dorsal temporal pole (dTP), which consists of granular and dysgranular areas, has extensive connections with surrounding auditory regions of STG, parabelt areas, limbic thalamus, amygdala, hippocampus, lateral, orbital and medial prefrontal cortices (Markowitsch et al., 1985; Moran et al., 1987; Yeterian and Pandya, 1989; Suzuki and Amaral, 1994; Carmichael and Price, 1995; Hackett et al., 1998; Barbas et al., 1999; Romanski et al., 1999a; Kondo et al., 2005; Saleem et al., 2008). The ventral portion of TP (vTP) shows neural coding for visual analysis of stimulus details like ITC. The spike activity is maintained by repetitions of visual stimuli during delay intervals, and vTP neurons appear to contribute in an active, working memory epoch like those in PFC (Nakamura and Kubota, 1995). The dorsal, auditory domain of TP may support complex analysis of auditory information, and further facilitate learning and memory functions when holding stimulus-/task- relevant events for working memory usage.

The present study examines response profiles of dTP neurons when monkeys listen to various sound types and perform an auditory delayed matching-to-sample task. A collection of 129 sound stimuli from eight distinct sound types is used to assess spike activity of dTP neurons. Each session consists of passive listening blocks and a behavioral DMTS task. The current passive listening experiment is the first study to examine neuronal activity of dTP across a wide range of auditory stimuli *in vivo.* It aims to elucidate how auditory neurons of dTP represent acoustic features of sounds, and how their activities distinguish sound types bearing different levels of acoustic complexity and biological/ethological significance, such as conspecific and heterospecific sounds to monkeys including music and man-made environmental sounds. Since neuronal activity in higher-order auditory and visual cortical areas is often difficult to be evoked, compared to the respective primary areas. It is expected that units of dTP are sound selective, and only respond to a few sound stimuli in general. Conspecific monkey vocalizations are shown to promote spontaneous categorization of sounds (Hauser, 1998; Gifford et al., 2005; Cohen et al., 2006; Ghazanfar et al., 2007) and induce significant activation in dTP (Poremba et al., 2004). This sound type may be more effective to evoke neuronal responses than other types. In contrast, the memory experiment surveys how neural representations of dTP neurons contribute to auditory recognition memory. Based on perceptual nature of ITC and vTP on visual details, dTP neurons are expected to show stimulus-selective activity, reflecting acoustic details of sounds like those neurons in rostral STG and lateral belt areas along the anteroventral stream of nonspatial information pathway (Rauschecker and Tian, 2000; Poremba and Mishkin, 2007; Russ et al., 2008).

General Methods

Subjects and Surgical Methods

Two adult rhesus macaque monkeys (*Macaca mulatta)* were used, a male and a female, weighing 10 and 6 kilograms, respectively. They were individually housed in Spence Laboratories at the University of Iowa (12-h light/dark cycle). All procedures were approved by the Institutional Animal Care and Use Committee at the University of Iowa. Monkey biscuits (Harlan Teklad, Madison, WI) were fed to animals daily with fruits, vegetables, and treats scheduled throughout the week. Monkeys had access to water ad libitum in home cages equipped with environmental enrichment. Each animal's weight was maintained above 85% of their starting weight with controlled daily feeding schedules, and weight was adjusted upward based on age. Prior to surgery, each monkey was first scanned with magnetic resonance imaging (MRI: 2T Sigma unit; GE Medical Systems, WI) for locating the precise coordinates of temporal pole. Placements of recording chambers were initially performed at the National Institute of Mental Health (Bethesda, MD). Monkeys were sedated with ketamine (10mg/kg i.m.) and anesthetized with isoflurane (1-2%). Using a stereotaxic apparatus (David Kopf Instruments, CA), recording chambers were implanted on the skull of the left hemisphere above the parietal cortex, and its position was secured with screws and dental acrylic. A stainless steel headpost was attached tight against the backside of the skull for restraining head movement during electrode recording. Antibiotics and analgesics were given to the animals after surgery. Recording chambers were cleaned routinely with antiseptics to inhibit infection once open for recording.

Auditory Stimuli

All auditory stimuli were presented through a speaker, placed 15 centimeters (cm) in front of the animal at 80 decibels (dB). Each sound stimulus, 220 to 500 milliseconds (ms) long, was digitized and processed with a sampling frequency of 44100 Hz, and consisted of 8-bit mono-recorded sound clips. A collection of 96 standard stimuli was used and classified into eight different sound types: animal vocalizations (Anivoc, n = 12), human vocalizations (Hvoc, n = 12), monkey vocalizations (Mvoc, n = 12), music clips (Music, n = 12), natural sounds (Nature, n = 12), synthesized clips (Syn, n = 12), band-passed white noises (WhiteN , n = 12), and pure tones (PT, n = 12). The 96 standard stimuli were then sorted into 12 sound folders, with each folder containing one sound sample from the eight sound types. Twenty-one additional exemplars of PT (250-ms long) were used for sampling spike activity of dTP across different sound frequency levels. They cover the frequency range between 100 Hz to 20 kHz (100 – 1000 Hz at 100-Hz incremental steps; 1 – 10 kHz at 1-kHz incremental steps; 10 – 20 kHz at 5-kHz incremental steps). These 21 sinusoidal pure tones were digitally generated and normalized with root-mean-square methods. Twelve additional monkey vocalizations (220 – 500 ms long) were recorded in a natural monkey reserve (South Carolina, USA; by the thesis supervisor A.P.), and created in a similar fashion as in Methods of Chapter 2's Experiment 1B.

Acoustic Analyses of Sounds

Analyses were performed to determine acoustic characteristics within each sound type, and helped reveal their possible relationship with the spike activity of dTP neurons. The harmonics-to-noise ratio (HNR, expressed in dB), i.e., 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 (τ) greater than zero.

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

Since sound stimuli contained rapid, changing frequencies at different energy levels, frequency and intensity should be taken into account for acoustic analysis. All calculations were done in the MATLAB programming environment (The Math Works; Natick, MA). It was first calculated the two-dimensional Fourier transform of each sound. If a sound (e.g., a monkey vocalization) was less than 500 ms, it was zero padded until its length was 500 ms. Sound amplitudes at each frequency were then converted to power in scales of dB. Spectral mean of a sound stimulus was defined as a weighted mean of frequency across intensity as followed:

$$\text{Spectral mean} = \sum(f_i \times P_i) / \sum(P_i)$$

i : frequency level at 1–Hz increment.

Values of each sound stimulus were also pooled into the eight sound types for further statistical analyses.

Recording Procedures

Single-unit activity of dorsal temporal pole was recorded by lowering tungsten microelectrodes (130-140 millimeters long, 1-3MΩ; FHC Inc., Bowdoin, ME) at appropriate angles through the parietal cortex to the dorsal region of temporal pole (Figure 11). Microelectrodes were inserted into a 23-g sterile guide cannula that was held by an x-y grip positioner attached to a micromanipulator. Electrode advancement was accomplished by a computer-controlled electrode drive system (NAN Instruments, Plexon Inc., TX). Spike activity was isolated, amplified and discriminated in real-time by the Multichannel Acquisition Processor with the SortClient program (Plexon Inc., TX).

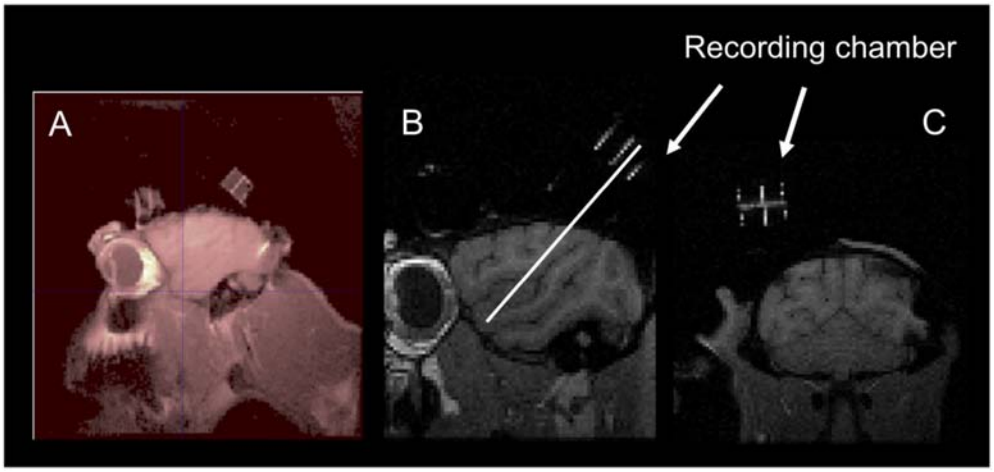


Figure 11. MRI images depicted the location of dorsal temporal pole of monkey AB. A. A sagittal view of monkey AB's head. The recording chamber is positioned above the left parietal cortex at an angle on the skull surface. B. A magnified, sagittal view of the temporal lobe with the three white lines representing the top, center, and bottom parts of the recording chamber. C. A coronal view from the back of the monkey brain shows the position of the recording chamber. The three white lines represent the left, center, and right parts of the recording chamber, and the horizontal line indicates the inner diameter of the chamber.

Corresponding data, as well as timelines of stimulus and behavioral events, were saved for offline analysis. The x-y position of the recording site was recorded every time, and was compared with MRI coordinates to determine locations of dTP cells recorded. Recordings were conducted in the left hemisphere of dTP, since consistent findings showed a lateralization of higher-order auditory perception in the left hemisphere in terms of faster acquisition and robust neuronal activity (Heffner and Heffner, 1984, 1986;

Petersen et al., 1984; Poremba et al., 2004). This was to maximize the probability of finding significant dTP activity during passive listening and the memory task. To minimize sampling bias, any neuron isolated was tested and no sound stimuli was presented to the animal when lowering electrodes and isolating a neuron. Once an isolated dTP neuron was obtained, the experimenter commenced the recording session.

Data Analysis

Spike activity was recorded with a sampling frequency of 40 kHz. Each isolated unit was sorted into waveforms for offline analysis (Offline Sorter, Plexon Inc., TX). On a neuron-by-neuron basis, the response profile of each dTP neuron was constructed for passive listening and the auditory DMTS task, in regard to different stimulus or task events. Peri-stimulus time histograms (PSTH) were created to visualize spike activity related to various stimulus and task events across time using NeuroExplorer (Nex Technologies, Littleton, MA). All spike activity values were sampled at 10 ms intervals. A baseline mean and standard deviation of each spike unit were calculated from the activity values during the 500 ms pre-trial period (passive listening and the auditory DMTS task). Activity values of each event were subtracted by the baseline mean, and then divided by the standard deviation of the baseline period. The resultant standardized scores were then used for statistical analysis. The standardization of spike activity helped eliminate variance of firing rate of each dTP unit when examining population activity of dTP across recording sessions. Spike activity of each unit was compared to baseline firing rate to determine whether it was responsive to sounds or task-relevant events. Baseline firing rate was estimated from recording intervals, 500 ms prior to stimulus presentation (passive listening) or prior to trial presentation (the DMTS task). The similar analysis was also applied to assess neuronal activity of various stimulus or task events during the DMTS task. Additional details of data analysis were provided in the experiment sections of the auditory DMTS task (Exp. 2A) and the passive listening task (Exp. 2B). Mean firing rate of an event was defined as the number of action potentials divided by the length of a specific time interval.

Experiment 2A: Auditory DMTS Task

Each animal was trained with the delayed matching-to-sample (DMTS) task with auditory stimuli (for more details, see Exp. 1B, Methods). The task employed go/no-go response rules for the auditory DMTS task. A pre-determined folder of eight sounds (1 stimulus per sound type, 12 folders in total) was chosen for each recording session. The ratio of match to nonmatch trials was 1, psuedorandomly 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 sounds presented 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. Sometimes, subjects received a 300 – 500 ms air puff (a mild punishment) when they pressed the button during nonmatch trials (no-go conditions). The mild air puffs were delivered semi-randomly during training to discourage erroneous or impulsive responding. A prolonged inter-trial interval (up to 30 seconds long) was introduced after the last incorrect button-press when the monkey kept responding to the previous five nonmatch trials successively. In each trial, the memory delay between two sounds (i.e., inter-stimulus intervals) was always five seconds long. The inter-trial interval (ITI) was randomized across sets of 6, 8 & 10 s and 8, 10 & 12 s for the monkeys OP and AB respectively, and a 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. Positions of sound presentations on nonmatch trials (i.e., cue 1 or cue 2) were completely counterbalanced among the eight sound types. Each sound stimulus had equal probability to appear in match and nonmatch trials, as well as become the sample (cue 1) or test (cue 2) stimulus. The behavioral task contained 200 trials in order to yield 8 – 10 repetitions of each sound stimulus at each trial and event condition. Trials were sometimes excluded for data analysis, for example, due to pre-mature response during ITIs or ISIs, poor memory performance, presentation error during sounds and events due to an occasional computer glitch, and movement artifact. A minimum of 10 trials were used for any trial-type analysis.

Percentage correct and response latency during match and nonmatch trials were collected by the LabView program. Eighty-six total sessions from the two monkeys (OP: n = 42; AB: n = 44) were used for data analysis where their memory performance for match and nonmatch trials was correct on an average of 60% for each trial type. Repeated-measures ANOVAs and linear regression analysis (SPSS 13.0; Chicago, IL) were used for memory performance of monkeys during match and nonmatch trials, respectively. In match trials, sound type was a within-subject factor for conducting repeated-measures ANOVAs. 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 (cue 1, the first position) or as the test stimulus (cue 2, the second position) on different trials. Percent correct of a given sound pairing was the dependent variable. Regressions were conducted hierarchically with sound type presented as the first sound or the second sound entered to account for within-subject variability. Paired-sample *t*-tests were used for pairwise comparisons, and were responsible for examining performance differences between any two of the eight 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).

For each recorded cell, mean firing rates regarding sound presentation (cue 1 and 2), memory delay, wait and response periods were calculated for task events. Separate tests of one-way ANOVAs ($p < 0.05$) verified whether any of these events was significantly different from baseline, with *post-hoc* Tukey's tests used for pairwise comparisons. Each trial was divided into a baseline period (500 ms; before cue 1), cue 1 and cue 1 offset events (500 ms each), cue 2 and cue 2 offset events (500 ms each), delay event (4500 ms), wait time (500 ms) and response period (1500 ms). For sound presentations, cue presentation (cue 1 and 2) and post-cue event (cue offset 1 and 2) were each binned into 100 ms intervals for ANOVAs. For the other task-relevant events, each of them was divided into epochs of 500 ms for ANOVAs; wait time (1 epoch), response period (3 epochs) and delay period (9 epochs). To examine trial influence on spike activity across task-relevant events, repeated-measure ANOVAs were used to evaluate whether dTP neurons would change due to matching rules (match versus nonmatch trials), behavioral outcome (correct versus incorrect) or a combination of both factors. Otherwise specified, *post-hoc* comparisons were conducted by paired sample *t*-tests with the Bonferroni procedure and Keppel's modification to correct for multiple trial-type comparisons (Keppel, 1982). The product of the number of degrees of freedom and the standard alpha level of 0.05 was divided by the number of *t*-test comparisons. The resultant value was then the adjusted critical probability level.

Experiment 2B: Passive Listening

A folder of 21 pure-tone stimuli and the 12 sound folders were presented in series of blocks. Within each block, the animal listened to only one sound folder and each stimulus was repeated at least 8 times. The order of sound presentations within each block was randomized by the LabView program (National Instruments, Austin, TX). Inter-stimulus intervals were randomized among 1, 1.2 and 1.5 s. In every recording session, one of the 12 sound folders (the target folder) was pre-selected for the auditory DMTS task in each recording session. Once a unit was isolated, the target folder and a folder of 21 pure tone stimuli were first presented before the memory task, and constituted the passive listening 1 session. After the auditory DMTS task, the remaining 11 folders were presented and constituted the passive listening 2 session. At times, an additional folder of 12 monkey vocalizations was presented to the subject if the isolated unit remained active after passive listening 2. All monkeys were restricted in a primate chair inside the same sound-attenuated chamber (background noise: 35dB) as during the memory task. The animal was not required to respond and no food reward was given throughout the passive listening experiment.

Each isolated unit was tested first with a one-way ANOVAs ($p < 0.05$) to determine whether it was auditory responsive relative to baseline activity. *Post-hoc* Tukey's tests were used with the similar procedures mentioned in Experiment 2A to control errors due to multiple comparisons. For the 12 sound folders and the additional folder of monkey vocalizations, spike activity of the baseline period for each unit was normalized into values of 100 ms intervals. Spike activity during stimulus period and stimulus offset period was then binned into 100 ms intervals (i.e., 5 – 100 ms intervals each). For the 21 pure-tone stimuli, spike activity during baseline, stimulus period and stimulus offset period was binned into 50 ms intervals (i.e., 5 – 50 ms intervals each).

Results

Behavioral Performance on the Auditory DMTS Task

The two monkeys performed, on average, 78% and 66% correct on match and nonmatch trial, respectively. The male monkey OP never vocalized during the memory task. The female monkey AB vocalized around 10 – 15 times per session, usually during the inter-trial intervals. The male monkey produced button-press responses with its right hand for greater than 99% of the time, while the female monkey primarily used its left hand to press the button. Average response latency for the two monkeys was 181 ms and 209 ms on match correct trials and nonmatch incorrect trials respectively.

Auditory memory performance for two matching sound stimuli was consistently good across the eight sound types (overall mean = 78, standard error = ± 1.5), indicating that no main effect of sound was present ($F_{(1,85)} = 0.94$; $p = 0.33$). In contrast, auditory memory performance when distinguishing two different sound stimuli apart was influenced by sound types. There was a main effect of sound type presented as the second sound ($R^2_{\text{change}} = 0.20$, $F_{\text{change}\,(7,\,97)} = 3.96$, $p < 0.005$), but no main effect of sound type presented as the first sound ($R^2_{\text{change}} = 0.09$, $F_{\text{change}\,(7,\,104)} = 1.46$, $p = 0.19$). When the second sound was a pure-tone stimulus, animal subjects yielded significantly better memory performance than any of the other seven sound types (paired-sample *t*-tests, $p < 0.05$; Figure 12).

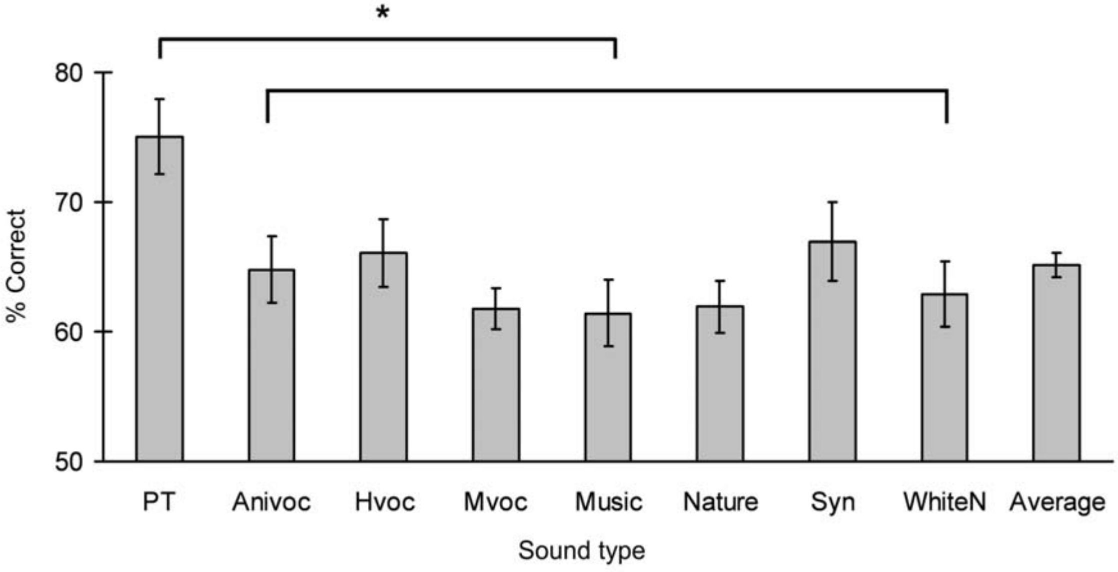


Figure 12. Auditory memory performance in the auditory DMTS task. Auditory memory performance for distinguishing two different sounds based on sound types. The graph shows average nonmatch performance at fixed 5-s delays when a particular sound type was presented as the second sound. Asterisks and brackets indicate a significant performance difference between two sound types (paired-sample *t*-tests, $^*p < 0.05$). Pure tones (PT) yielded better memory performance than those associated with animal vocalizations (Anivoc), human vocalizations (Hvoc), monkey vocalizations (Mvoc), music clips (Music), natural sounds (Nature), 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 pure tones as the second sound (test).

Memory improvement due to pure tones presented as the second sound may depend on which sound type was presented as the first sound. Thus, this factor was entered into the regression model: $R^2_{\text{change}} = 0.05$, $F_{\text{change }(6,\ 91)} = 1.05$, $p = 0.39$. The result showed no interaction between a given sound type and pure tones when considering the first and second sound position. The results suggest that auditory memory performance was improved accordingly when a pure-tone test stimulus (the second position) was compared against a sample stimulus of any sound type.

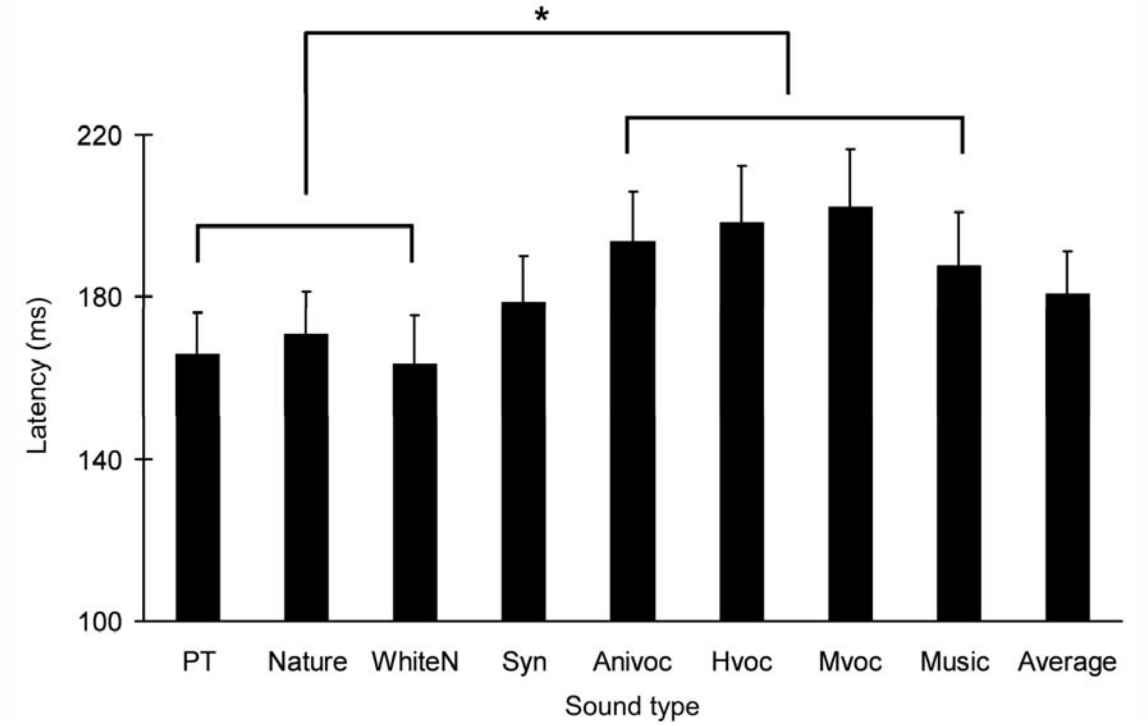


Figure 13. Effects of sound type on response latencies during auditory memory performance of two monkeys. The graph illustrates average response latencies when monkeys showed correct go-responses for match trials. Asterisks and brackets indicate that monkeys responded faster to pure tones, natural sounds or white noises than the other four sound types during match trials (paired-sample *t*-tests, $^*p < 0.05$). Response latencies were similar across the eight sound types during nonmatch trials.

On match trials, there was a main effect of sound on response latency ($F_{(7,581)} = 5.79$; $p < 0.05$). Figure 13 illustrates average response latencies across the eight sound types during match-trial conditions and indicates effects of sound mainly due to the three sound types: pure tones, natural sounds and white noises. Subjects showed significantly faster go-responses (correct) for these three sound types than for any of the music clips and vocalizations from animals, humans or monkeys (paired sample *t*-tests, $p < 0.05$). Regression analysis of nonmatch trial conditions for response latency showed neither effect of sound type presented as the first sound ($R^2_{\text{change}} = 0.01$, $F_{\text{change}\,(7,\,104)} = 0.11$, $p = 0.99$) or the second sound ($R^2_{\text{change}} = 0.01$, $F_{\text{change}\,(7,\,97)} = 0.11$, $p = 0.99$) when monkeys produced incorrect button presses.

Single-Unit Activity of dTP during the Auditory DMTS Task

The study collected 225 units at the dorsal temporal pole when subjects were participating in the memory task (Monkey OP: N = 112; Monkey AB: N = 113). The majority of these recorded cells (80%) were responsive to at least one of the task-relevant events and considered task-related: the sample sound (cue 1, 500 ms long), test sound (cue 2, 500 ms long), their respective stimulus offset periods (cue 1 and 2 offset, 500 ms long each), memory delays (divided into early, middle and late portions), wait time (500 ms long) and response window (three epochs of 500 ms intervals).

Correct Trials

The following paragraphs summarize response profiles of the 225 units in dTP during match and nonmatch correct trials. Among these 225 units, 13% and 16% of them, respectively, showed activity change in one or more of the 5 – 100 ms intervals during presentation of cue 1 and cue 2 in match correct trials (Table 2). For cue offset period, 8% and 10% of recorded units showed activity changes during the offset periods of cue 1 and 2. On nonmatch correct trials, 11% and 12% of the units showed activity change in one or more of the 5 – 100 ms intervals during presentation of cue 1 and cue 2. And 8% and 10% of them showed activity changes during the offset periods of cue 1 and 2. These cue-evoked activity changes were mainly increases or decreases in firing rate against the baseline period, and rarely were there both significant increases and decreases within the 500 ms cue presentation. Figures 14 to 17 demonstrate various examples of dTP units showing cue-related activity during match and nonmatch correct trials.

Table 2. Percentage of units showing activity change (increase +, decrease -, firing rate FR or both) during cue events for match correct and nonmatch correct trials (N = 225).

		Cue 1	Cue 1 offset	Cue 2	Cue 2offset
Match Correct (MC)	+FR	8	3	8	4
	-FR	4	4	8	5
	Both	1	1	-	1
	Total	13	8	16	10
Nonmatch Correct (NC)	+FR	6	6	7	6
	-FR	5	2	5	3
	Both	-	-	-	1
	Total	11	8	12	10

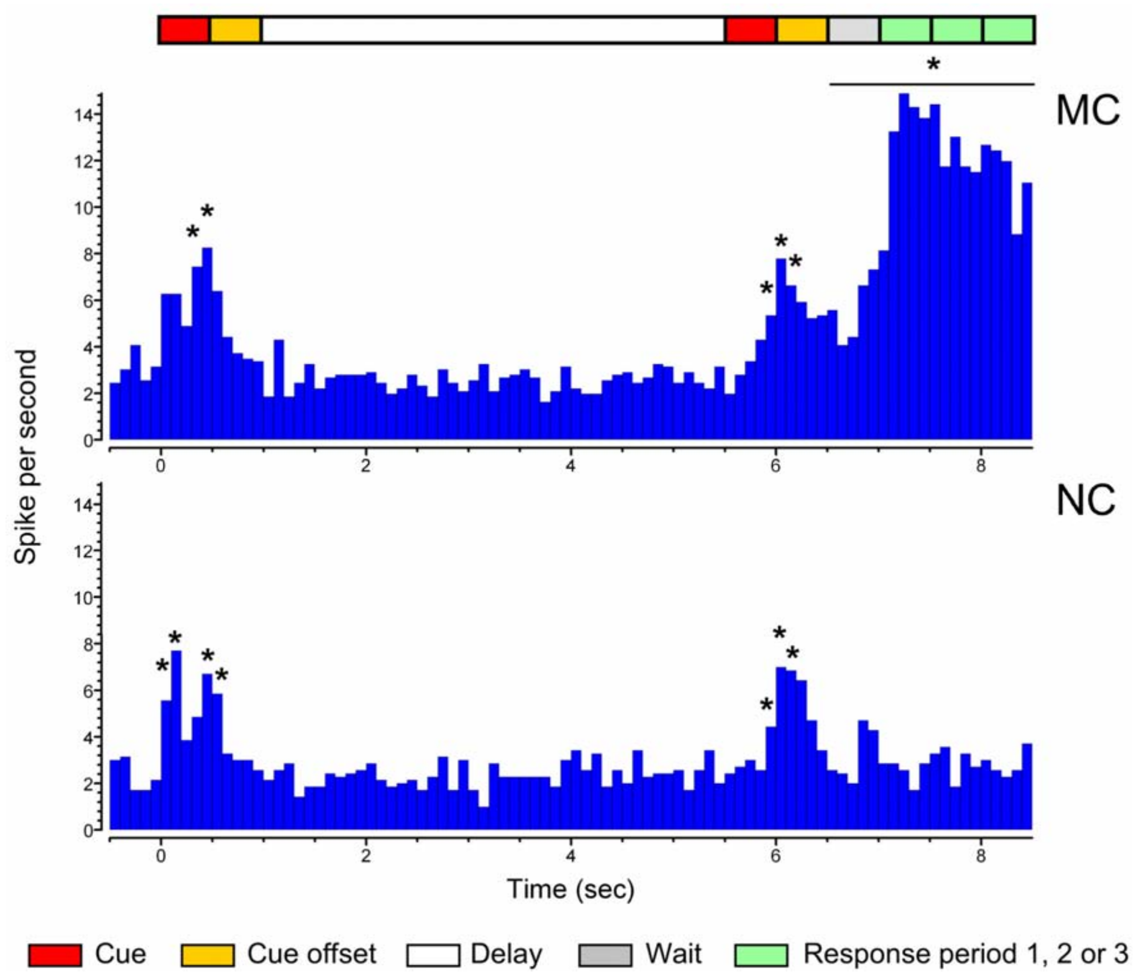


Figure 14. An example of a unit (# 1021092a) responsive to the events of cue presentations and offsets, wait and response periods during match and nonmatch correct trials. Asterisks denote significant activity change against baseline during a given 100 ms interval or a given 500 ms interval. Each bin is 100 ms. Compared to nonmatch correct trials, this unit had sustained firing from wait toward response periods 1 – 3 during match correct trials. This effect was modulated by press responses and food rewards.

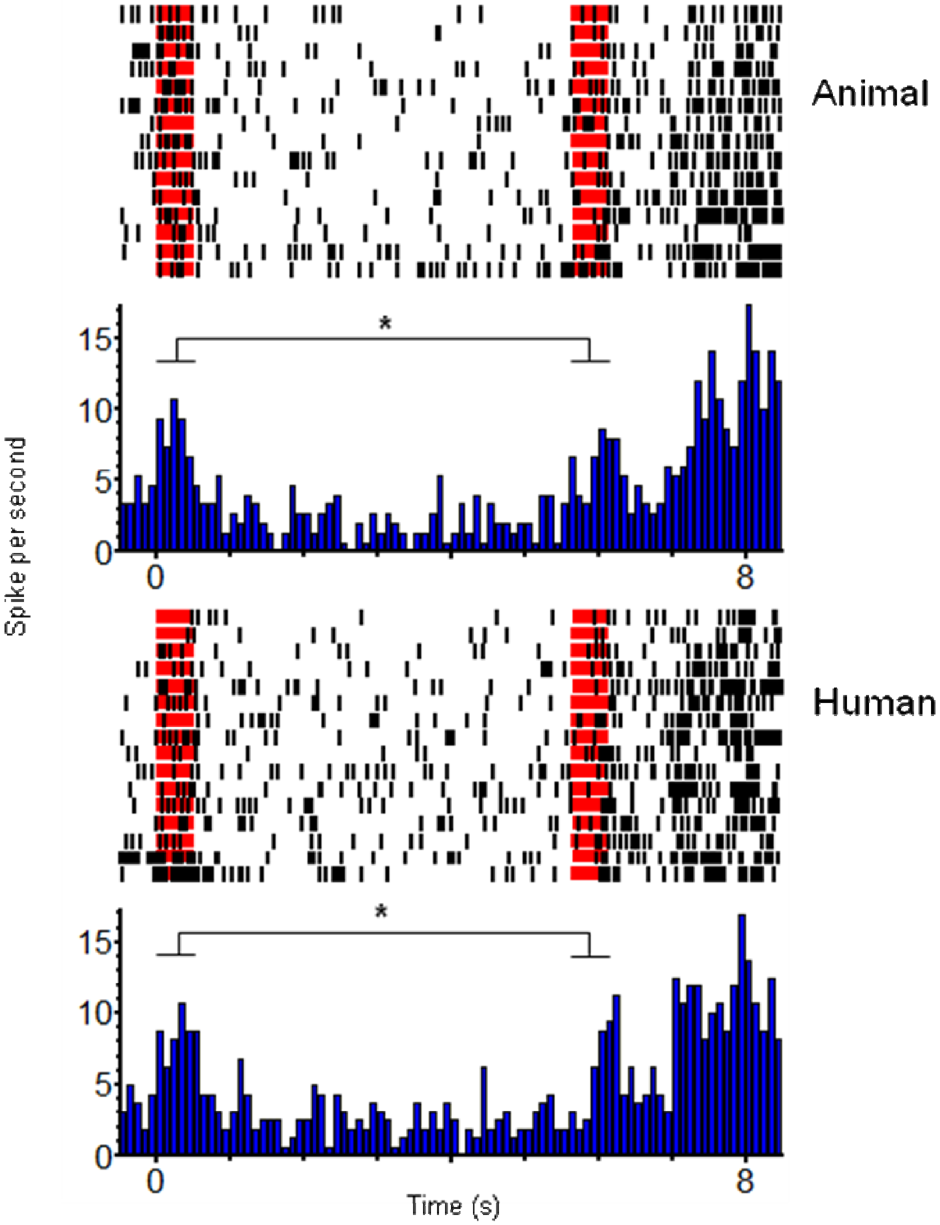


Figure 15. An example of a unit (# 1028091c) showing match suppression when the monkey correctly identified two matching sounds. Asterisks and brackets denote significant activity difference between cue 1 and cue 2 (paired sample *t*-test, $*p < 0.05$). Each bin is 100 ms. Red bars represent sound presentations, and black ticks represent action potentials in a raster plot of spike activity. The unit was responsive to four sound stimuli during the DMTS task: an animal vocalization, a human vocalization, a monkey vocalization, and a synthesized clip. Match suppression occurred when the two matching sound were an animal vocalization or a human vocalization, but not the other two sounds.

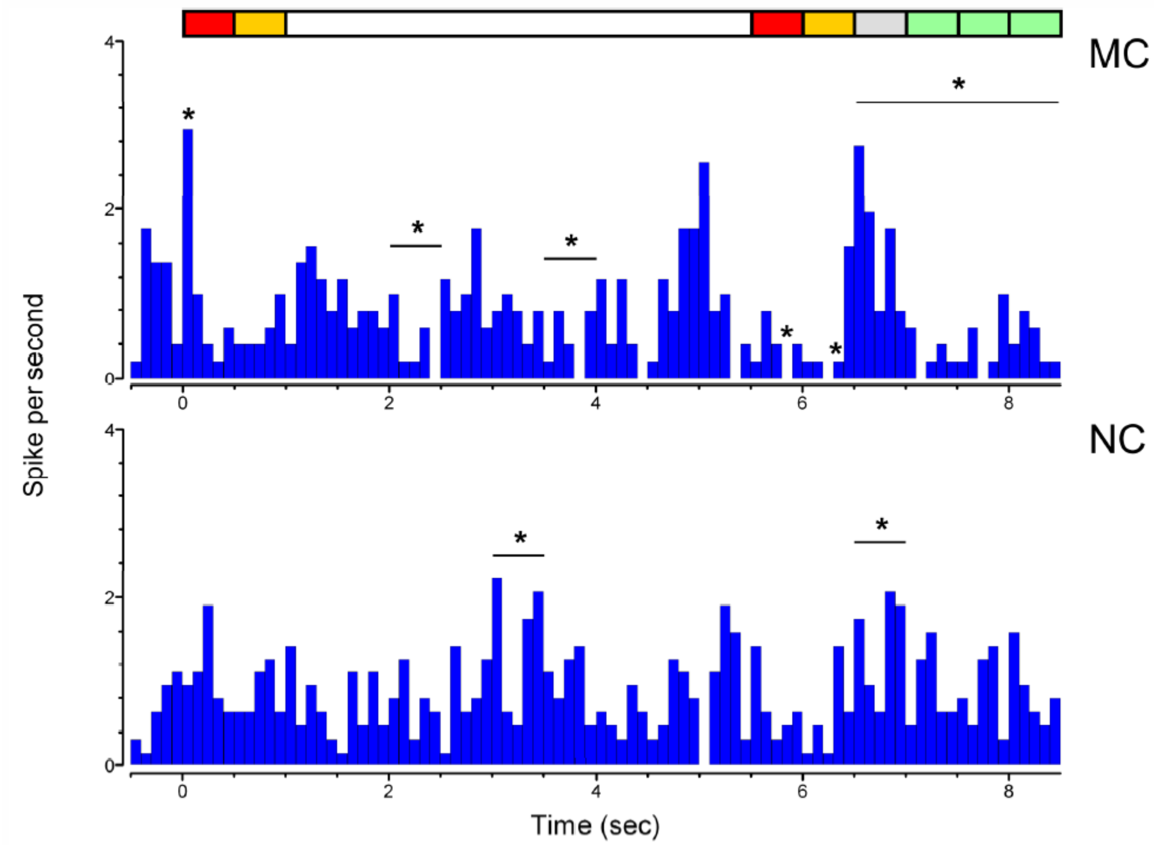


Figure 16. An example of a unit (# 0614102a) responsive to the events of cue 1, cue 2, cue 2 offset, delay, wait and response periods during match and nonmatch correct trials. Asterisks indicate significant activity change from baseline for 100 ms bins or 500 ms intervals. Each bin is 100 ms. During match correct trials, this unit increased firing to cue 1 and decreased firing to cue 2 relative to baseline.

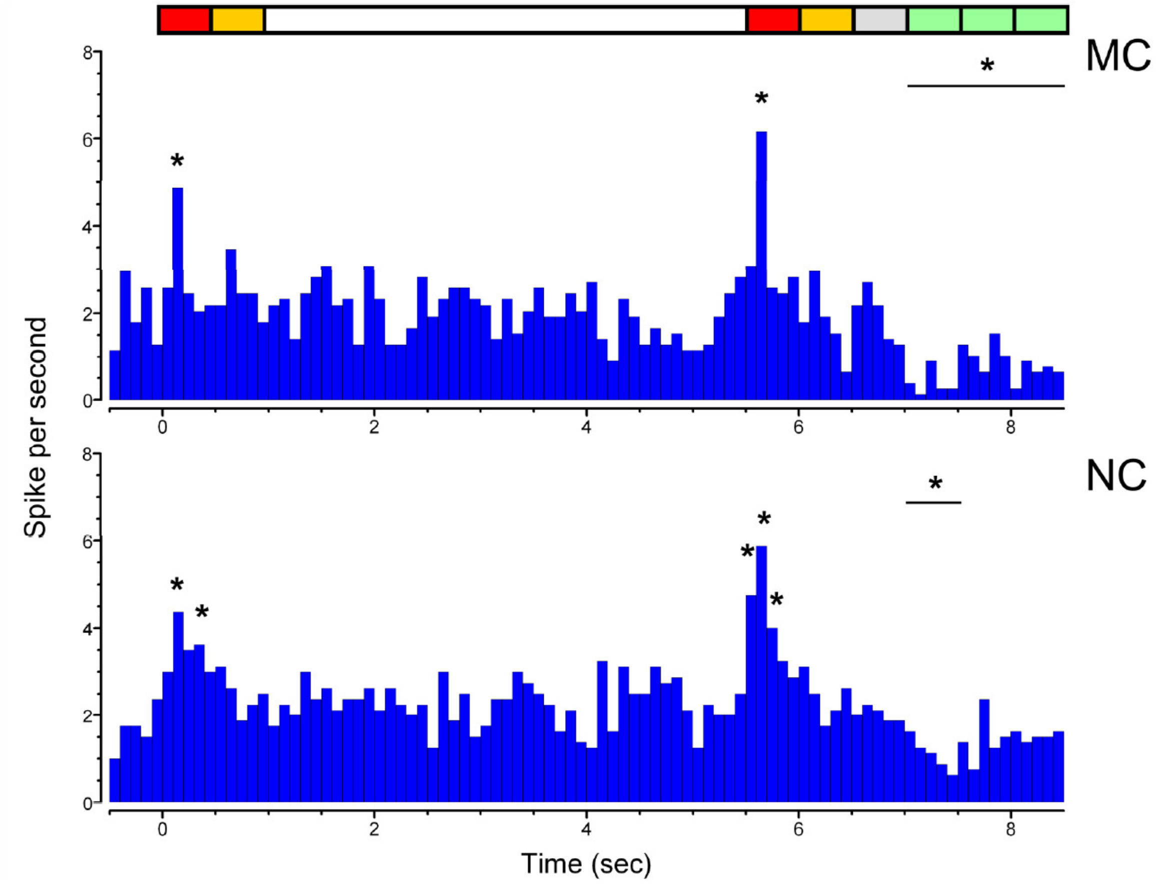


Figure 17. An example of a unit (# 0420102b) responsive to the events of cue 1, cue 2, and response periods during match and nonmatch correct trials. Each bin is 100 ms. This unit increased firing to cue 1 and 2 for both trial types.

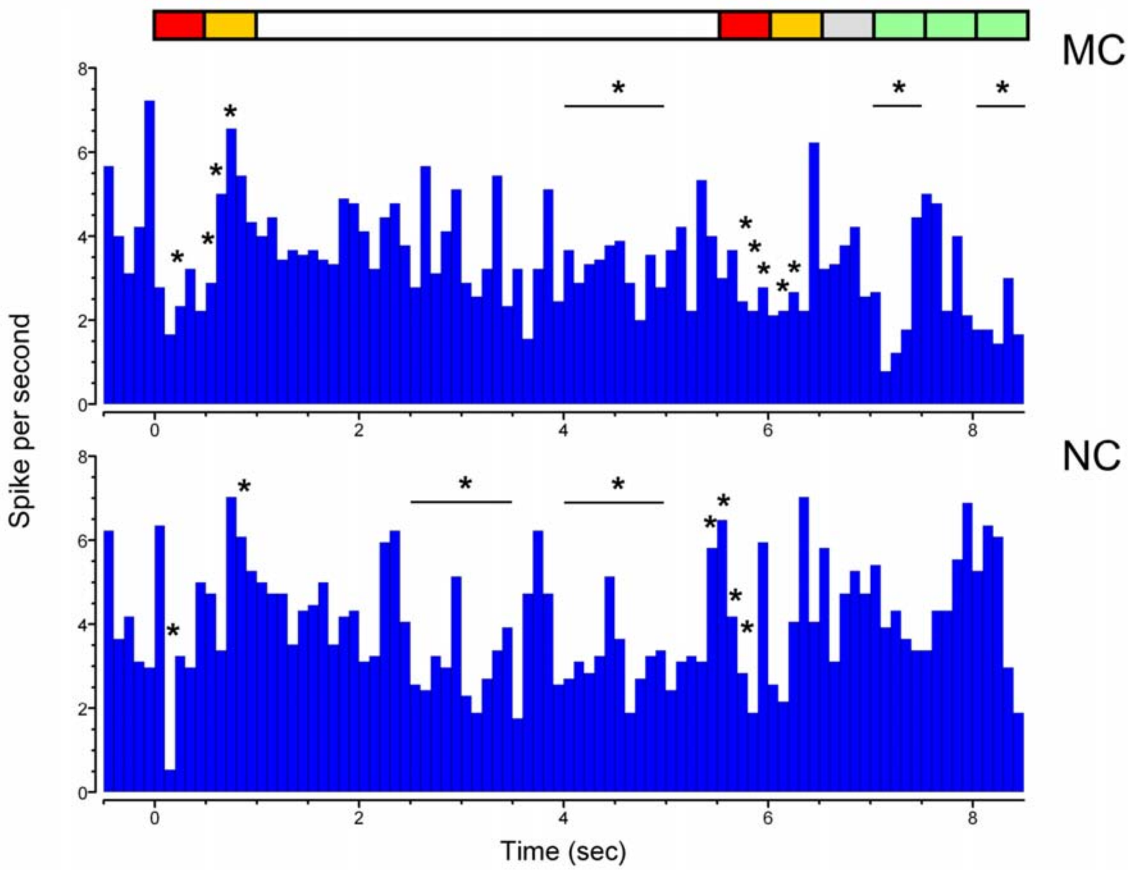


Figure 18. An example of a unit (# 1005091a) responsive to the events of cue 1, cue 2, delay and response periods during match and nonmatch correct trials. Each bin is 100 ms. This unit showed decreased firing to middle and late portions of memory delay in both trial types.

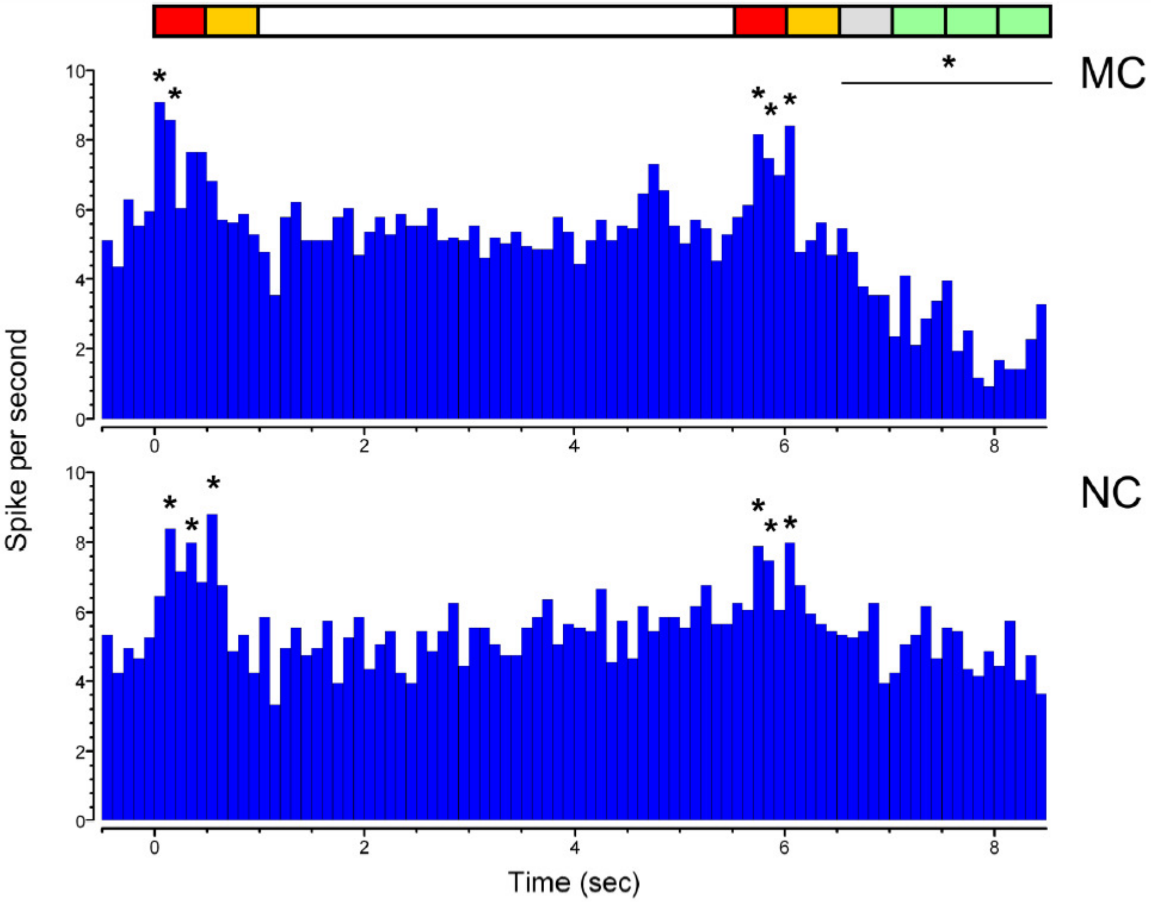


Figure 19. An example of a unit (# 1028092a) responsive to the events of cue presentations and offsets, wait and response periods during match and nonmatch correct trials. Each bin is 100 ms. Compared to nonmatch correct trials, this unit reduced firing from wait toward response periods 1 – 3 during match correct trials. This effect was modulated by press responses and food rewards.

During memory delays, 8 – 10% of dTP units showed activity change across the early, middle and late epochs (three epochs of 500 ms) of the memory delay period during match correct trials (Table 3). For nonmatch correct trials, about 8% of them showed activity change across the three memory delay intervals. Figures 16, 18 and 22 illustrate examples of dTP units with delay-related activity in each trial type.

Table 3. Percentage of units showing activity change during three epochs of the delay period for match correct and nonmatch correct trials (N = 225).

		Early	Middle	Late
	+FR	5	4	7
Match Correct	-FR	3	4	3
(MC)	Total	8	8	10
	+FR	4	3	2
Nonmatch Correct	-FR	3	5	5
(NC)	Total	7	8	7

During the 500 ms wait period, 17% and 8% of the recorded units showed activity change at match and nonmatch correct trials respectively. Particularly in match correct trials, more than half of the wait-related activity was due to an increase in firing rate against baseline, rather than a decrease (Table 4). During the 1500 ms response period, there were consistently more task-related units in match correct trials than those in nonmatch correct trials (21 – 28% versus 8 – 11%). This difference likely reflects the presence of press response and reward after the subjects successfully respond to the match condition. Figures 14, 19 and 24 show various examples of task-related activity during the wait and response period for correct trial types.

Table 4. Percentage of units showing activity change during the wait and response periods (R1, 2 and 3) for match correct and nonmatch correct trials (N = 225).

		Wait	R1	R2	R3
	+FR	12	14	16	12
Match Correct	-FR	5	12	12	9
(MC)	Total	17	26	28	21
	+FR	3	3	4	6
Nonmatch Correct	-FR	5	8	4	5
(NC)	Total	8	11	8	11

Incorrect Trials

Response profiles of dTP activity were examined for their responses during incorrect performance and these analyses were based on match incorrect trials (N = 204) and nonmatch incorrect trials (N = 225). Numbers of units with significant cue-evoked activity during presentation of cue 1 and 2 were similar between match incorrect and nonmatch incorrect trials (5 – 9%, see Table 5 and, Figures 20 and 21). This was also true when focusing on evoked activity during periods of cue 1 offset and cue 2 offset on the two incorrect trial types (6 – 7%, see Table 5). In general, the percentage of neurons responsive to cues during the match incorrect trials (5% for cue 1 and 8% for cue 2) tended to be lower than that during match correct trials (13% for cue 1 and 16% for cue 2). Similar to correct trials, evoked spike activity most likely increased rather than decreased in firing rate (see Tables 2 and 5). Cue-related activity was also rare during match and nonmatch incorrect trials due to both increased and decreased firing rates within cue presentations.

Table 5. Percentage of units showing activity change during cue events for match incorrect (N= 204) and nonmatch incorrect trials (N = 225).

		Cue 1	Cue 1 offset	Cue 2	Cue 2offset
Match Incorrect (MI)	+FR	4	5	7	5
	-FR	1	1	1	1
	Both	-	-	-	-
	Total	5	6	8	6
Nonmatch Incorrect (NI)	+FR	7	5	4	5
	-FR	2	2	2	2
	Both	-	-	-	-
	Total	9	7	6	7

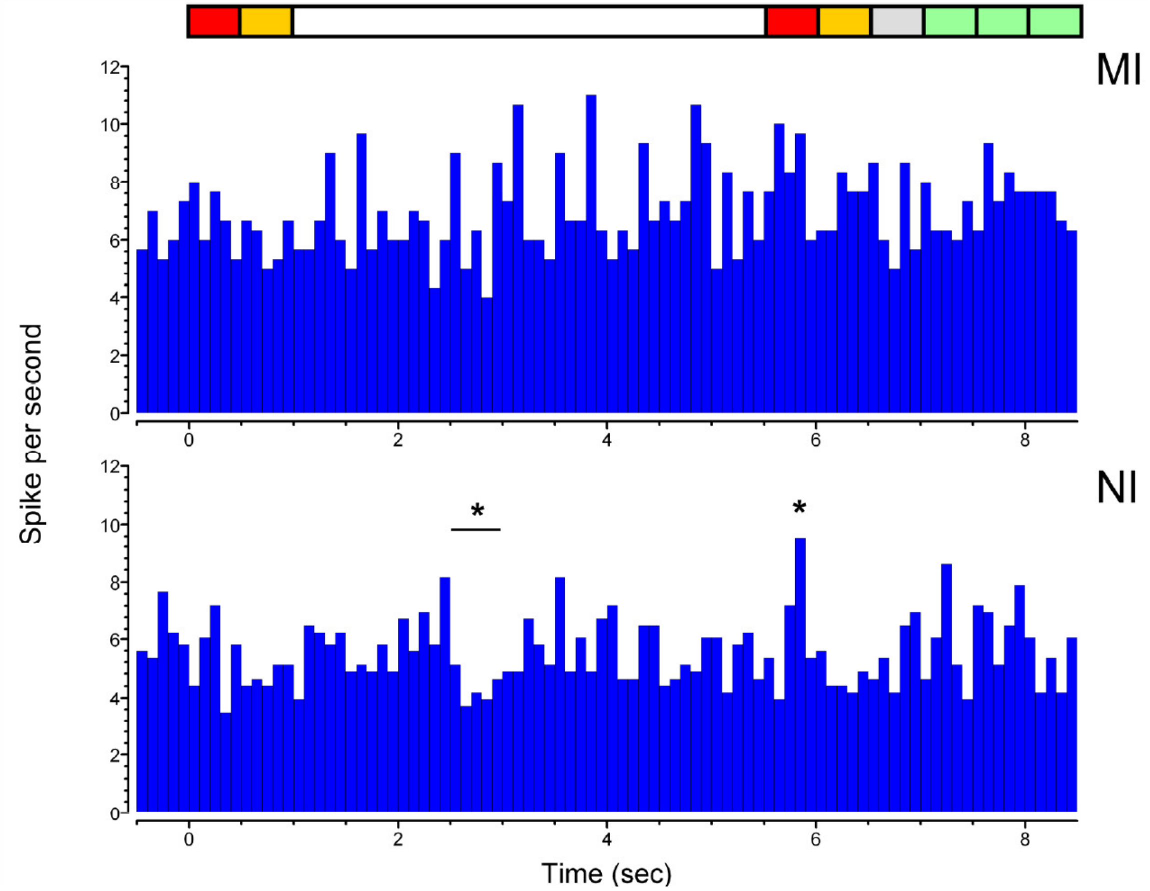


Figure 20. An example of a unit (# 0705102a) responsive to the events of cue 2 and delay during match and nonmatch incorrect trials. Each bin is 100 ms. This unit only increased firing to cue 2 during nonmatch incorrect trials. It also showed a transient delay-related activity. The unit was also responsive to the events of cue 1 and 2 during nonmatch correct trials.

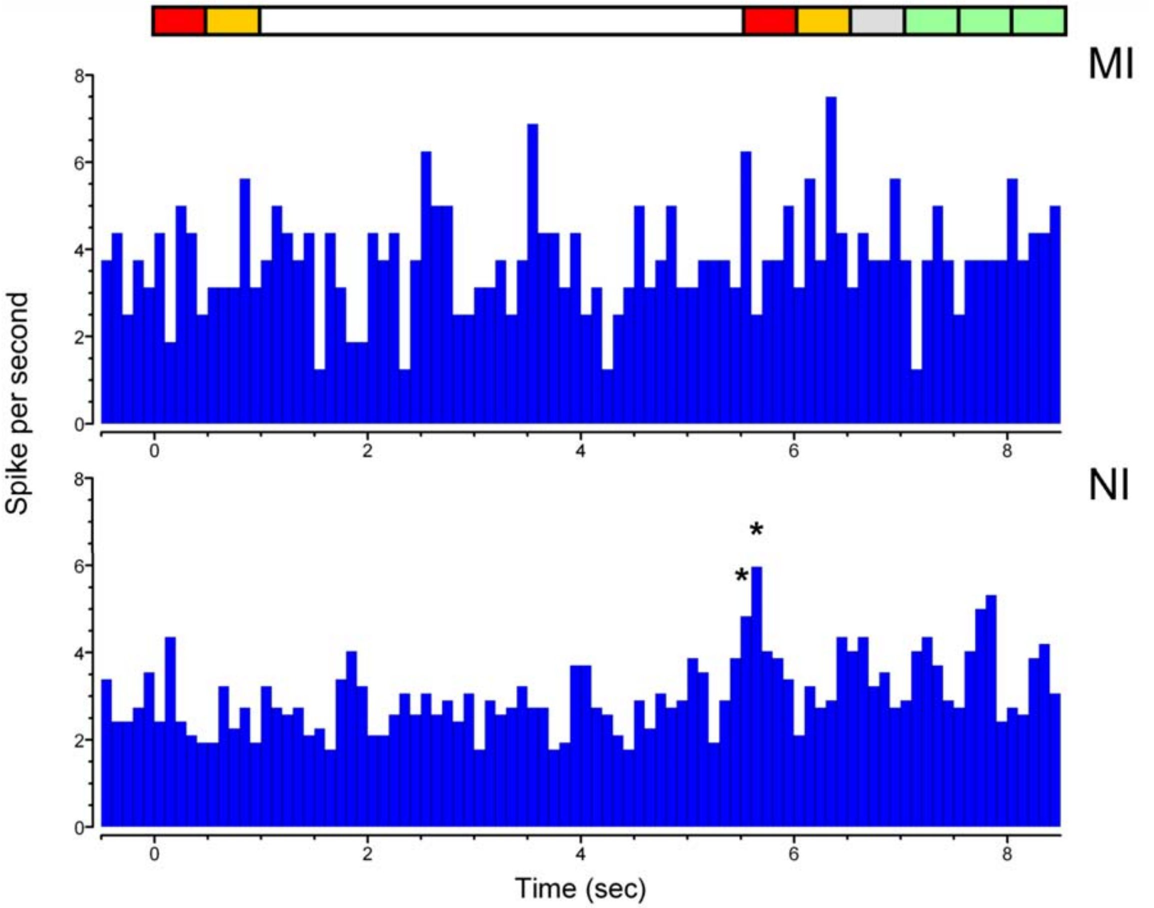


Figure 21. An example of a unit (# 0627101b) responsive to events of cue 2 during nonmatch incorrect trials. Each bin is 100 ms. Compared to nonmatch incorrect trials, spike activity was irregular with remarkable signal fluctuations during match incorrect trials.

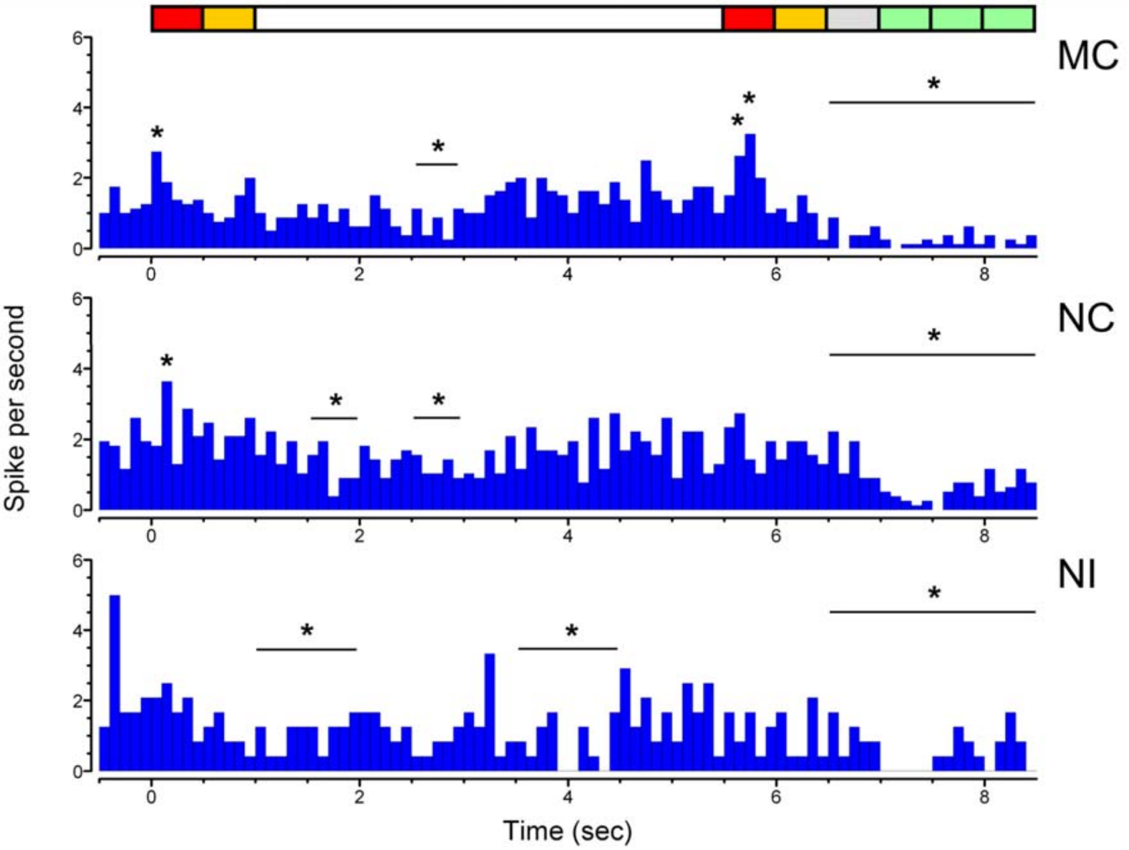


Figure 22. An example of a unit (# 0415102a) responsive to the events of cue 1, cue 2, delay, wait and response periods during match correct, nonmatch correct and nonmatch incorrect trials. Each bin is 100 ms. Reduced, intermittent activity was present in various epochs of early, middle and late delays.

Delay-evoked activity during match and nonmatch incorrect trials was evenly present across the three memory delay intervals (5 – 7%, see Table 6 and Figures 20 and 22). During the wait period, 5% and 7% of units showed event-related activity at match and nonmatch incorrect trials, accordingly (Table 7). Compared to match incorrect trials, monkeys produced erroneous button-presses during nonmatch incorrect trials. The number of units with event-related activity during nonmatch incorrect trials tended to be higher than the number of units encoding events during match incorrect trials across three response period intervals (8 – 13% on NI versus 5% on MI, Table 7). Figures 23 and 24 display examples of task-related activity during the wait and response periods of incorrect trials. Behavioral responses often modulated spike activity during match correct and nonmatch incorrect trials, but the presence of food rewards further influenced evoked activity on match correct trials only (Figure 24).

Table 6. Percentage of units showing activity change during three epochs of the delay period in match incorrect (N = 204) and nonmatch incorrect trials (N = 225).

		Early	Middle	Late
	+FR	2	2	3
Match Incorrect	-FR	3	3	3
(MI)	Total	5	5	6
	+FR	4	3	4
Nonmatch Incorrect	-FR	1	3	3
(NI)	Total	5	6	7

Table 7. Percentage of units showing activity change during the wait and response periods (R1, 2 and 3) in match incorrect (N = 204) and nonmatch incorrect trials (N = 225).

		Wait	R1	R2	R3
	+FR	1	1	2	1
Match Incorrect	-FR	4	4	3	4
(MI)	Total	5	5	5	5
	+FR	4	10	6	6
Nonmatch Incorrect	-FR	3	3	2	4
(NI)	Total	7	13	8	10

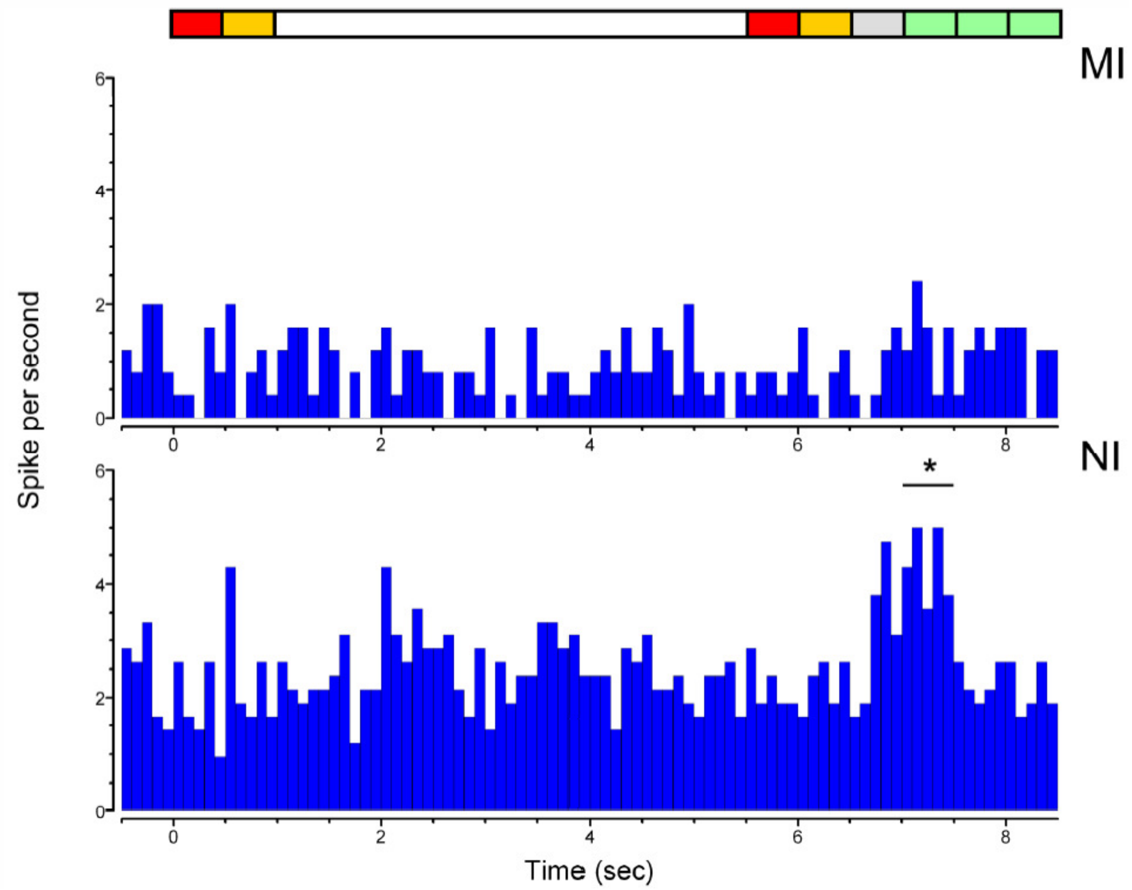


Figure 23. An example of a unit (# 0601101b) associated with button-press responses during nonmatch incorrect trials. Each bin is 100 ms.

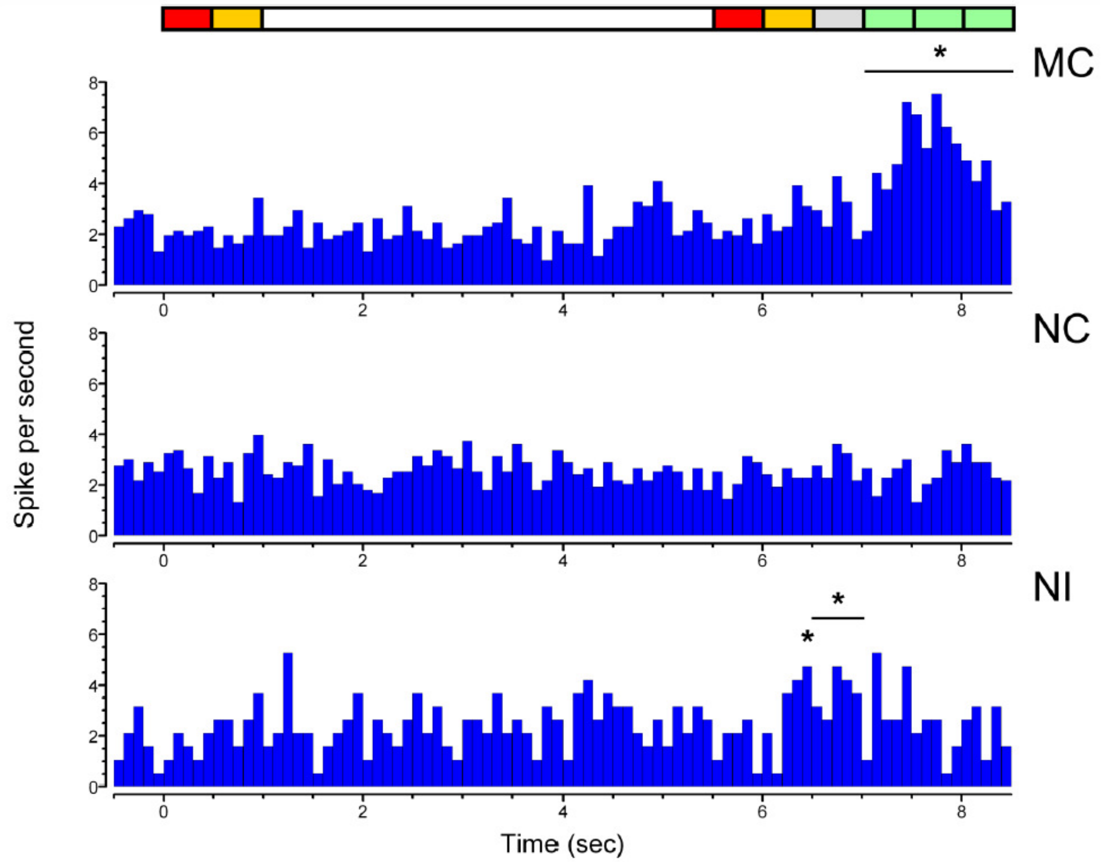


Figure 24. An example of a unit (# 0521101c) associated with wait and response periods during match correct, nonmatch correct and nonmatch incorrect trials. Each bin is 100 ms. The unit had sustained firing from cue 2 offset toward the wait period during nonmatch incorrect trials in contrast to sustained firing along the entire response interval during match correct trials. The effect may be due to behavioral outcomes associated with these two trial types (correct/error or reward/no reward).

Trial Conditions Associated with Button-Press

In order to assess how dTP activity correlates with button-press responses during match correct and nonmatch incorrect trials, firing rate was assessed in relationship to the exact timing of the button press. The intervals analyzed were before and after a button-press was recorded during four 500 ms intervals: pre-response period, post-response periods 1, 2 and 3. Spike data were rearranged and aligned with the first button-press produced during match correct trials (correct go-response) and nonmatch incorrect trials (erroneous go-response in no-go trials). Although both trial conditions involved button pressing, there were more responsive units present for match correct trials than for nonmatch incorrect trials. This difference was consistently present across the pre-response period, post-response periods 1 and 2 when comparing match correct trials to nonmatch incorrect trials (19 – 25% versus 9 – 11%; see Table 8 and Figure 25 for examples with activity change during the three intervals associated with button-press responses).

Table 8. Percentage of units showing activity change during the occurrence of button pressing (match correct and nonmatch incorrect trials; N = 225).

		Pre-response	Post-response 1	Post-response 2	Post-response 3
	+FR	10	13	13	11
Match Correct	-FR	9	10	12	10
(MC)	Total	19	23	25	21
	+FR	6	8	6	6
Nonmatch Incorrect	-FR	3	3	4	3
(NI)	Total	9	11	10	9

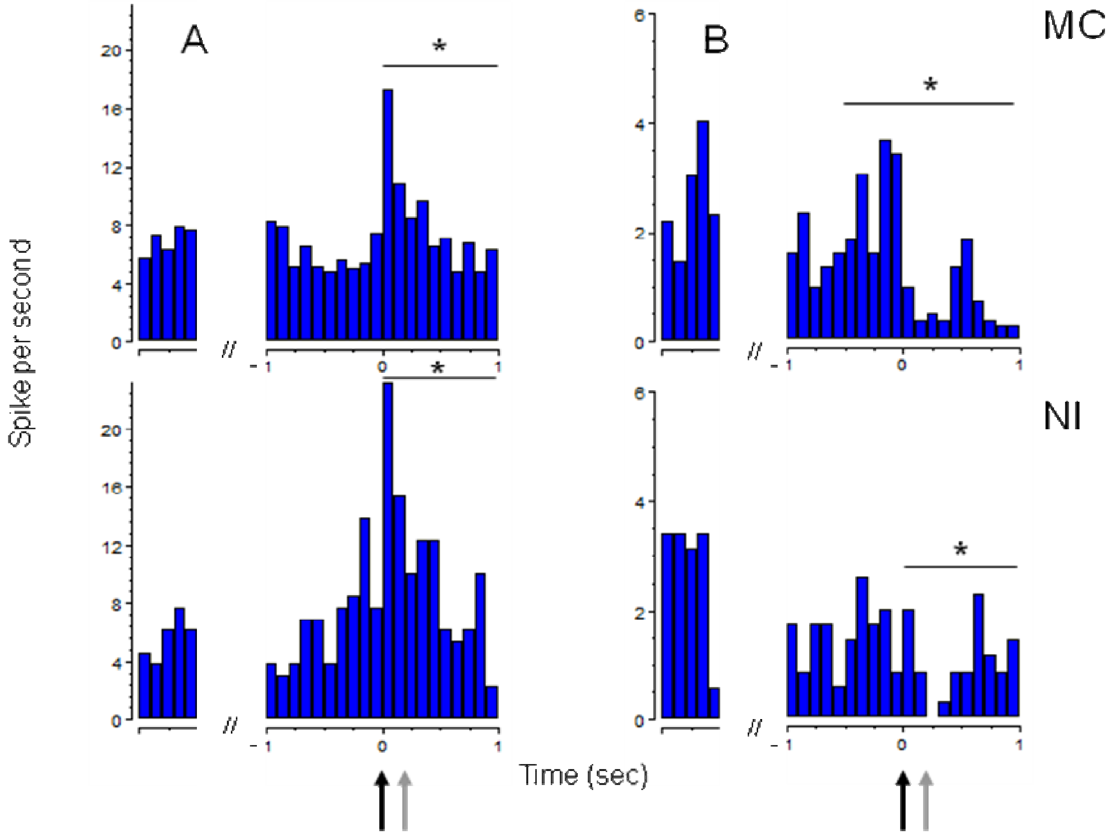


Figure 25. Two examples of units (# 0911011a & 0426101a) associated with button-press responses during match correct and nonmatch incorrect trials. Each bin is 100 ms. Button-press responses (black arrows) were followed by food rewards (grey arrows) at match correct trials, compared to when an erroneous response was made on nonmatch incorrect trials and no reward was presented. A food reward followed the first button-press during MC trials by approximately 200 ms. A. The unit showed a rapid increase in firing that immediately followed responses (at time = 0). B. The unit increased firing 500 ms before button-press responses, and then decreased firing during match correct trials. A combination of response and reward modulated the activity of this unit.

General Summary of 225 dTP Units during the Auditory DMTS Task

The summary reported here is based on the criteria that a particular unit was considered responsive to a target event if the unit yielded a significant activity change during the event compared to the baseline for at least one trial type. In general, more than a quarter of recorded units were responsive to presentations of cue 1 (27%) and cue 2 (29%); (Table 9). During cue periods, 36% of the recorded units were responsive to cue 1, cue 2 or both. Among these cue-responsive units, more than half of these units (54%) responded to both cue 1 and cue 2 periods, compared to those that responded to either cue only (46%). Approximately 6% and 9% of these cue-responsive units showed either enhanced or suppressed firing rate when comparing cue 1 to cue 2 at each trial type.

Table 9. Percentage of units showing activity change during cue events for at least one trial type.

	Cue 1	Cue 1 offset	Cue 2	Cue 2 offset
+FR	16	13	17	14
-FR	8	5	9	6
Both	3	2	3	4
Total	27	20	29	24

During memory delays, about 21% of recorded units were responsive during the three memory delay intervals for at least one trial type (Table 10). Among this sub-population, the majority of the responsive units showed significant activity change during

1 epoch of 500-ms interval in any of the three delay portions (i.e., early, middle and late), and these delay-related activity were short-lived, primarily lasting for 0.5 to 1 second (i.e., 1 – 2 epochs) (Table 11). During the wait period, 28% of units fired in one trial type or more, and over 30% of recorded units were responsive during the three response period intervals for at least one trial type (Table 11). Figure 26 provides a summary of the numbers of task-relevant events encoded by the 225 units of dTP. More than half the units encoded at least one out of the nine events during the memory task (i.e., cue 1, cue 1 offset, cue 2, cue 2 offset, delay, wait period, and response periods 1-3). The majority of these task-related units encoded one to three events (46%).

Table 10. Percentage of units showing activity change during three epochs of the delay period for at least one trial type.

	Early	Middle	Late
+FR	11	9	13
-FR	9	11	9
Total	20	20	22

Table 11. Percentage of responsive units showing delay-related activity across early, middle and late portions of fixed 5-second delays in which significant activity change is either present in 1, 2 or 3 successive epochs for each delay portion.

	Trial type	1 epoch	2 epochs	3 epochs
Early Delay	MC	29	11	6
	MI	39	6	-
	NC	50	8	4
	NI	33	4	4
Middle Delay	MC	34	11	3
	MI	39	6	-
	NC	50	12	-
	NI	38	8	-
Late Delay	MC	51	9	3
	MI	39	22	-
	NC	23	31	-
	NI	33	13	8

Table 12. Percentage of units showing activity change during the wait and response periods for at least one trial type.

	Wait	R1	R2	R3
+FR	16	21	22	19
-FR	12	18	15	14
Total	28	39	37	33

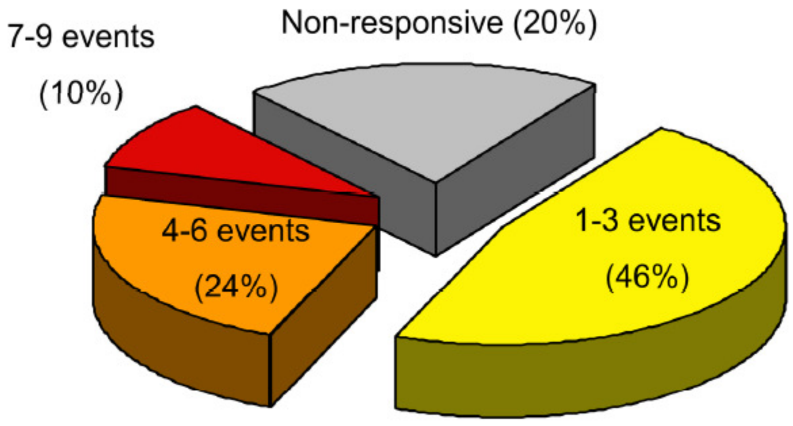


Figure 26. Distribution of task-related units encoding multiple events throughout the auditory DMTS task (N = 225).

Sound Evoked Activity during the Auditory DMTS Task

The neuron-by-neuron analysis for evoked activity change was based on spike data associated with presentations of cue 1 but not cue 2, since neuronal activity during cue 2 would not only be influenced by auditory stimulation, but could also be confounded by decision making (match or nonmatch rules) and perhaps anticipation of behavioral response (go or no-go response). Overall, more than one-third of the 225 units (34%) were responsive to at least one of the eight sound stimuli used in the memory task. When stimuli were organized into their respective sound types (animal, human, monkey, music, nature, pure tone, synthesized clip and white noise), a low percentage of dTP units was active across all eight sound types (3 – 9%, Table 12). These findings suggest that these auditory responsive units did not seem to exhibit firing preferences on certain sound types over others during the memory task. Among these auditory responsive units, 83%

of responded to only one out of the eight sound stimuli used during the DMTS task. This finding suggests that these units were greatly selective to particular sounds during the memory task (Figure 27).

Table 13. Percentage of units responsive to a sound stimulus of the eight sound types during the auditory DMTS task.

	+FR	-FR	Both	Total
Animal	3	-	-	3
Human	7	2	-	9
Monkey	4	1	-	5
Music	4	1	-	5
Nature	4	3	-	7
Pure Tone	5	2	-	7
Synthesized Clip	3	1	-	4
White Noise	5	2	1	8

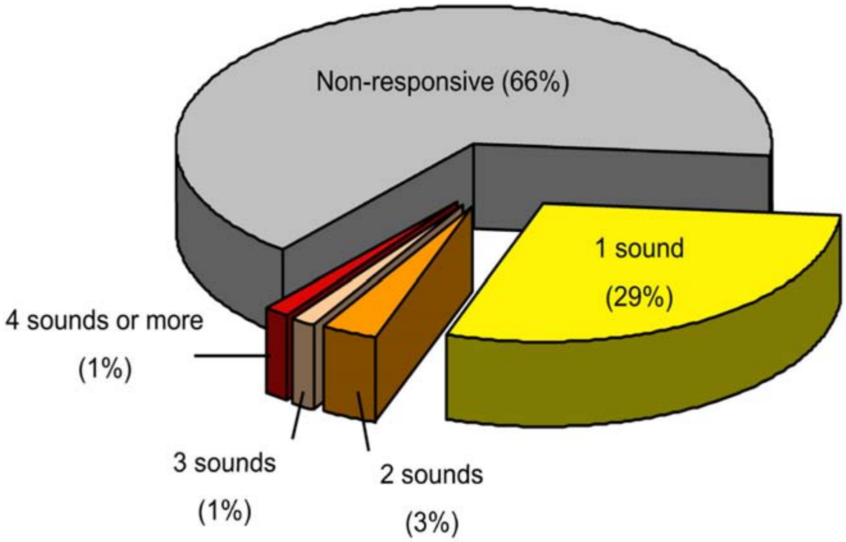


Figure 27. Distribution of auditory-responsive units encoding multiple sounds throughout the auditory DMTS task.

Population Analysis of Spike Activity across Task-Relevant Events

The population results (one-way ANOVAs, $p < 0.05$) are based on 225 recorded units during match correct trials, nonmatch correct trials and nonmatch incorrect trials. The match incorrect trials were excluded from population analyses as nearly 80% of total units associated with this trial type were non-responsive to any task-relevant event. The current subjects have been trained in this memory task for several years and perform well during match trials (78% correct on average). Number of trials and spike activity associated with the match incorrect condition is limited by good behavioral performance. Spike firing during match incorrect trials seems to be low and irregular across trials within a session, compared to the other three trial types. Thus, interpretation of results on this trial type needs to be cautious, and the in-depth analysis only emphasizes spike activity of dTP on match correct, nonmatch correct and nonmatch incorrect trials.

For match correct trials, population activity was significantly higher than the baseline level at the first 100 ms interval during presentations of cue 1, and was significantly lower than the baseline level at the fourth 100 ms interval during cue 1 (Figure 28, top). The population activity was also significantly higher than the baseline level at the last 100 ms interval of cue 2 offset, which was immediately followed by the wait period. Evoked activity during cue 1 offset and cue 2 was similar to the baseline level. For nonmatch correct trials, population activity was significantly higher than the baseline level at the first 100 ms interval during presentations of cue 1 and cue 2 (Figure 28, center). Evoked activity was significantly lower than baseline at the last 100 ms interval of cue 1. For nonmatch incorrect trials, increased spike activity was present at the first 100 ms interval of cue 2 only (Figure 28, bottom). In short, population analyses of dTP units revealed discrete trial-type differences of activity change during events of cue 1 and cue 2. Evoked activity change of the 225 units was associated with the involvement of behavioral responses and reward during wait and response periods. During match correct trials, increased spike activity was present for all the intervals for the wait and response periods. For the two nonmatch trial types, significant activity change was found at two out of three epochs of the response period. On the contrary, there was an absence of significant activity change for any portion of the memory delay at any trial type.

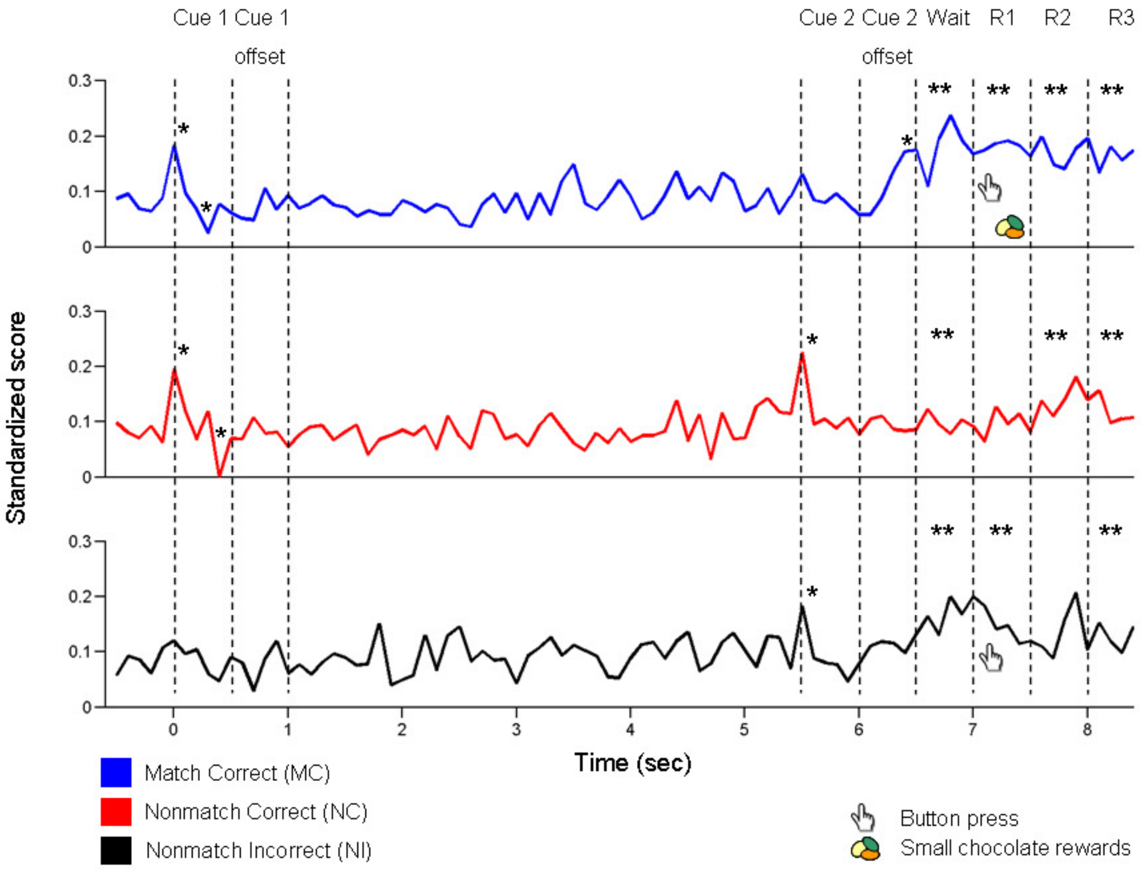


Figure 28. Population activity of 225 units across discrete task events for the three trial types. Single asterisks denote significant activity change against baseline during a 100 ms interval of cue presentation or cue offset. Double asterisks denote significant activity change against baseline during a 500 ms interval of a target event (i.e., wait time, response periods R1, R2, and R3). Normalized spike data shown on the graph was binned into 100 ms intervals. Note that button-press responses were present during response period R1 of MC trials (correct go-response), which were followed by a food reward. Behavioral responses were also present during response period R1 of NI trials (erroneous go-response).

Population Analyses of Trial Types among Discrete Task Events

Sound Presentations of Cue 1 and Cue 2

Effects of the three trial types at cue and cue offset periods were analyzed by repeated-measures ANOVAs, in which within-subject factors were cue (i.e., cue 1 and 2 or cue 1 offset and cue 2 offset) and bin (5, 100 ms intervals), and the between-subject factor was trial (match correct, nonmatch correct and nonmatch incorrect). A main effect of bin was found ($F_{(4, 26988)} = 20.26$, $p < 0.01$) in that population activity at the first 100 ms interval was always the highest compared to the remaining 4 intervals, regardless of cue position. A significant trial * cue * bin interaction was also present ($F_{(8, 26988)} = 2.61$, $p < 0.01$). At the fourth 100 ms interval of cue 1, population activity of match correct trials was higher than that of nonmatch correct trials and the effect was reversed at the fifth 100 ms interval. During cue 2, the critical decision point, population activity of nonmatch correct trials was significantly higher than that of match correct trials at the first 100 ms interval of cue 2.

The study further examined effects of trial type associated with activity change during presentations of cue 1 and 2. The repeated-measures ANOVAs for cue presentations demonstrated that the critical difference between the two correct trial types was an absence of increased firing at cue 2 during match correct trials, particularly at the first 100 ms interval of cue presentation (Figure 29). The next question was to ascertain the latency of this trial type effect. Among 225 units, population activity during the first 90 ms was re-sampled into three 30 ms intervals for detailed analyses. Population data for match and nonmatch correct trials during cue 1 presentation were combined (Cue 1 - MCNC) as they had similar increases in firing rate. The combined cue 1 data was then compared to match and nonmatch correct trials during cue 2 (Cue 2 - MC and Cue 2 - NC) with paired-sample *t*-tests. At the second 30 ms bin of cue presentation (30-60 ms from cue start), there was higher activity to the combined cue 1, and cue 2 during nonmatch correct trials compared to the match correct trials of cue 2 (Figure 30). No trial difference was present at the first or last 30 ms interval analysis. These findings suggest that trial differences regarding the activity change during presentations of cue 2 occurred within 30 – 60 ms of the cue 2 start.

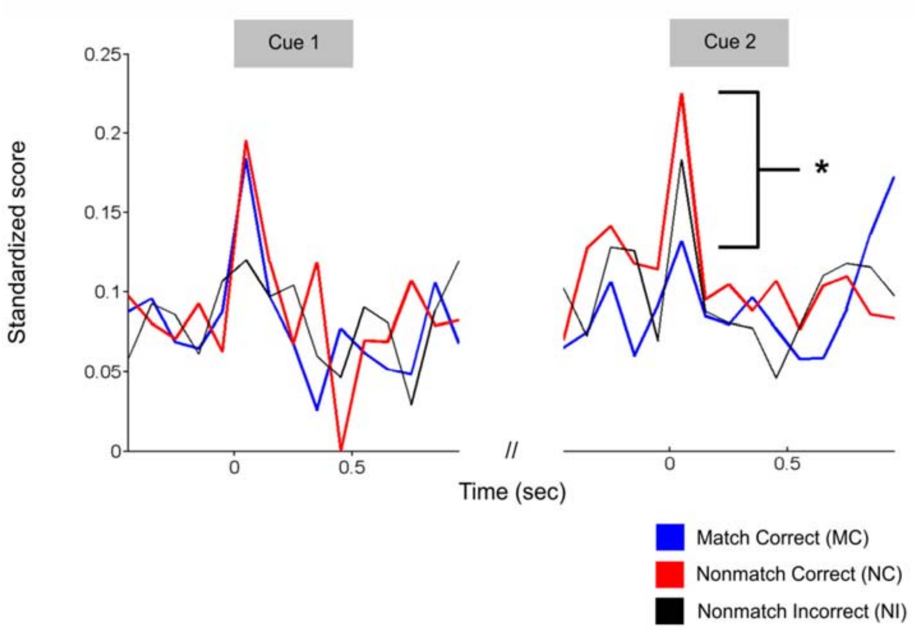


Figure 29. Trial effects associated with spike activity change during the presentations of cue 1 and 2. Here, population spike activity of the three trial types was emphasized during cue presentations. The bracket and asterisk denote a significant difference between MC trials (blue lines) and NC trials (red lines) during the 1st 100 ms interval of cue 2.

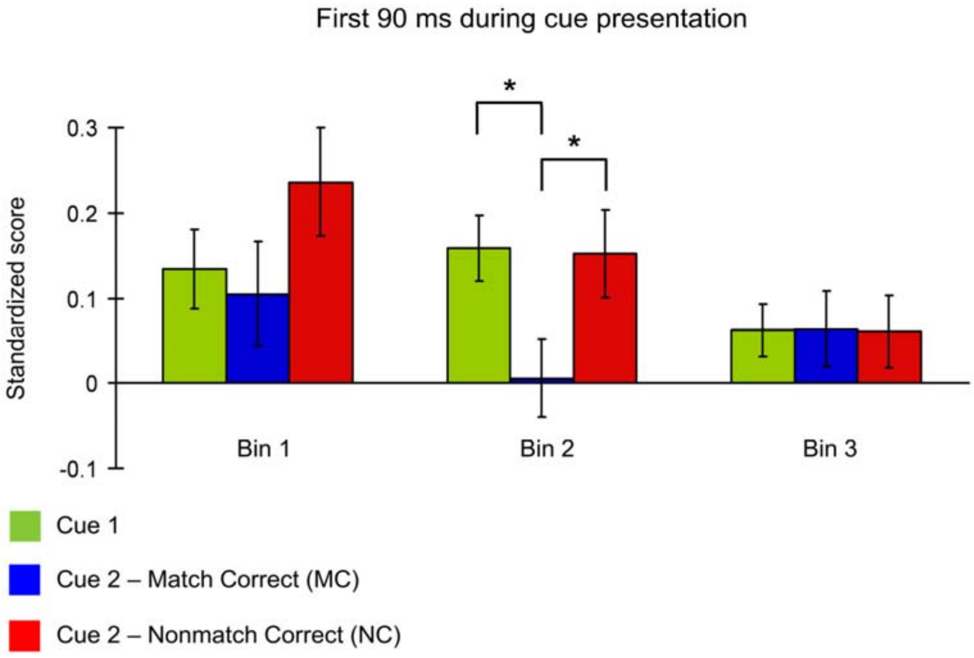


Figure 30. Detailed analysis of population activity during cue 2. The brackets and asterisks denote a significant trial difference in discrete 30 ms intervals. Results suggest that active suppression may be present during presentations of cue 2 during match correct trials. The effect was very early during the cue 2 presentation, within 60 ms from cue onset.

When focusing on population response during cue 1 and 2 offset, there was a main effect of bin ($F_{(4,\ 26988)} = 2.37$, $p < 0.05$) and cue ($F_{(1,\ 26988)} = 7.74$, $p < 0.01$). Population activity during the last two 100 ms intervals was higher than that of the first interval of the general offset period. Population activity during cue 2 offset was always higher than that during cue 1 offset, regardless of trial types. The later finding was potentially related to the fact that in two out of the three trial types, the cue 2 offset period was immediately followed by the wait period preceding a button-press response. Therefore, a detailed analysis between trial types was used to clarify trial differences at cue 2 offset. Parallel to the detailed analysis of cue 2, all spike data during events of cue 2 offset was re-sampled into 30 ms time bins during match correct trials, nonmatch correct trials, and nonmatch incorrect trials. Significant trial effects existed during the last 90 ms intervals of cue 2 offset (Figure 31). Population activity at match correct trials was higher than the two nonmatch trial types during the first and last 30 ms intervals. Although behavioral responses were involved in both match correct trials and nonmatch incorrect trials, the findings suggest that population of dTP units showed activity difference between the two trial types during the event of cue 2 offset.

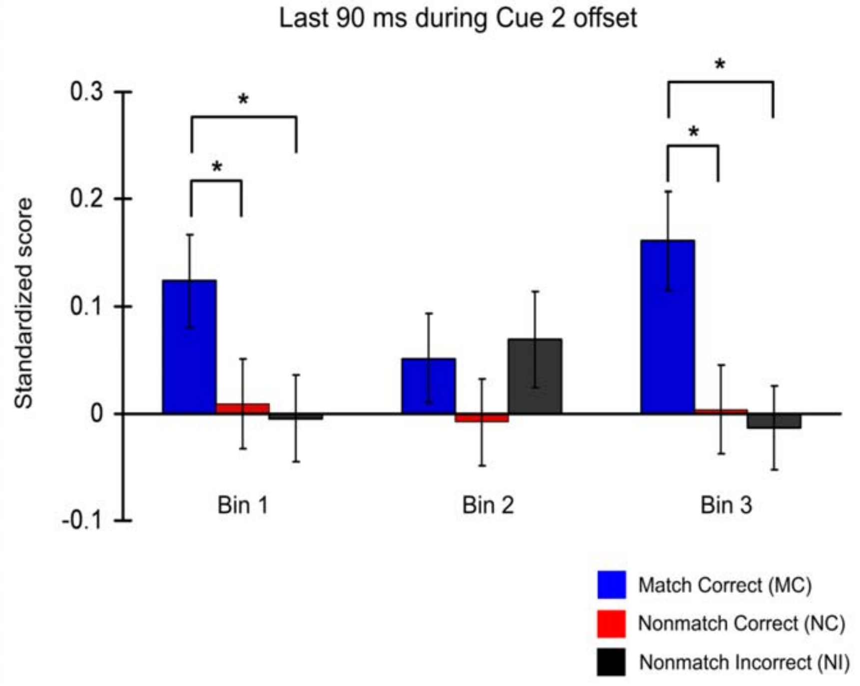


Figure 31. Detailed analysis of population activity during cue 2 offset. The brackets and asterisks denote a significant trial difference in discrete 30 ms intervals. The findings suggest that spike units during MC trials generally fired more often than during the other two trial types within the 90 ms interval immediately preceding the wait time.

Wait and Response Periods

After processing the cue 2 presentations during the cue and offset period, the subject was required to wait to make a response and neuronal activity was also recording during the possible response and reward periods. Four events (wait, and response periods 1, 2 and 3) and trial type were used as within-subject and between-subject factors respectively for ANOVAs. There was a main effect of trial type ($F_{(2, 13797)} = 17.91$, $p < 0.01$) showing that the population activity of match correct trials and nonmatch incorrect trials were higher than that of nonmatch correct trials (Figure 32). There was also a trial * event interaction ($F_{(6, 41391)} = 2.32$, $p < 0.05$) showing that population activity during match correct trials and nonmatch incorrect trials at the wait period and response period 1 were higher than that of nonmatch correct trials. Spike data was then aligned to the first button-press produced during match correct and nonmatch incorrect trials. No trial difference was present between the two trial types in any of the four 500 ms events (Figure 33). Since trial effects occurred as early as the wait period and were sustained throughout the three response periods, population activity during wait time was further analyzed into 100 ms intervals by repeated-measures ANOVAs. A main effect of trial ($F_{(2, 6747)} = 12.93$, $p < 0.01$) confirmed that population activity of match correct trials and nonmatch incorrect trials were higher than that of nonmatch correct trials. A trial * bin interaction ($F_{(8, 26988)} = 2.16$, $p < 0.05$) revealed that population activity at match correct trials was higher than that at nonmatch correct trials at all five 100 ms intervals, while population activity at nonmatch incorrect trials was higher than nonmatch correct trials only at the fourth 100 ms interval. Trial effects on population activity within the wait period temporally differentiated match correct trials and nonmatch incorrect trials, even if behavioral responses were both associated with these two trial types.

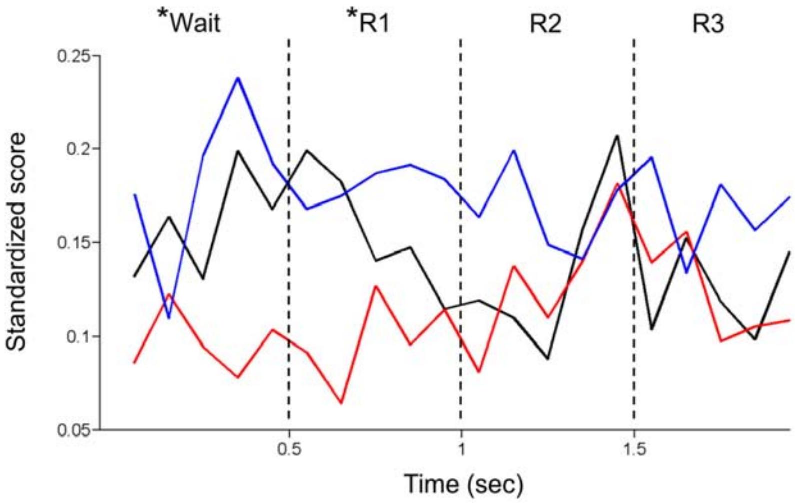


Figure 32. Trial effects associated with population activity change during wait and response periods. Each interval was 500 ms long. Button-press responses were followed by food rewards at MC trials, compared to when an erroneous response was made on NI trials and no reward was presented. Asterisks denote a significant increase in firing rate for a particular 500 ms interval during MC and NI trials (blue and black lines), relative to NC trials (red line). Results show that spike activity at MC trials is higher than at NC trials occurring as early as the 1st 100 ms interval, while NI trials are higher than NC trials at the 4th 100 ms interval.

Memory Delay Periods

Although population activity was not shown to be significant from baseline level during the memory delay period for each of the three trial types individually, trial differences between the three trial types for population response during the memory delay period was revealed in the study. Within-subject factors here were three delay epochs (early, middle and late) and bin (three 500 ms interval). A main effect of trial ($F_{(2,\ 13797)} =$ 23.94, $p < 0.01$) showed that population activity of nonmatch incorrect trials was higher than those of the two correct trial types during the entire delay period. A main effect of delay epoch ($F_{(2, 27594)} = 3.53$, $p < 0.05$) revealed that spike activity during the late delay epoch was higher than that during the middle delay portion.

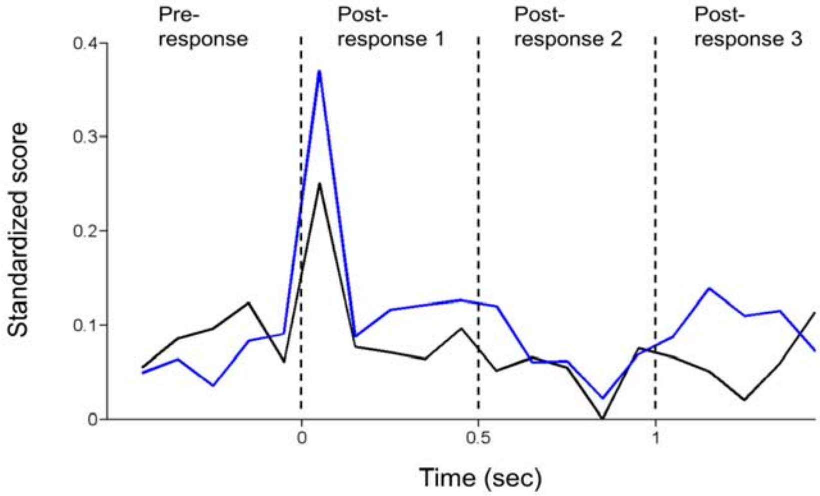


Figure 33. Population activity changes during button-press responses. A food reward followed the first button-press during MC trials by approximately 200 ms. Spike activity for each trial was aligned when the first button-press was produced. There was no trial difference between MC (blue line) and NI (black line) across all four 500 ms events, but reward related activity was observed at the single unit level.

Sound Evoked Activity of dTP during Passive Listening

The passive listening experiment was divided into two parts. During passive listening 1, eight different sounds (the target folder, one per sound type), soon to be used for the auditory DMTS task, were initially presented to monkey subjects, followed by a block of 21 pure tones (250 ms long per sound). After completing the memory task, the remaining 11 sound folders (eight sounds per folder, one per sound type) from the eight sound types were presented to subjects until the isolated unit was no longer present (passive listening 2) or the sounds had all been presented at least 8-12 times. When isolated units were still active after passive listening 2, a folder of 12 additional monkey vocalizations was then presented to the animal subjects. Results are reported according to the placement of the passive listening experiments.

Passive Listening 1

Among the recorded 176 units during passive listening 1, nearly a third of them (30%) were responsive to at least one of the eight sounds of a target folder. These auditory responsive units fired in a similar manner to the eight sound types (5 – 9%), and firing rate tended to increase during stimulus presentations, similar to the activity reported above for the DMTS task (Table 13). These units were sound-selective as the majority responded to one stimulus only (Figure 34).

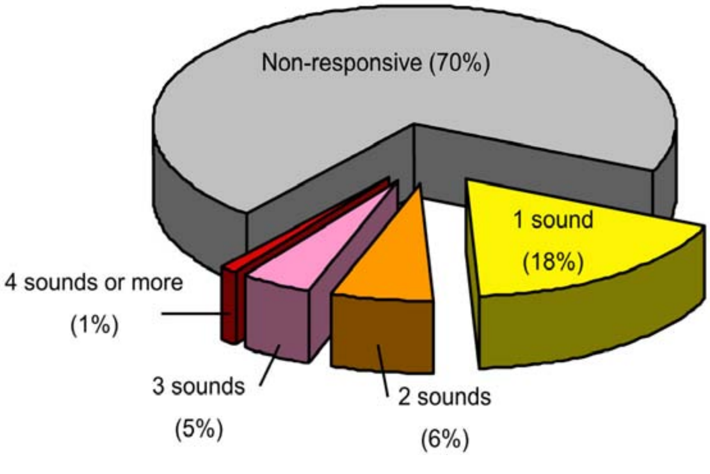


Figure 34. Distribution of auditory-responsive units encoding multiple sounds throughout passive listening 1 for a target folder (N = 176).

Table 14. Percentage of units responsive to a sound stimulus of the eight sound types during a target folder of passive listening 1 (N = 176).

	+FR	-FR	Both	Total
Animal	7	1	1	9
Human	4	1	-	5
Monkey	7	1	-	8
Music	7	2	-	9
Nature	5	-	-	5
Pure Tone	5	2	-	7
Synthesized Clip	6	1	-	7
White Noise	5	2	-	7

When presenting 21 pure tones, 35% were responsive to at least one pure-tone stimulus. Table 14 illustrates a fairly uniform response profile of dTP units across the 0.1 – 20 kHz frequency range. Again, the units were sound-selective within the pure tone presentations in that the majority responded to one stimulus only (Figure 35). Figure 36 illustrates examples of eight units responding to two pure tones or more. These eight units together responded to 16 out of 21 pure tones presented during passive listening 1. In general, the units that responded to multiple pure tones were evoked in a distributed manner across the frequency range of 0.1 – 20 kHz. There were intervening frequencies between the ones effectively evoking spike activity for a given unit. It appears that these distributed, firing patterns, shown by a few units, could cover a wide range of frequency bandwidths.

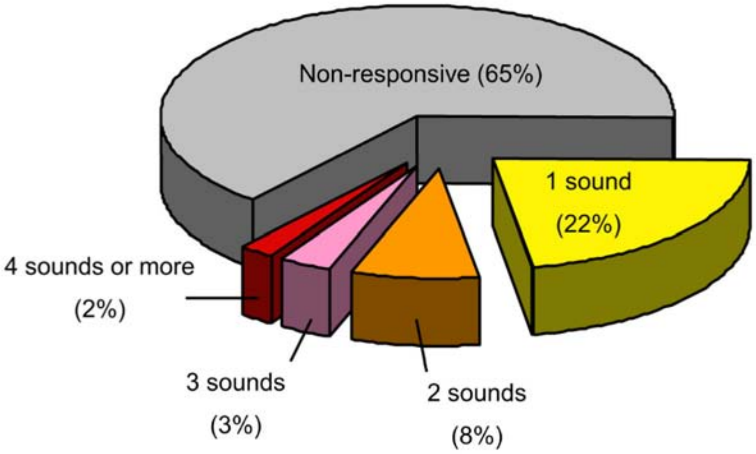


Figure 35. Distribution of auditory-responsive units encoding multiple pure tones throughout passive listening 1 (N = 176).

Table 15. Percentage of units responsive to a sound stimulus of the 21 pure tones during passive listening 1 (N = 176).

Frequency (Hz)	+FR	-FR	Both	Total
100	2	1	1	4
200	3	-	-	3
300	2	-	-	2
400	3	1	-	4
500	2	1	-	3
600	1	-	-	1
700	3	1	-	4
800	2	-	-	2
900	3	-	-	3
1000	3	-	-	3
2000	2	-	-	2
3000	2	1	1	4
4000	2	1	-	3
5000	2	-	-	2
6000	2	-	-	2
7000	3	-	-	3
8000	2	-	-	2
9000	3	1	-	4
10000	2	1	-	3
15000	3	1	-	4
20000	2	-	-	2

Figure 36. Examples of eight units (1-4: page 112; 5-8: page 113) responsive to various pure tones ranging between 0.1 and 20 kHz. Asterisks indicate a significant activity change from baseline at a particular pure-tone frequency. Note that 16 out of 21 pure tones were encoded by just these eight units.

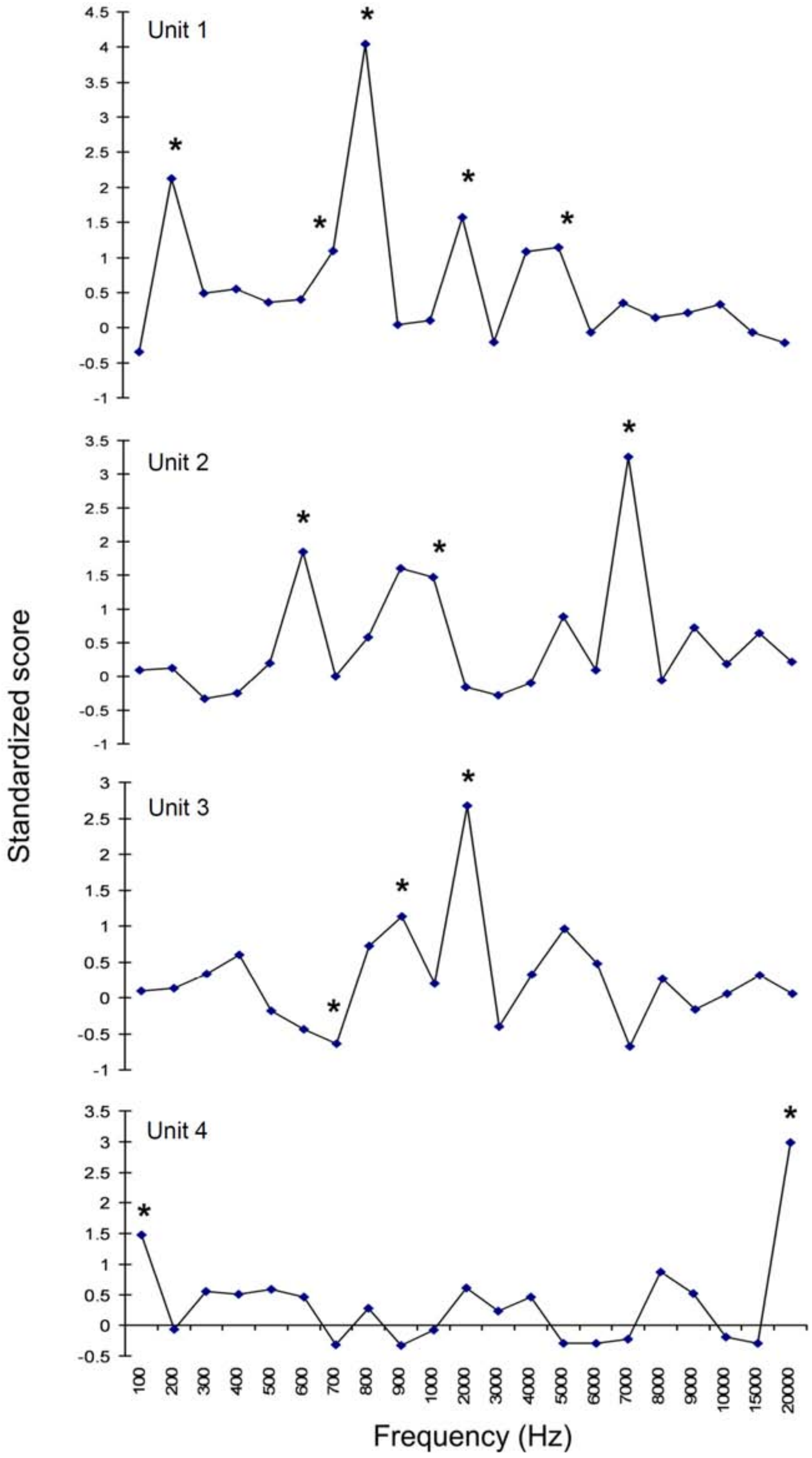
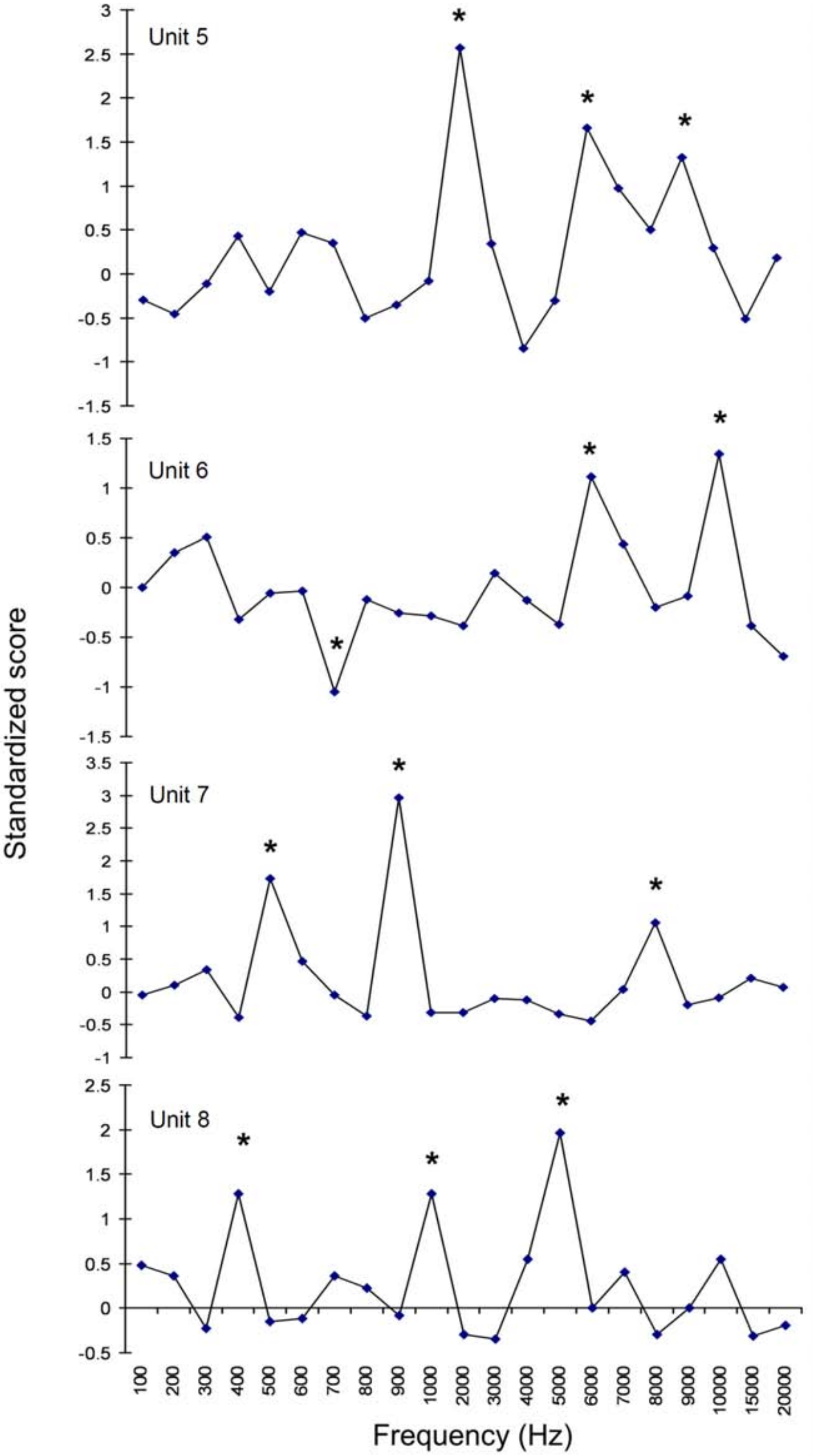


Figure 36 --- continued



Passive Listening 2 and Additional Monkey Vocalizations

There were 161 units maintained after the auditory memory task and analyzed. Over two-thirds of the units (69%) were responsive to at least one out of 96 standard stimuli. Similar to passive listening 1, these units mostly responded to one to four discrete stimuli (Figure 37), and 6% of recorded units responded to more than 9 sounds. Table 15 summarizes response profiles of dTP units responsive to sounds, separated by sound types.

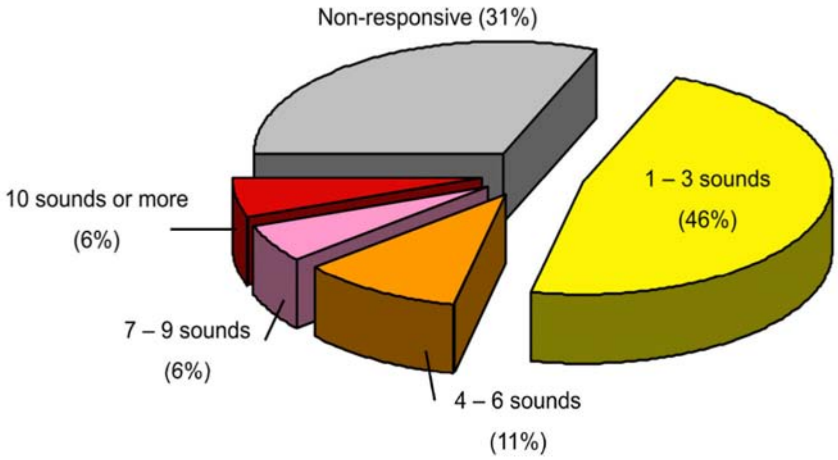


Figure 37. Distribution of auditory-responsive units encoding multiple sounds throughout passive listening 2 (N = 161).

After passive listening 2, 124 units were maintained when subjects listened to 12 additional monkey vocalizations, which were not part of the pool of 96 sounds. About one third of the recorded units (33%) were responsive to at least one of the 12 monkey vocalizations. Table 16 summarizes units of dTP responsive to the 12 monkey vocalizations, separated by call types. Similar to passive listening 1 and 2, these units were sound-selective as 60% of them responded to one stimulus only.

Table 16. Percentage of units responsive to the eight sound types during passive listening 2 (N = 161).

	+FR	-FR	Both	Total
Animal	24	8	1	33
Human	26	3	-	29
Monkey	27	8	-	35
Music	26	5	1	32
Nature	24	3	3	30
Pure Tone	40	5	-	45
Synthesized Clip	29	6	-	35
White Noise	36	5	1	42

Based on passive listening 1 and 2, most auditory responsive units in dTP only responded to one to three out of 96 sounds, and were therefore considered selective. Figure 38 illustrates examples of units responsive to sounds during various experimental contexts.

Table 17. Percentage of units responsive to the additional 12 monkey vocalizations (N = 124).

	+FR	-FR	Both	Total
Coo 1	3	2	-	5
Coo 2	3	2	-	5
Coo 3	2	1	-	3
Coo 4	2	-	-	2
Grunt 1	2	3	-	5
Grunt 2	4	2	1	7
Harmonic Arch 1	1	2	-	3
Harmonic Arch 2	3	-	-	3
Scream 1	3	2	-	5
Scream 2	3	-	-	3
Scream 3	6	2	-	8
Other	3	3	-	6

Figure 38. Examples of evoked activity change across various sound stimuli between experimental contexts. Units A (# 0829101a) and B (# 0628101a) were responsive to various sounds during the memory task and passive listening 2. Asterisks denote a significant activity change from baseline at that 100 ms interval. Grey bars represent sound presentations (500 ms duration). Unit A increased and decreased firing rate within the same sound (‘water plops’ and ‘Thunder’). Unit B not only increased firing rate to certain sounds (e.g., ‘Plucked string’ and ‘Bell’) but also decreased firing rate to others (e.g., ‘Coo 2’ and ‘Roar of fire’). Unit C (# 1028091c) was responsive to the same sound (human voice ‘Boo’) as well as other sounds during the memory task and passive listening 1 and 2. Note that unit C responded at different 100 ms intervals to the same sound across different experimental contexts (C1, 2 and 3).

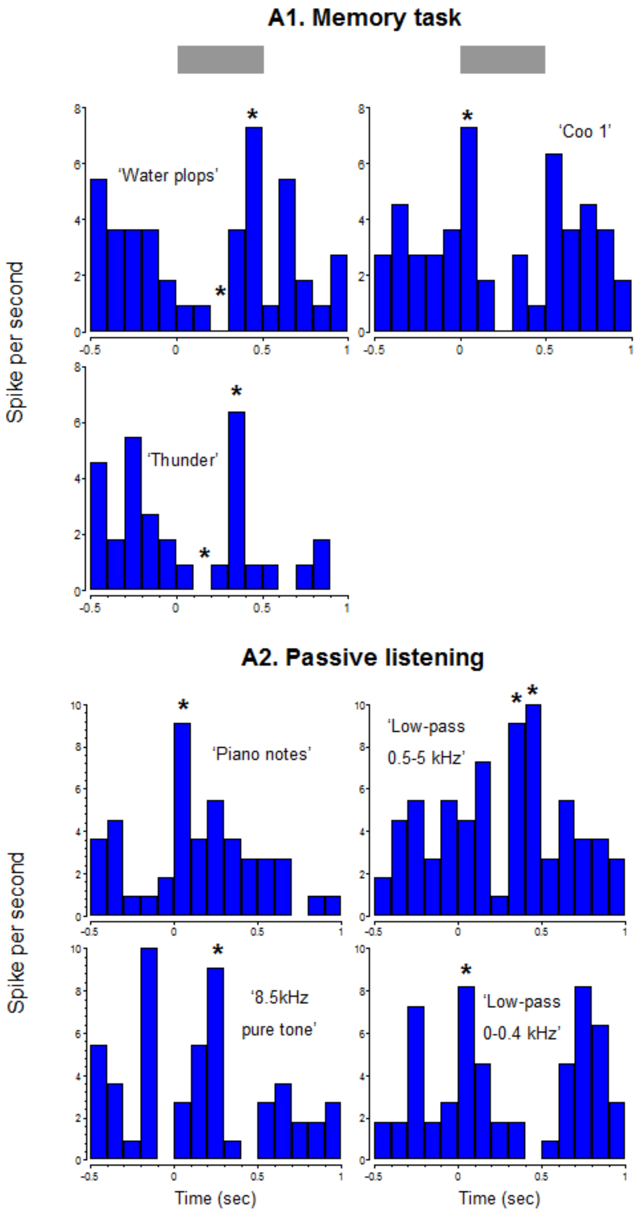


Figure 38 --- continued

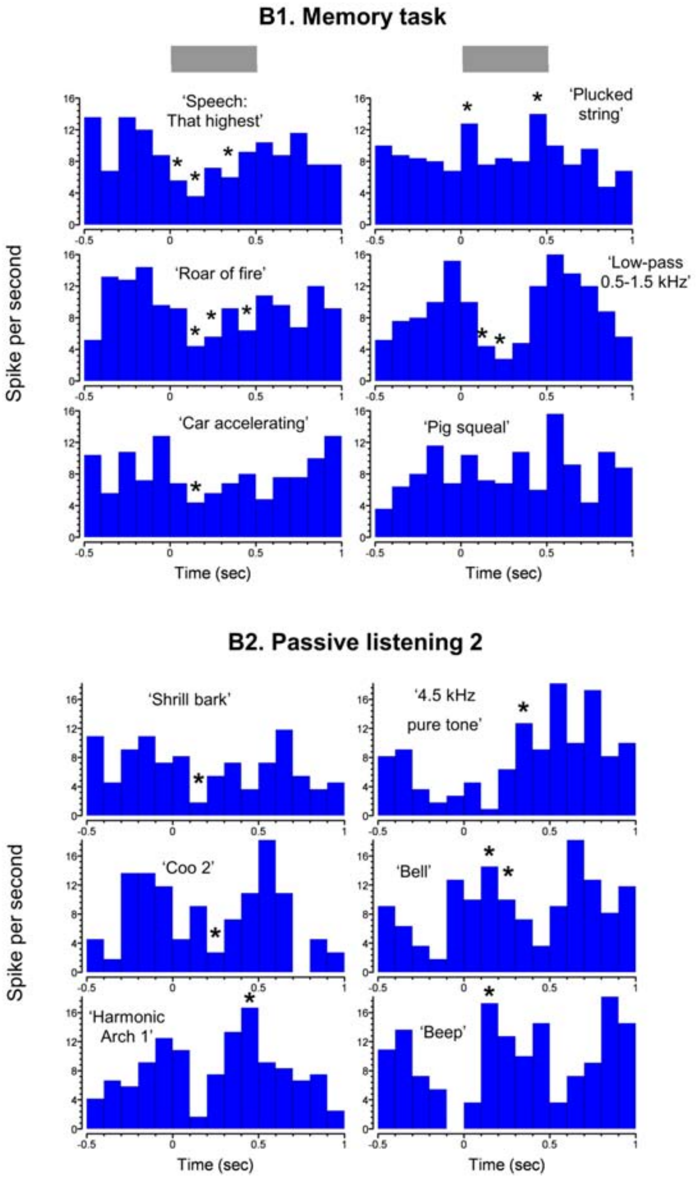
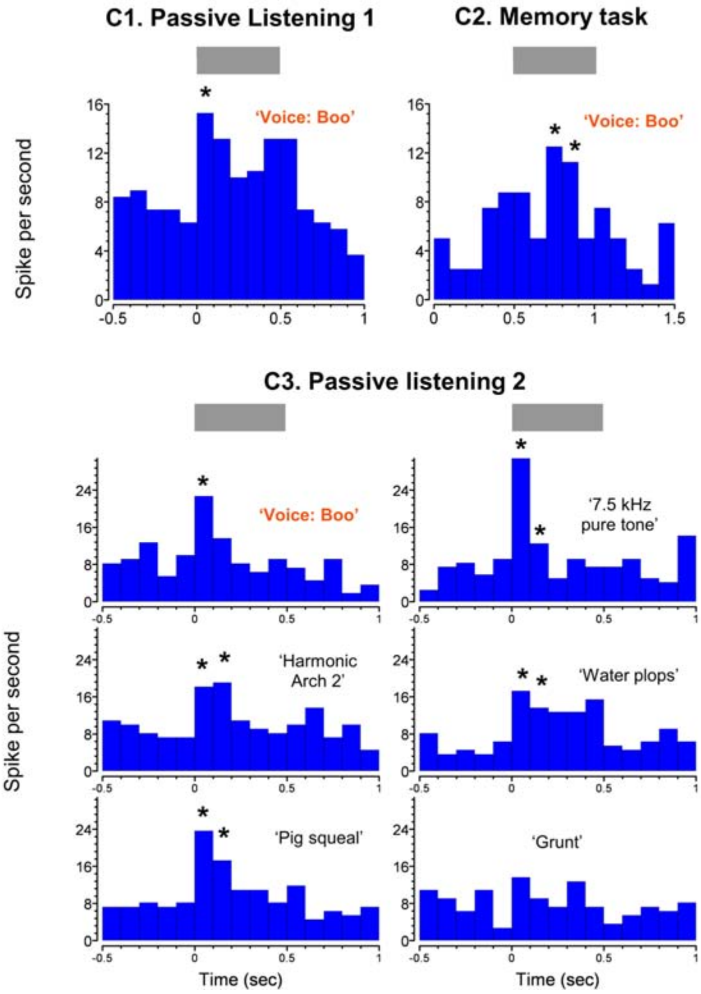


Figure 38 --- continued



Various factors were evaluated to assess how they might contribute to sound responsiveness and these analyses required that the units to be analyzed with these methods be responsive to many sound stimuli. For these analyses each unit responsive to 10 or more different sound stimuli was utilized. Acoustic parameters related to spectral and amplitude envelopes were estimated as mean, variance, skewness and kurtosis for each stimulus. Linear regression analysis was used to correlate between these acoustic parameters and mean firing rate during the significant interval of a stimulus presentation. These parameters of each sound included those estimated from spectral and amplitude envelopes, and harmonic-to-noise ratios (HNR). Mean firing rate of a unit responsive to a given sound was the dependent variable, and was referred to the average value within each significant 100 ms interval during a stimulus presentation. Regressions were conducted hierarchically for each responsive sound as spectral values, amplitude values, HNR and sound type entered to account for within-subject variability. Results did not indicate that any of these acoustic estimates were reliably correlated with spike activity of auditory responsive units. The author further inspected the data with scatter plots, and spectral mean of sounds was sometimes better for particular units when differentiating potential clusters of sound stimuli. Figure 39 illustrates observable clusters of sound stimuli from three units during the entire passive listening experiment. These three units respond to several stimuli of monkey vocalizations and pure tones, plus a small portion of stimuli from other types. Units D and E increased and decreased to potential clusters of sounds (in red and blue colors), even though spectral means of these two clusters were within a comparable range and thus spectral mean is not the only dimension of sound being coded for by these units. In the case of unit D, sound duration can be a potential factor for relating members within a cluster. Pure tones of passive listening 1 and all monkey vocalizations were generally shorter stimuli (average = 325 ms) compared to other types (average = 497 ms). It was found that unit D increased firing to the cluster of longer sounds and decreased firing to the cluster of shorter sounds, while both clusters mainly consisted of monkey vocalizations. For the other six units responsive to 10 sounds or above, potential clusters of sounds for a given unit seem not to be distinguishable by sound type, sound duration or other factors mentioned in the current experiment

When those units evoked by nine sound stimuli or less were examined, some of them could be linked to a particular sound type or a combination of sound types. There were three units and five units responsive exclusively to monkey vocalizations and pure tones respectively. Five units responded to a combination of monkey vocalizations and pure tones, but not other types. There were two units of each case where they responded to multiple sounds, in which monkey vocalizations or pure tones was the dominant sound type compared to others. Four units seemed to be evoked by sounds containing a wide frequency bandwidth and these sounds were a combination of white noises, music clips and nature sounds. Another two units responded to a monkey vocalization and a human vocalization exclusively.

.

Figure 39 Observable, separate clusters of sound stimuli of units D, E and F during passive listening experiments (# 0628101a, # 0426102aa & # 1026091a). Both units responded to several monkey vocalizations (M) and pure tone stimuli (P), plus a small portion of stimuli from other types. Units D and E increased and decreased firing to different monkey vocalizations, though both clusters (in red and blue colors) were mainly composed of sounds from the same types. Clusters in red, blue, green, brown or grey colors stand for responsive stimuli of each unit during passive listening. Grey boxes represent non-responsive sounds for each unit. Each letter represents the sound type of a responsive sound (A = animal; H = human; N = natural sound; S = synthesized clip; U = music; W = white noise).

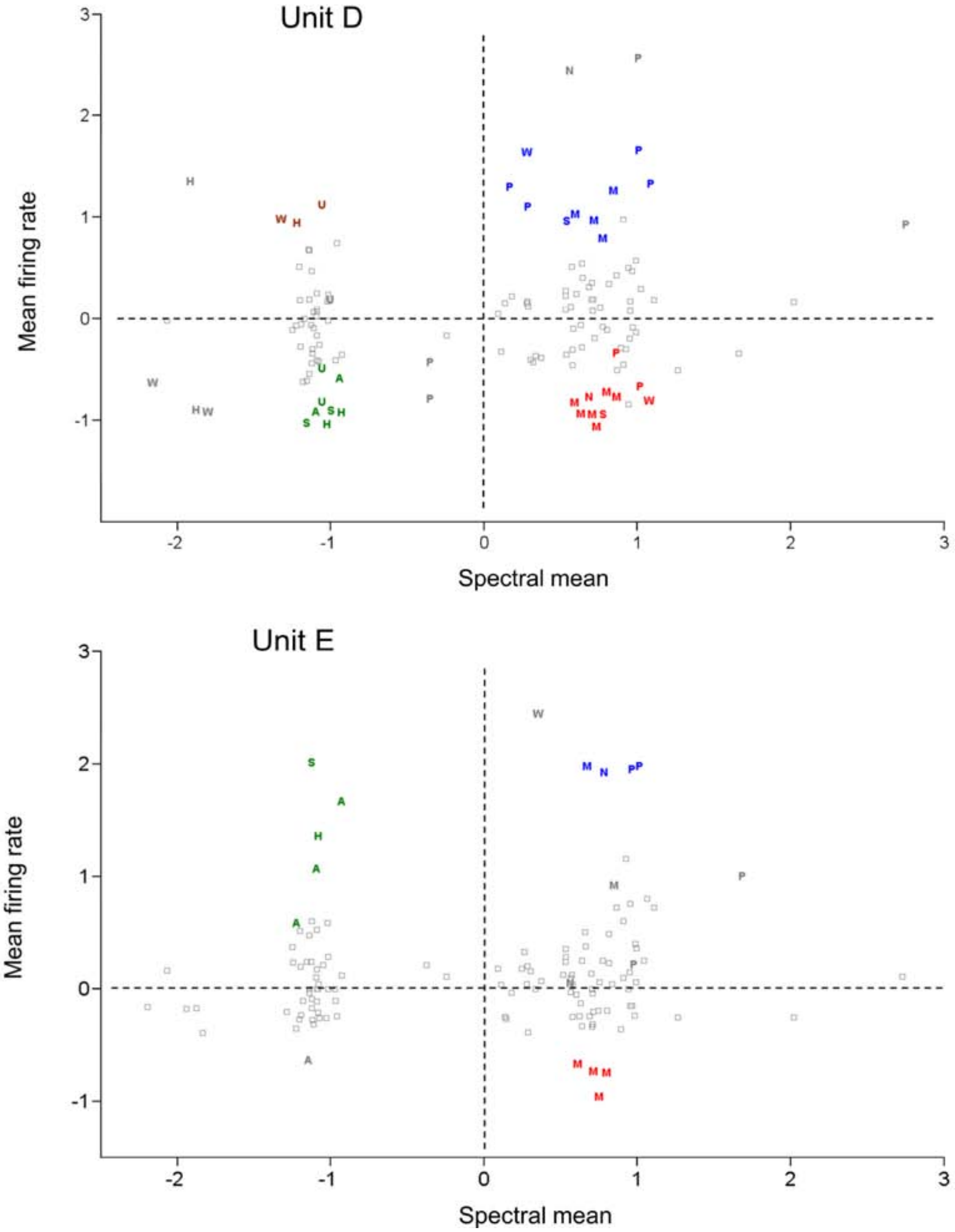
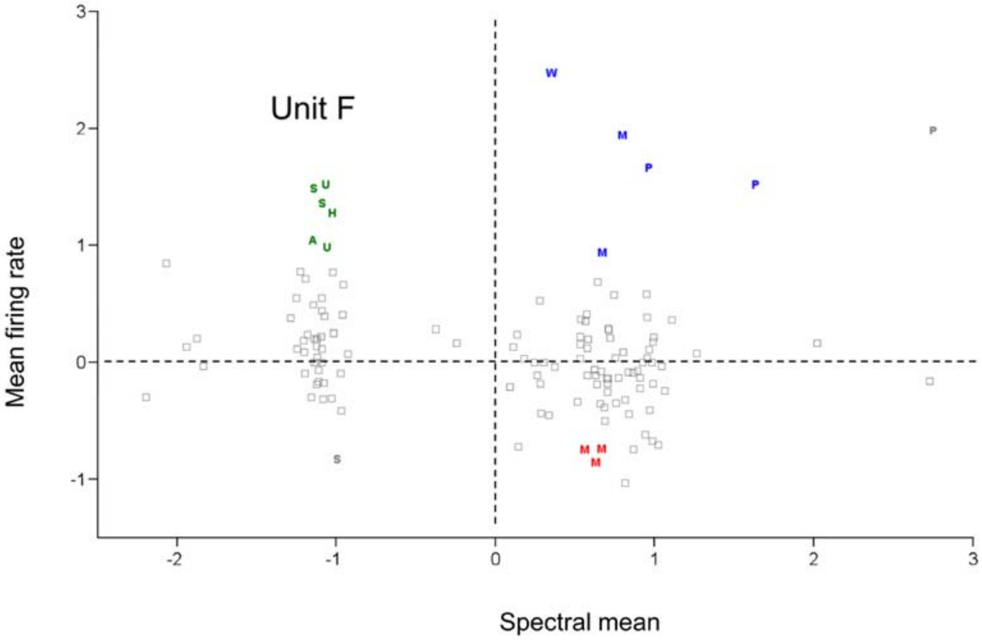


Figure 39 --- continued



Comparisons of Sound-Evoked Activity between Passive Listening and the Memory Task

The section examined how sound-evoked activity of dTP units was modulated across the three experimental contexts: passive listening 1, the memory task and passive listening 2. Among 176 units held during passive listening 1, 18% were auditory responsive during passive listening only. Initially non-responsive units during passive listening (22%) were later evoked by stimuli used in the memory task (Figure 40). A small number of units (7%) were consistently responsive to the same sound stimulus at both passive listening 1 and the memory task.

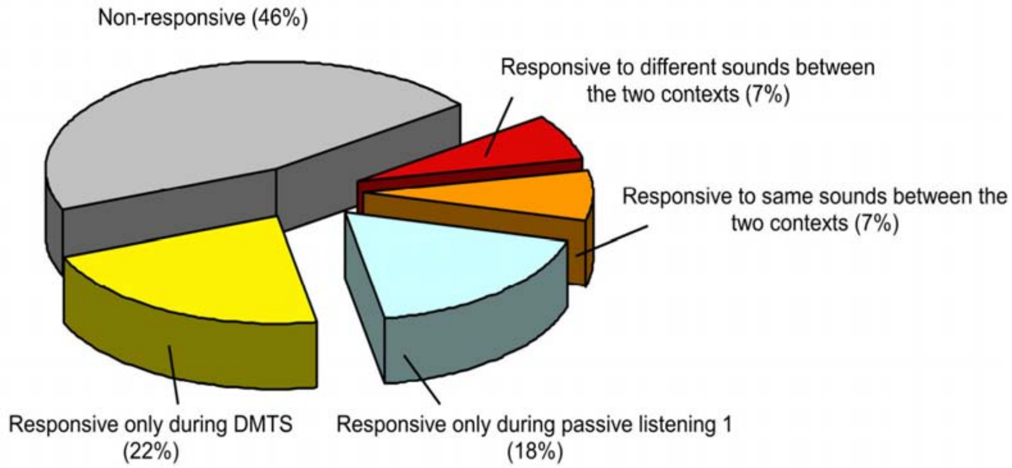


Figure 40. Distribution of dTP units responsive to sounds between passive listening 1 and DMTS task.

When the recording procedure switched from the memory task to passive listening 2, a few initially responsive units during the memory task (9%) ceased to respond to sounds. The majority of the non-responsive units held during the memory task (47%) became auditory responsive in passive listening 2 (Figure 41). Those responsive units between the two conditions primarily responded to different sounds (20%) instead of firing to the same sound stimulus (2%). These findings illustrated that units of dTP responded to different sound stimuli when switching between experimental contexts.

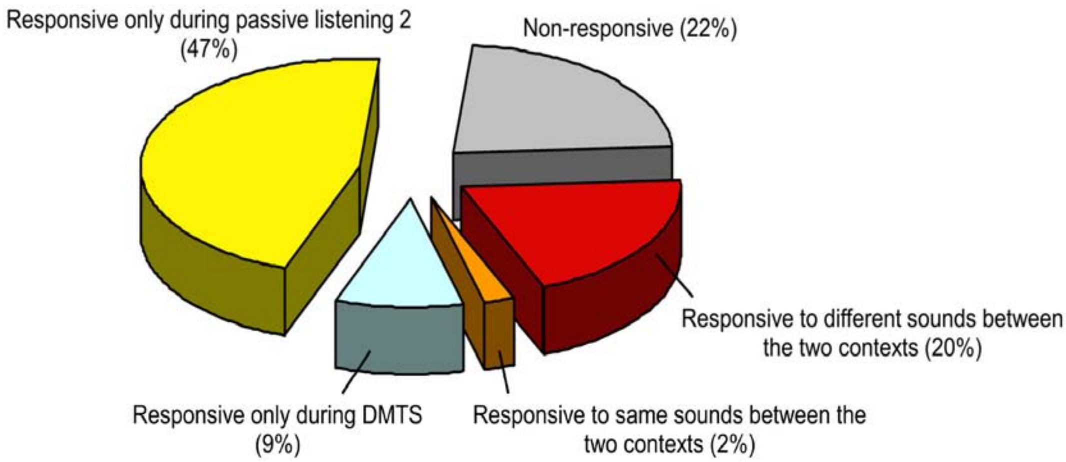


Figure 41. Distribution of dTP units responsive to sounds between the DMTS task and passive listening 2.

Recording Placement

Electrode placement inside the recording chamber is illustrated in Figure 42. Within the 225 units, there were more recorded units located on the lateral side of dTP (N = 120) than the medial side of dTP (N = 105). The depth of electrode penetration ranged between 33.502 – 42.395 mm, from the surface of the dura mater. Electrodes were sometimes lowered further down, beyond 45 mm from the surface of the dura mater, 13 recorded units (not included in the present study) that were more ventral than the intended dorsal temporal pole location and these units did not show any task- or stimulus-related activity at both the single-unit and population levels. There were 20% and 21% of total recorded units from the lateral and medial areas of dTP, respectively, that were not responsive to any task-relevant event during the DMTS task. Table 17 illustrates response profiles of dTP units across various task events, separated by anatomical locations. Within the lateral area of dTP, more units tended to be responsive during the delay and response periods 1 – 3 (34% or above). Within the medial area of dTP, more units had evoked activity during the response period 2 (40% versus 29% on average for other task-related events). Table 18 illustrates unit responsiveness to sounds during different experimental contexts, separated by anatomical locations. Patterns of auditory responsive units at each context were fairly similar between the lateral and medial areas of dTP.

Table 18. Percentage of units responsive to discrete task events during the DMTS task, separated by anatomical locations (Lateral: N = 120; Medial: N = 105).

	Cue 1	Cue 1 offset	Cue 2	Cue 2 offset	Delay	Wait	R1	R2	R3
Lateral	28	20	29	18	35	29	43	34	34
Medial	24	22	30	31	31	28	32	40	31

Table 19. Percentage of auditory responsive units at the three experimental contexts, separated by anatomical locations (Lateral: N = 120; Medial: N = 105).

	Passive Listening 1	DMTS	Passive Listening 2
Lateral	23	38	50
Medial	24	31	49

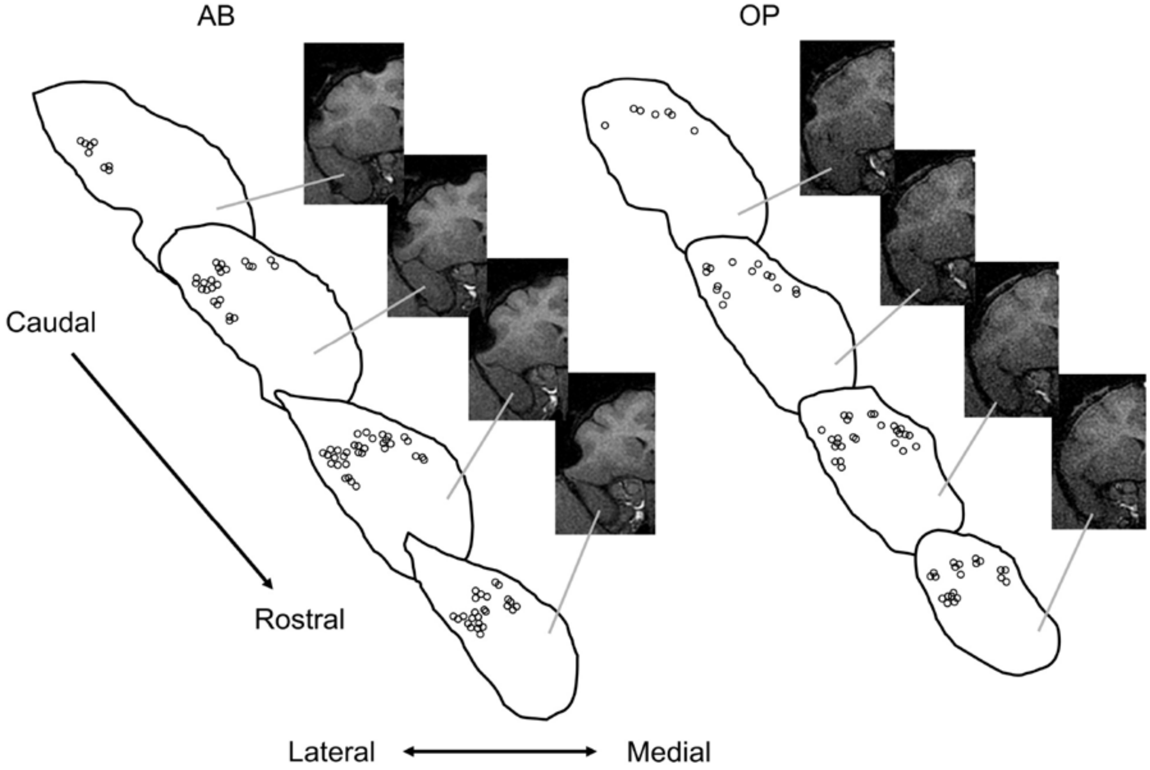


Figure 42. Electrode placements at dorsal temporal pole in the two monkeys AB and OP. Magnetic resonance images were obtained from the two monkeys at the left hemisphere of dorsal temporal pole (slice thickness is 1 mm). Schematic diagrams show recording sites at dTP, represented by circles.

Discussion

The present study investigates neuronal activity of left hemisphere dorsal temporal pole when two monkeys listened to a variety of sounds and performed an auditory delayed matching-to-sample task. In the DMTS task with a small sample set, pure-tone stimuli serve as a better cue for memory performance and fast response latency across 5-second memory delays, compared to other sound types. Population activity of dTP during cue events reveals important differences associated with accurate memory performance in the present task. Increased firing rate to cue 1 is related to memory accuracy, regardless of learning rules (matching or nonmatching sounds). This is in contrast to an absence of activity change during the same event in nonmatch incorrect trials. Neural encoding of cue 1 in dTP is therefore crucial for accurate auditory recognition of monkeys in the current experiment. Moreover, a suppression effect on population activity to identical stimuli exists during match correct trials, compared to nonmatch correct trials. The effect is quick and transient, lasting between 30 to 60 ms from the start of cue 2. Behavioral responses remarkably modulate spike activity of dTP and evoke the highest number of units with task-related activity during wait and response periods. Evoked activity during memory delays is rather limited, and sustained delay activity is rarely found in dTP. The findings may provide preliminary evidence as to why monkeys show poor auditory recognition memory and thresholds of forgetting, compared to those experiments testing visual memory performance of the same species (D’Amato and Colombo, 1985; Murray and Mishkin, 1998; Wright, 1998, 1999; Buffalo et al., 1999; Zola et al., 2000; Fritz et al., 2005). Passive listening experiments demonstrate that neurons of dTP seem to be highly sound-selective, and mostly respond to 1 to 4 sounds across all experimental contexts. Among those auditory responsive units, the study does not reveal substantial relationships between acoustic properties of sounds and spike activity of a given unit, at least based on those generally assessed by others and estimated in the current study (i.e., frequency range, spectral and amplitude envelopes, and harmonic-to-noise ratio). These results fit into the proposal that a hierarchical organization of information flow exists in a rostro-caudal axis along the superior temporal gyrus in the primate auditory system (Poremba and Mishkin, 2007; Rauschecker and Scott, 2009; Hackett, 2010).

Behavioral Performance on an Auditory Delayed Matching-To-Sample Task

The current findings reveal significant effects of sound type influencing auditory recognition memory performance of monkeys. Memory performance on nonmatch trials is improved when the test stimuli are pure tone stimuli and a decreased response latency to match trials is associated with the same sound type. Compatible results between memory accuracy and response latency suggest the monkeys perceptually and behaviorally distinguish pure tone stimuli more efficiently than other sound types.

Besides pure tone stimuli, natural sounds and white noises yield decreased response latencies relative to the other four sound types (i.e., animal vocalizations, human vocalizations, monkey vocalizations and music clips). Simple acoustic factors of these sound types may be efficiently utilized by monkeys in a limited sound context, compared to Experiment 1B of Chapter 2. Compared to the other seven sound types, most of the energy of a pure tone dominates at its characteristic frequency, and each stimulus contains a very positive, high value of harmonic-to-noise ratio (Figure 43). White noise bands contain a variety of frequency bandwidths and are perceptually monotonous compared to dynamic, rhythmic sounds (e.g., vocalizations and music clips). These stimuli also contain negative, low values of harmonic-to-noise ratio. The majority of natural sounds shares acoustic similarity with white noises, for example, burning fire, wind, thunder, and streaming water. Therefore, distinctiveness of the three sound types seems to serve as better cues for recognition in monkeys when the sounds are repeated multiple times across a single recording session. The pure tone stimuli, being the most acoustically simple type, robustly modulate memory performance at both match and nonmatch conditions.

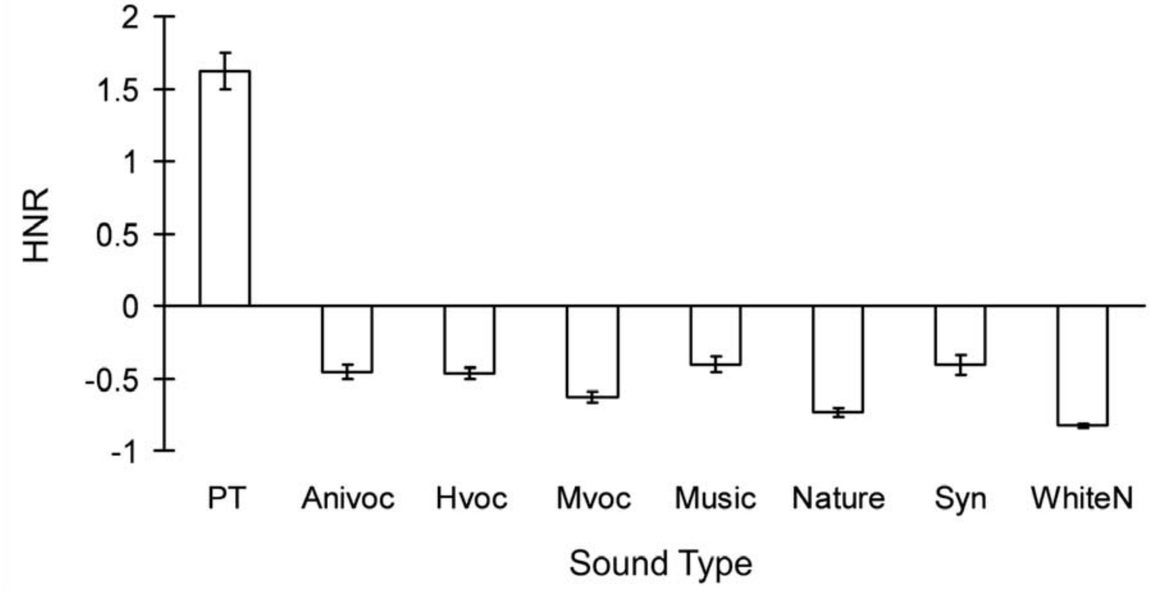


Figure 43. Harmonic-to-noise ratios of eight sound types used in the auditory DMTS task. Harmonic-to-noise ratio (HNR) indicates degree of acoustic periodicity for each sound type. Values of 96 sounds for HNR were transformed into standardized scores and pooled into the eight sound types.

This is in contrast to the findings of Chapter 2 as monkey vocalizations were shown to be better cues for recognition in monkeys during nonmatch memory performance. Prior behavioral and electrophysiological studies, as well as Chapter 2, commonly suggest that species-specific vocalizations should be better cues for mediating discrimination performance and evoking neuronal activity of auditory regions (e.g., the primary auditory core, rostral superior temporal gyrus and prefrontal cortex) in non-human primates (Zoloth et al., 1979; Petersen et al., 1984; 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). Additionally, important details (e.g., familiarity and biological significance, etc.) derived from conspecific cues are favorably attended by non-human primates (Hauser 1998; Fitch, 2000). These findings may not necessarily disagree with the present results. The behavioral context of sounds may be another important factor for influencing auditory recognition memory expression, and yet has not been studied extensively. In Chapter 2, Experiment 1B is a trial-unique DMTS task in which each sound stimulus is used once and never repeated within an experimental session. Within a recording session of Experiment 2A, eight different sounds, one from each type (a target set) are used for the auditory DMTS task. Twelve target sets are cycled across every recording session. Thus each stimulus of a sound type is novel between sessions, and is repeated within a given session. The behavioral context of sounds is not only constructed by interactions between each individual sound, but also repetitions of each sound occur at different trial conditions. The context may then affect monkeys' strategies on attending acoustic stimuli and modulate effects of sound type upon auditory memory performance in the present study. Future studies may focus on how different sound contexts modify expressions of auditory recognition memory. For instance, experimenters can manipulate degree of novelty or interference of a sound type relative to others during the memory task. Manipulations of different sound types may help elucidate whether effects of sound type on recognition memory are static or context-dependent.

Neuronal Activity of Dorsal Temporal Pole

The majority of the 225 dTP units (80%) exhibit significant activity changes to at least one event during the auditory DMTS task. When considering unit responsiveness within a trial type, task events during match correct trials generally evoked the highest number of units (56%) followed by the two nonmatch trial types (nonmatch correct: 34.7%, nonmatch incorrect: 36.4%). Those during match incorrect trials evoked the fewest number of units (20%).

Spike Activity Change during Cue Presentations

Cue presentation during the auditory DMTS task is shown to be an important event in neural encoding of dorsal temporal pole. More than one third of dTP units were responsive to at least one cue presentation (cue 1 or 2) of a trial type during the memory task. The population analysis was informative regarding differences between the three trial types during the recognition memory portion of the task, from cue 2 onset through the response periods. Population activity for dTP units remarkably increased firing rate within the first 100 ms interval of cue 1 for match and nonmatch correct trials, but not for nonmatch incorrect trials. This phenomenon suggests that correct memory performance is associated with attentive response of dTP on sample stimuli, regardless of learning rules.

Proper encoding of a sample stimulus may then mediate auditory recognition and facilitate memory performance of the two monkeys.

Another critical finding is that the population response to cue 2 was stronger to correct nonmatch compared to match trials. The current findings suggest that a suppression effect upon identical stimuli exists at the level of dTP during the auditory DMTS task. This is consistent with the similar processing termed match or repetition suppression observed in the visual object identification pathway located ventrally of dTP, inferior temporal cortex (ITC) and ventral temporal pole (vTP). A number of visual studies recorded single-unit activity of these visual cortical areas when monkeys were either performing a visual version of a delayed matching-to-sample task or a continuous recognition task (e.g., the *ABBA* paradigm). The studies show that individual neurons of ITC and vTP have a reduced firing rate for match stimuli, compared to nonmatch stimuli (vTP: Nakamura and Kubota, 1995, 1996; ITC: Baylis and Rolls, 1987; Miller et al., 1993; Miller and Desimone, 1994) and the phenomenon is also present in population responses of ITC neurons (Woloszyn and Sheinberg, 2009). Match suppression is often referred to in experimental conditions when behavioral choices were tied with explicit recognition of perceived events or objects (Miller et al., 1993; Persson et al., 2002; Woloszyn and Sheinberg, 2009). In addition, the suppression effect is also observed in situations when repeated stimuli were nonmatch stimuli or non-target distractors (Miller and Desimone, 1994). This phenomenon is suggested to be a repetition-induced suppression, because it was also shown in studies during which experimental subjects passively viewed stimuli (e.g., novel, familiar or morphed images) without overt responding tied with stimulus choice (Sawamura et al., 2006; McMahon and Olson, 2007; Liu et al., 2009) though some recent studies use the two terms interchangeably (Buchsbaum and D’Esposito, 2009; Buchsbaum et al., 2010). The author examined situations of stimulus repetition in which a sound stimulus of cue 2 from a previous trial was identical to that of cue 1 from the next trial. However, there were insufficient trials of data for conclusive analysis, as the experiment purposefully controlled and balanced sound pairings and their presentation orders during the memory task and passive listening blocks. Therefore, the experiment was designed to minimize effects of habituation or interference on spike activity due to stimulus repetition. Considering the setup of a delayed matching-to-sample task, match suppression seems more appropriate for describing the current finding. Future studies should verify this dampened effect on population response to familiar sounds. Other studies could systematically vary the number of sound repetitions, or numbers of intervening stimuli that separated the same sound can be manipulated during passive listening. During the memory task, certain sounds can be purposefully presented successively from trial to trial. Population patterns of spike activity during stimulus repetition could then be compared between passive listening and the memory task and help clarify the nature of reduced activity associated with identical sound presentations. Another alternative experiment could reverse the present task into a delayed nonmatching-to-sample task in which monkeys were required to produce responses for nonmatching sounds. This method would examine whether the suppression effect would occur on the match trials, and thus the phenomenon would be independent of a motor response.

In humans, reduced brain activations were also found in anterior superior temporal gyrus and sulcus when subjects passively listened to repeated stimuli or performed short-term memory tasks with verbal stimuli (Dehaene-Lambertz et al., 2006; Buchsbaum and D'Esposito, 2009; Buchsbaum et al., 2010). Suppression on matching sound stimuli revealed here may indicate that a familiar sound is more quickly and efficiently processed than a novel sound. Such a mechanism during recognition memory may then be common at higher-order brain areas of auditory and visual systems Priming might be related to the suppression effect with match stimuli in the current experiment. The hypothesis proposes that neuronal responses get sharpened, and therefore the encoding of match stimuli within a given trial can be supported by fewer, selective neurons of a given brain region (Desimone, 1996; Grill-Spector et al., 2006). In vision, match/repetition suppression is also revealed in the medial temporal memory system of humans and monkeys and is particularly robust in perirhinal and entorhinal cortices (Suzuki et al., 1997; Brown and Xiang, 1998; Ranganath and Rainer, 2003; Pedreira et al., 2010). Converging evidence between higher-order sensory association areas and medial temporal lobe may suggest a functional linkage between sensory perception and memory formation. A form of response fatigue or adaptation within dTP itself cannot explain the present finding. Population activity differences between the two correct trial types did not occur until at least 30-60 ms from the start of cue 2 presentations. Such trial differences at the population level were transient and ceased within the first 100 ms of cue 2 presentations consistent to those suppression effects shown in visual studies of non-human primates (Liu et al., 2009). The suppression shown in dTP is due to specific processing in a time-limited fashion, not general fatigue of the neuron.

Spike Activity Change during the Delay Period

Among the 225 recorded units of dTP, a low percentage of units (22%) show delay-related activity changes. There is a lack of consistent patterns of spike activity during the delay period correlated with behavioral performance. Population response during the three epochs (early, middle and late) of the memory delay is similar to the baseline activity level for each trial type. Sustained activity during delay is rare and seldom lasts longer than 1.5 s (i.e., one epoch of a memory delay). Unlike the current findings, neurons of ITC and vTP show sustained delay activity when monkeys performed a visual version of a DMTS task (Miller and Desimone, 1994; Nakamura and Kubota, 1995, 1996; Woloszyn and Sheinberg, 2009). One-third of recorded vTP units showed robust, steady firing activity up to delay conditions of 5 s. This delay-related activity is also correlated with response selectivity of visual stimuli and correct memory performance. Sustained delay activity of neurons is one of the important hallmarks for retention of visual information during working memory, and it is present in prefrontal and inferior temporal cortices (Fuster and Jervey 1982; Miyashita and Chang, 1988; Desimone, 1996; Miller et al., 1996). Non-human primates express poor memory performance and short forgetting thresholds for auditory stimuli compared to visual stimuli (D'Amato and Colombo, 1985; Wright, 1998, 1999; Fritz et al., 2005). Behavioral correlates of auditory recognition may then be related to transient, intermittent delay-related activity of dTP. More experiments are required to clarify the relationship between delay activity and auditory memory performance in non-human primates. Future studies may focus on those sound-selective neurons and use effective and non-effective sounds that a given unit responds to during the memory task. This recording method may maximize the likelihood of spike activity change related to stimulus encoding and retention during memory delays.

Spike Activity Change during the Wait and Response Periods

The subjects seem to anticipate or prepare for a possible button-press response during the 500 ms wait period on button-press trials as there were 28% of dTP units responsive to this event. Similar to cue events, evoked neurons expressed either an increase or decrease in firing rate. In fact, there were more units increasing than decreasing in firing rate during match correct trials, in contrast to those for the two nonmatch trial types. The population analysis of spike activity reveals something slightly different in that increased firing rates were present during the wait period for both match correct and nonmatch incorrect trials, and their activity was stronger than nonmatch correct trials. Trial differences between the two correct trial types occurred as early as the first 100 ms interval of the wait period and extended into the first response period (R1). These results are consistent with the detailed analysis of population responses to cue 2 offset. Trial differences between match and nonmatch correct trials occurred during the last 90 ms of cue 2 offset, and then extended into the wait period. A combination of decision-making processes and execution of a correct go-response seems to play a significant role in modulating neuronal activity of dTP.

Trials associated with behavioral responses distinguish patterns of spike activity during the first response period (R1) from those with an absence of responses (nonmatch correct trials). When spike activity was aligned to the instance when the first button-press response was produced, dTP population activity increased firing rate only during the post-response period 1 during match correct and nonmatch incorrect trials. The evoked population activity during button pressings may indicate feedback about the motor response rather than motor initiation. These findings suggest that feedback information about motor response may modulate neuronal activity of dTP. At the single-unit level there are also detectable differences between match correct and nonmatch incorrect trial types. The 1500 ms response period evoked the highest number of dTP units (56.9%) compared to other events during the memory task. Among these responsive units, match correct trials (25%) evoked more dTP units than nonmatch correct (9%) and nonmatch incorrect (10%) trial types. A button-press response (match correct and nonmatch incorrect trials) and food rewards (match correct trials only) could be present at R1. As illustrated by Figures 14, 19, and 24, spike activity of a dTP unit can be robustly strengthened or weakened by food rewards. Implications from the above analyses suggest that some dTP units not only encode behavioral response, but also food rewards. The current findings are parallel to the role of temporal pole for processing reward outcome and emotional state. Lesions to the temporal pole, amygdala or orbitofrontal cortex in monkeys induce Kluver-Bucy syndromes, for example, diminished fear, hypersexuality, social withdrawal, and impulsive behavioral actions. The similar social and emotional impairments are presented in human patients with bilateral temporal lobotomies (Olson et al., 2007; Blaizot et al., 2010). In human imaging studies, the temporal pole, as well as the amygdala and orbitofrontal cortex, respond to auditory, visual and olfactory stimuli associated with emotional contents (e.g., pleasant and unpleasant valence), compared to neutral stimuli (Royet al., 2000). Negative emotional states are often reported to be associated with the temporal pole in humans, for example, anger, anxiety, sadness and negative evaluation of gambling outcomes (Liu et al., 2007; Olson et al., 2007; Jimura et al., 2009; Blaizot et al., 2010). A neural network connecting the three areas may therefore mediate information processing of reward outcome and emotion (Markowitsch et al., 1985; Barbas et al., 1999; Kondo et al., 2003; Höistad and Barbas, 2008; Saleem et al., 2008; Reser et al., 2009). During auditory object processing, temporal pole may help in facilitating recognition and retrieval using emotional or episodic details to a familiar stimulus perceived before.

Sound Responsiveness during Passive Listening 1 and 2

Neurons of dTP generally respond to up to four different sound stimuli. Auditory responsive units often modified encoding across different stimuli for different experimental contexts instead of continuing to encode the same sound stimulus. Parallel to the current findings, numbers of responsive units along the superior temporal cortex decrease, as auditory units were located more rostrally in the direction toward temporal pole (Kikuchi et al., 2010). Units recorded from the rostral superior temporal plane (rSTP, 2 – 6 mm from the rostral tip of temporal pole) were more selective to sounds than those recorded from the primary auditory core (A1) and RT (i.e., the anterior portion of A1). Neurons of rSTP respond to up to three specific stimuli and often include monkey vocalizations. The reported sound selectivity of rSTP is comparable to dTP revealed in the present study. Onset response latency of rSTP units were on average 112 ± 9.4 ms, while dTP units showed on average 257 ± 6.7 ms. Increase in sound selectivity and onset response latency of dTP suggests that the area is situated at a higher-order region of auditory processing from the primary area A1 (Kaas and Hackett, 2000; Poremba and Mishkin, 2007; Kikuchi et al., 2010). The hierarchical organization of the auditory processing pathway of non-spatial attributes, from A1 rostrally extending to rSTG and dTP, thus bears a resemblance to that of ventral visual information pathway (i.e., V1 to ITC) in humans and monkeys (Rauschecker and Scott, 2009; Hackett, 2010).

Auditory responsive units of dTP seem to show fair responsiveness toward all eight sound types. A few units (N = 15, at least responsive to two sounds or more) tended to show sound responsiveness prone to certain sound types. At least one third of sounds each unit responded to can be associated with one of the eight sound types. The current results indicate a lack of simple relationships between effective sounds (which evoked neurons of dTP) and their associated acoustic properties mentioned above (i.e., harmonic-to-noise ratios, parameters of spectral and amplitude envelopes), with the more complex spectral mean underlying some portion of the separation whereas the rest of the parameters may vary by neurons. Higher-order auditory areas, like dTP, rSTG and the prefrontal cortex, may work as neuronal ensembles and flexibly encode certain complex features or significance embedded in a wide variety of sound stimuli (Romanski et al, 2005; Russ et al., 2008; Kikuchi et al., 2010). Future studies can initially assess stimulus-selectivity of dTP units with a collection of less complex sounds, for example, pure tones, tone bursts, ripple noises, white noises with different frequency bandwidths, and amplitude/frequency-modulated noises. After searching common sound-driven factors, prototype sounds composed of these factors may then be feasible for elucidating how higher-order auditory regions encode complex stimuli.

When switching between experimental contexts (see Figures 40 and 41), only a few units (7% or less) are responsive to the same sound stimulus throughout the recording experiments. About 20% of recorded units either became non-auditory responsive or switched firing preferences to different sound stimuli. The implication is unique in that dTP is flexible in selective encoding across different behavioral contexts. The notion of flexibility for stimulus selectivity has not been studied in the visual system, and is not inherent in the typical visual experimental designs. Prior studies often look for effective features that best evoke spike activity when sampling across a pool of standard images within a given context of passive viewing or visual fixation tasks. Hypotheses and models for visual object analysis are primarily based on stimulus selectivity shown in ITC, vTP, and other higher-order visual areas. The current recording procedures do not have an a priori criterion to study neurons of dTP with their preferred sounds or sound features, and they assess sound responsiveness across experimental contexts. Future studies could investigate sound encoding of dTP and other auditory regions during different experimental contexts with the same pool of stimuli and evaluate degrees of encoding flexibility that may vary with attention, behavioral requirement and/or task complexity.

Response profiles of auditory responsive units revealed in dTP is reminiscent of the model of sparse coding for visual stimulus processing shown in medial temporal lobe (Olshausen and Field, 2004; Quiroga et al., 2005, 2007). The model suggests that a small, selective group of neurons are able to explicitly represent specific features of a given stimulus, therefore leaving the remaining neurons in a brain region to be quiet and non-responsive. This is in contrast to the model of distributed population coding in which a large group of neurons are broadly tuned to discrete features of stimuli and thus a given stimulus can be represented by large, collective neuronal ensembles. In this model, a large portion of neurons are responsive to a wide variety of sensory stimulation. In the present study, neurons in dTP often show low firing activity, around 6 – 10 spikes per second. The majority of these auditory responsive units are very sound-selective, but their firing preferences do not respond to sounds based on purely acoustic properties. These findings suggest that dTP neurons do not faithfully encode acoustic features of sounds compared to auditory regions at earlier stages of auditory processing (Wang, 2007). Dorsal temporal pole neurons may instead demonstrate a possibly abstract, sparse coding of sounds at a later stage of auditory encoding in non-human primates.

Neural Correlates of Auditory Recognition Memory:
Prefrontal Cortex versus Dorsal Temporal Pole

The only other brain region where single-unit activity has been recorded from during the auditory DMTS task is the prefrontal cortex. The prefrontal cortex (PFC) has been suggested to be involved in working and short-term memory functions in visual studies of humans and non-human primates (Miller et al., 1996; Owen et al., 1998; Levy and Goldman-Rakic, 2000; Curtis and D'Esposito 2004). Neurons of PFC encode visual stimuli and show sustained delay activity in various working memory tasks. The frontal region receives neural inputs from various brain areas of the auditory nervous system including belt, lateral belt and parabelt regions of the auditory association cortex (Romanski et al., 1999a, b). In humans and monkeys, PFC responds to auditory stimuli, including species-specific vocalizations (Alain et al., 2001; Romanski and Goldman-

Rakic, 2002; Poremba et al., 2003; Romanski et al., 2005; Cohen et al., 2007), and the area is also activated when human subjects performed auditory recognition memory tasks (Rämä et al., 2004; Rämä and Courtney, 2005). Since dorsal temporal pole has reciprocal connections to the prefrontal cortex (Carmichael and Price, 1995; Barbas et al., 1999; Romanski et al., 1999a; Kondo et al., 2003, 2005), interactions between the two areas may mediate auditory recognition memory performance. Here the author compares the electrophysiological results of dTP with those obtained from PFC conducted by another experimenter in the same laboratory (Plakke et al., under peer review). The study of Plakke et al. utilized the same two monkeys for single-unit recordings during the auditory DMTS task and passive listening blocks. Isolated neuronal activity (N = 215) was collected from the lateral prefronal cortex (lPFC). Experimental methods for analyzing spike activity of lPFC across discrete task events and sound stimuli were similar to the current study.

Similarities in Wait, Response, Delay Periods and Multi-Encoding of Events

During the auditory DMTS task, task events during match correct trials often evoked more units from both dTP and lPFC than those during other trial types, particularly at wait and response periods (R1 – 3). Parallel to the current findings, behavioral responses exert a dominant influence on neuronal activity of lPFC during match correct and nonmatch incorrect trials. Although more units of lPFC (46%) showed activity changes during the three epochs of delay period than dTP (21%), sustained delay activity shown by lPFC is as limited as those by dTP. Low levels of delay activity from these two areas during the auditory DMTS task may provide joint evidence to explain why rhesus monkeys express poor auditory memory performance and short forgetting thresholds for sounds, compared to the visual counterparts. Across the nine discrete task events, the majority of units in lPFC (over 50%) responded to more than one event. Both dTP and lPFC encode multiple events occurring during the auditory memory task, and this effect may contribute to mediate performance accuracy during the behavioral task.

Differences during Cue Presentation and Button-Press Response

The critical difference between the two areas lies in population response to the presentations of cue 1 and 2. Compared to dTP, the population response of lPFC to the difference between the cue 1 to cue 2 presentation was stronger at match correct trials than nonmatch correct trials. While population activity of dTP shows match suppression to the identical sound (i.e., match correct trials), that of lPFC shows match enhancement instead. Reduced population activity of dTP occurred transiently, around 30 to 60 ms from the start of cue 2, while enhanced population activity of lPFC started around 200 ms of cue 2 and lasted for another 200 ms interval before cue 2 offset. In short, dTP and lPFC illustrate effects of match suppression and enhancement at population level, respectively, when distinguishing correct memory performance between match and nonmatch rules. During presentations of cue 2, information about identical sounds seems to first be processed by dTP, followed by lPFC in the current DMTS task. Another critical difference between the two areas is present during cue 2 offset. For lPFC, population activity of cue 2 offset during match correct trials is higher than those for nonmatch correct trials at the entire 500 ms interval. For dTP, the respective trial difference appears at the last 200 ms of cue 2 offset. The match enhancement effect during cue 2 presentation may be carried over to the cue offset period, and then contribute to latency differences on the two correct trial types shown in lPFC and dTP. Compared to lPFC, the late onset latency of dTP activity may be closely related to the anticipation of an eventual response. When focusing on the pre-response period, population activity of lPFC during match correct trials is stronger than those during nonmatch incorrect trials. This is in contrast to an absence of trial differences revealed in dTP. Expectations of behavioral outcome (correct/error, reward or both) may further delineate population encoding of task events between the two areas.

Comparison between dTP and lPFC for Sound Evoked Activity during Passive Listening and the Memory Task

Generally, there is a consistent trend showing that more units of lPFC were auditory responsive to sounds in general, relative to dTP (Table 19). A higher percentage of units with evoked activity in lPFC and dTP were identified when more sounds were presented during passive listening 2. There is a lack of firing preference to sound types in dTP and 1PFC when unit responsiveness is classified into eight sound types. This phenomenon suggests that these two higher-order auditory regions respond similarly to various sound types. Yet, dTP is more sound-selective than lPFC with the neurons there responding to fewer sounds across different experimental contexts.

Table 20. Percentage of units in dTP and lPFC responsive to at least one sound stimulus throughout passive listening blocks and the auditory DMTS task.

	dTP	lPFC
Passive Listening 1	30 (N=176)	62 (N=70)
DMTS	34 (N=225)	57 (N=215)
Passive Listening 2	69 (N=161)	98 (N=157)

In the visual nervous system, a division of labor exists between ITC and PFC for visual categorization and memory functions. Inferior temporal cortex is involved in analyzing stimulus attributes of visual images, and PFC is involved in encoding behavioral-relevant details, such as stimulus categories, response selection and memory delay (Freedman et al., 2003; Muhammad et al., 2006; Meyers et al., 2008). The current results also imply differential roles of auditory processing in recognition and memory functions between dTP and lPFC. Dorsal temporal pole is more prone to encoding specific features of sounds, and lPFC is more liberal and encodes sounds as events pertinent to a behavioral situation. The division of labor in object analysis and identification may then be domain specific between sensory association areas and prefrontal cortex, and conserved between the two sensory systems.

In short, the present study demonstrates evidence that dorsal temporal pole is a higher-order auditory region for auditory information processing, and neurons of this area tend to be very sound-selective to specific stimuli as well as encode discrete task-relevant events when monkeys are performing an auditory version of a delayed matching-to-sample task. Its proper encoding of sample stimuli during the current task is crucial to memory performance accuracy. Match suppression on neuronal activity in dorsal temporal pole suggests familiar stimuli are quickly and efficiently processed, and in turn facilitates auditory recognition memory in non-human primates.

CHAPTER 4. GENERAL DISCUSSION

Auditory memory is important for survival of animals in daily life. Discriminating and recognizing auditory information regarding identity, sex, kinship, emotional states and communicative significances (e.g., bird songs, monkey calls and human speech) involves basic to complex processing in the auditory nervous system of birds, non-human primates and humans and requires memory as well. However, studies about auditory recognition memory are very limited in non-human primates and other animals or birds. In fact, monkeys have difficulty acquiring complex behavioral learning and recognition with auditory stimuli, and maintaining auditory information over memory delays (Wright et al., 1990, 2000; Wright, 1998, 1999; Fritz et al., 2005) in contrast to their rapid learning and excellent retention of visual information (Murray and Mishkin 1998; Buffalo et al. 1999; Zola et al. 2000). Short-term auditory recognition memory is dependent on rostral superior temporal gyrus, but not the rhinal cortex (i.e., entorhinal and perirhinal cortices), which is necessary for visual memory performance in non-human primates (Fritz et al., 2005). Yet, there are similarities on neural mechanisms of information processing for both sensory modalities, for example, topographical representation of stimulus attributes, hierarchical organization of information processing along primary, secondary and sensory association areas, and separate streams of information flow regarding spatial and non-spatial attributes (Poremba and Mishkin, 2007; Rauschecker and Scott, 2009; Hackett, 2010). The author is interested in investigating behavioral and neural correlates of auditory recognition memory in non-human primates. The multi-disciplinary approach provides implications for how monkeys process sounds and perform memory functions. It is particularly important to seek evidence at the neuronal level of the primate auditory system in order to account for why auditory memory of monkeys is not as robust as their visual counterpart. This information may tell us about mechanisms of sensory memory and the increased communication that humans show that relies heavily on the auditory system. Dorsal temporal pole (dTP), a cortical region at the rostral portion of the superior temporal gyrus, has been associated with auditory perception with complex sound stimuli (Poremba and Mishkin, 2007; Olson et al., 2007) that is particularly sensitive to species-specific vocalizations in humans and monkeys (Belin et al., 2002; Poremba et al., 2004; Petkov et al., 2008). It is also implicated by lesion and imaging studies in discrimination of, and recognition of, complex sounds (Heffner and Heffner 1984, 1986; Nakamura et al., 2001; Jimura et al., 2009). This thesis thus examined auditory recognition performance of monkeys across sound types and neural correlates of auditory processing in dTP.

Experiments in Chapters 2 and 3 both demonstrate auditory recognition memory of monkeys varied by sound types. In a trial-unique design when all sound stimuli were always novel within a session, monkeys yielded better memory performance when distinguishing a sound from a monkey vocalization. When only limited to one sound from each sound type utilizing a total sample set of only eight sounds/session and more trials, pure tones became better cues for memory performance. Results from the two experiments do not necessarily contradict each other, but instead provide critical implications for expressions of auditory recognition memory. Compared to Experiment 1B, the behavioral context of sounds is dramatically different in the recording experiment. The same eight sounds were used and repeated throughout each session as necessitated by the recording experiment, unlike the situation presented in Experiment 1B in which each sound was presented once. Interactions between different sounds and repetition-induced effects from each sound establish a more complex context for auditory memory processing. When each stimulus is always novel (Experiment 1B), attending to common features shared within a sound type may be efficient for recognition and discrimination during a context of varying sound identity. Chapter 2 thus demonstrates that familiarity, biological and ethological significances of monkey vocalizations may contribute to behavioral performance, as these conspecific sounds are more readily processed by subjects (Hauser, 1998; Fitch, 2000; Gifford et al., 2003). This is analogous to the benefits of using face stimuli on visual perception and memory performance in humans and monkeys (Diamond and Carey, 1986, Parr and Heintz, 2008). During the recording experiment, attending particular features of a given sound may be sufficient during nonmatch memory performance in a relatively limited sound context. Pure tones may be better cues for evoking activity and discriminating sounds due to their distinctive harmonic structure and characteristic frequencies from other types. Therefore, the behavioral context of sounds can affect monkeys' strategies in auditory processing, as well as play a pivotal role in modulating expressions of auditory recognition memory. Future studies should examine dynamics of sound context and its impact on auditory memory functions of monkeys at various memory delay durations. Though effects of sound type on auditory memory expression depend on the context of sounds, the resultant manifestations on behaviors seem to produce the same improvements on nonmatch memory performance as test stimuli and short response latencies for match memory performance. Subjects in the current study perceptually and behaviorally distinguished the "effective" sound type better than others. These parameters can serve as one of the underpinning measures to study auditory recognition memory of monkeys in the go/no-go version of a DMTS task. Tracking response latencies associated with manipulations of sound features and context may then be another important factor to investigate short- and long-term auditory memory in non-human primates.

Memory improvement due to the "effective" sound type is only manifested during nonmatch trials of two different sounds. These acoustic advantages may be prone to memory judgment upon distinct auditory cues perceived in a closely timed manner. Monkeys in both Experiment 1 and 2 perform uniformly well across all sound types, reflecting a possible asymptotic effect at fixed 5-second delays. Whether similar acoustic advantages would be present during match memory performance may require future studies using long memory delays (e.g., 10 s, 15 s or longer) to increase task difficulty. Comparison between auditory and visual memory performance using humans and monkeys proposes that auditory memory is not as robust as the visual counterpart (Wright, 1998, 1999; Fritz et al., 2005; Cohen et al., 2009). Performance discrepancies may be due to differences in nature of their sensory qualities. When visual images of words, faces and pictures are presented, visual details are all available to subjects simultaneously. In contrast, auditory presentations of the same information are temporally organized. Acoustic details fade in and out continuously and vary rapidly from time to time. This may lead to a difference in information processing, and thus precipitate as a divergence for memory capability. Future studies are required to elucidate origins of memory differences between auditory and visual information present in humans and monkeys.

Chapter 2 also demonstrates a sexual difference on auditory memory performance of monkeys, i.e., males generally express better memory accuracy than females. Prior studies suggest a reversal of this sexual difference on performances in non-spatial tasks. However, one male and one female were tested in the DMTS task of Experiment 2A, and memory performance between match and nonmatch trials was similar to each other. This discrepancy may be due to both the limited sample size of subjects in the current experiment, and experimental designs that do not purposefully address the sexual issue. Future studies could test sexual differences on auditory memory with more animal subjects and employ a series of auditory discrimination and memory tasks to clarify the nature of sexual differences.

Chapter 3 examined neural correlates of auditory processing and recognition memory in the left-hemisphere dorsal temporal pole. The study confirms that dorsal temporal pole is a higher-order area for auditory processing. Neurons of dTP show higher sound specificity and longer onset response latencies than those revealed in the primary and secondary auditory regions (Recanzone et al., 2000; Tian et al., 2001; Bendor and Wang, 2008; Kusmierek and Rauschecker, 2009; Kikuchi et al., 2010). Though the available parameters that estimated acoustic properties of sounds do not seem linked to spike activity in dTP, a few neurons tend to show firing preferences prone to certain sound types. The phenomenon may be related to the notion that higher-order brain regions receive large amounts of information from the primary and secondary auditory areas, and thus broadly tuned neurons in dTP respond better to complex stimulus features of sounds (Hackett et al., 2010). In order to elucidate the ‘token’ attributes dTP encoded from complex sounds, future studies may initially sample spike activity from a wide variety of sounds to search for the effective ones and then decompose them into common features necessary to evoke neuronal activity of dTP. This is analogous to the stimulus simplification process used in visual electrophysiological studies of monkeys in which neurons of ITC were shown to selectively encode shapes of visual images, e.g., a face or a hand (Desimone et al., 1984; Tanaka et al., 1991; Kobatake and Tanaka, 1994).

During the memory task, units of dTP encode various task events including: epochs of cue presentation, memory delay, wait time and response periods. Encoding of sample stimuli (cue 1) effectively differentiates correct memory performance (match and nonmatch correct trials) from the incorrect one (nonmatch incorrect trials), regardless of match and nonmatching rules. Attending to sample stimuli may then be associated with improved accuracy for auditory recognition in a delayed matching-to-sample task. Encoding of test stimuli (cue 2) furthermore distinguishes the two types of correct memory performance; reduced population response to identical sounds is in contrast to increased population response to different sounds. This match suppression effect occurs rapidly and transiently. These findings are parallel to neural correlates of visual short-term memory shown in inferior temporal cortex and ventral temporal pole (Baylis and Rolls, 1987; Miller et al., 1993; Miller and Desimone, 1994; Nakamura and Kubota, 1995, 1996; Freedman et al., 2003; Muhammad et al., 2006; Meyers et al., 2008; Woloszyn and Sheinberg, 2009). In general, these higher-order sensory association areas are able to analyze attributes of currently perceived stimuli, process familiar information efficiently, and facilitate recognition and memory functions. It is still unknown whether or not match suppression would be originated within dTP or contributed by other brain areas at an earlier stage of auditory processing, as dTP and lPFC are the only regions recorded from during this auditory memory task. Future studies may record neurons from the primary and secondary auditory areas (e.g., A1 and lateral belt areas) to assess neural correlates of match suppression in a DMTS task.

At the single-unit level, spike activity of dTP during the memory delay is rather limited and intermittent. Sustained delay activity, one of the hallmarks for retention of information in visual working memory models (Desimone, 1996; Miller et al., 1996; Ungerleider et al., 1998) is rather scarce in dTP. Congruent evidence from dTP and lPFC may jointly explain why monkeys have poor auditory recognition and weak thresholds for forgetting, relative to those examined in vision. Compared to visual studies, the recording protocol of the current project does not have an a priori selection to study neurons with their preferred sound stimuli. Future studies may initially assess the effective and non-effective sound stimuli to evoke an isolated unit before the monkey subject performs a memory task. This a priori selection in auditory electrophysiology may maximize the likelihood to obtain better delay-related activity in dTP.

Lastly, implications of the current electrophysiological project suggest that information processing at higher-order sensory areas is fairly similar between the auditory and visual nervous systems, which are a hierarchical organization of cortical representation for stimulus attributes. It is also coherent with the proposed model of the ‘what’ information processing stream for object identification and recognition (Mishkin et al., 1983; Rauschecker, 1998; Romanski et al., 1999b; Kaas and Hackett, 2000; Poremba and Mishkin, 2007; Recanzone and Cohen, 2010). The present study reveals neural correlates of auditory encoding and memory functions in dorsal temporal pole focused on data from the left hemisphere of two monkeys. Hemispheric specialization of auditory processing in humans and monkeys demonstrates that the right hemisphere, e.g., posterior portions of STG, specializes in processing a wide variety of sounds, and the left hemisphere specializes in processing species-specific vocalizations, for example, temporal pole and the rostral portion of STG (Poremba et al., 2004; Poremba and Mishkin, 2007; Rauschecker and Scott, 2009). Interhemispheric interactions between both sides of auditory cortical regions, as well as temporal pole, may then collaborate for neural representations of acoustic events and further mediate learning and memory functions.

One crucial difference between the sensory systems lies on expression of short-term recognition memory, and the concern over whether auditory memory is organized differently from the visual system is still unclear. A recent review employed the anatomical perspective to compare and contrast the information flow from sensory cortical areas to the medial temporal memory system (i.e., hippocampal formation and parahippocampal cortices) between audition and vision (Munoz-Lopez et al., 2010). Compared to the visual nervous system, auditory information from superior temporal gyrus reaches the medial temporal areas through indirect connections with polysensory areas including temporal pole, superior temporal sulcus, insula and medial frontal cortex. The hypothesis suggests that the direct and indirect flows of sensory information to the medial temporal areas may then be associated with different formation and expression of recognition memory present in the two systems. More studies like the current project are required to examine the neural correlates of auditory memory functions along the primary, secondary and association cortical areas, in order to elucidate the organization of auditory memory in non-human primates.

REFERENCES

Alain, C., Arnott, S.R., Hevenor, S., Graham, S., Grady, C.L., 2001. "What" and "where" in the human auditory system. Proc Natl Acad Sci U S A. 98, 12301-6.

Amrhein, P.C., McDaniel, M.A., Waddill, P., 2002. Revisiting the picture-superiority effect in symbolic comparisons: do pictures provide privileged access? J Exp Psychol Learn Mem Cogn. 28, 843-57.

Andics, A., McQueen, J.M., Petersson, K.M., Gál, V., Rudas, G., Vidnyánszky, Z., 2010. Neural mechanisms for voice recognition. Neuroimage. 52, 1528-40.

Bachevalier, J., Mishkin, M., 1994. Effects of selective neonatal temporal lobe lesions on visual recognition memory in rhesus monkeys. J Neurosci. 14, 2128-39.

Barbas, H., Ghashghaei, H., Dombrowski, S.M., Rempel-Clower, N.L., 1999. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. J Comp Neurol. 410, 343-67.

Barrett, D.J., Hall, D.A., 2006. Response preferences for "what" and "where" in human non-primary auditory cortex. Neuroimage. 32, 968-77.

Baylis, G.C., Rolls, E.T., 1987. Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. Exp Brain Res. 65, 614-22.

Baylis, G.C., Rolls, E.T., Leonard, C.M., 1987. Functional subdivisions of the temporal lobe neocortex. J Neurosci. 7, 330-42.

Belin, P., 2006. Voice processing in human and non-human primates. Philos Trans R Soc Lond B Biol Sci. 361, 2091-107.

Belin, P., Zatorre, R.J., 2000. 'What', 'where' and 'how' in auditory cortex. Nat Neurosci. 3, 965-6.

Belin, P., Zatorre, R.J., Ahad, P., 2002. Human temporal-lobe response to vocal sounds. Cogn Brain Res. 13, 17-26.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature. 403, 309-12.

Bélizaire, G., Fillion-Bilodeau, S., Chartrand, J.P., Bertrand-Gauvin, C., Belin, P., 2007. Cerebral response to 'voiceness': a functional magnetic resonance imaging study. Neuroreport. 18, 29-33.

Bendor, D., Wang, X., 2008. Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. J. Neurophysiol. 100, 888–906.

Betke, K., 1991. New hearing threshold measurements for pure tones under free-field listening conditions. J. Acoust. Soc. Am. 89, 2400-2403.

Blair, R.J., Morris, J.S., Frith, C.D., Perrett, D.I., Dolan, R.J., 1999. Dissociable neural responses to facial expressions of sadness and anger. Brain. 122, 883-93.

Blaizot, X., Mansilla, F., Insausti, A.M., Constans, J.M., Salinas-Alamán, A., Pró-Sistiaga, P., Mohedano-Moriano, A., Insausti, R., 2010. The human parahippocampal region: I. temporal pole cytoarchitectonic and MRI correlation. Cereb Cortex. 20, 2198-212.

Blake, D.T., Heiser, M.A., Caywood, M., Merzenich, M.M., 2006. Experience-dependent adult cortical plasticity requires cognitive association between sensation and reward. Neuron. 52, 371-81.

Boersma, P., Weenink, D., 2007. Praat: doing phonetics by computer (Version 4.6.38) [Computer program]. Retrieved November 19, 2007, from http://www.praat.org/.

Brown, M.W., Xiang, J.Z., 1998. Recognition memory: neuronal substrates of the judgement of prior occurrence. Prog Neurobiol. 55, 149-89.

Buchsbaum, B.R., D'Esposito, M., 2009. Repetition suppression and reactivation in auditory-verbal short-term recognition memory. Cereb Cortex. 19, 1474-85.

Buchsbaum, B.R., Padmanabhan, A., Berman, K.F., 2010. The neural substrates of recognition memory for verbal information: spanning the divide between short- and long-term memory. J Cogn Neurosci. [Epub ahead of print]

Buffalo, E.A., Ramus, S.J., Clark, R.E., Teng, E., Squire, L.R., Zola, S.M., 1999. Dissociation between the effects of damage to perirhinal cortex and area TE. Learn Mem. 6, 572-99.

Bulthoff, I., Newell, F.N., 2006. The role of familiarity in the recognition of static and dynamic objects. Prog Brain Res. 154, 315-25.

Carmichael, S.T., Price, J.L., 1995. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. J Comp Neurol. 363, 615-641.

Cheung, S.W., Nagarajan, S.S., Bedenbaugh, P.H., Schreiner, C.E., Wang, X., Wong, A., 2001. Auditory cortical neuron response differences under isoflurane versus pentobarbital anesthesia. Hear Res. 156, 115-27.

Chung, D.Y., Mason, K., Gannon, R.P., Willson, G.N., 1983. The ear effect as a function of age and hearing loss. J. Acoust. Soc. Am. 73, 1277-1282.

Cohen, M.A., Horowitz, T.S., Wolfe, J.M., 2009. Auditory recognition memory is inferior to visual recognition memory. Proc. Natl. Acad. Sci. U S A. 106, 6008-10.

Cohen, Y.E., Hauser, M.D., Russ, B.E., 2006. Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. Biol Lett. 2, 261-5.

Cohen, Y.E., Russ, B.E., Gifford, G.W. 3rd., Kiringoda, R., MacLean, K.A., 2004. Selectivity for the spatial and nonspatial attributes of auditory stimuli in the ventrolateral prefrontal cortex. J Neurosci. 24, 11307-16.

Cohen, Y.E., Theunissen, F., Russ, B.E., Gill, P., 2007. Acoustic features of rhesus vocalizations and their representation in the ventrolateral prefrontal cortex. J Neurophysiol. 97, 1470-84.

Colombo, M., D'Amato, M.R., Rodman, H.R., Gross, C.G., 1990. Auditory association cortex lesions impair auditory short-term memory in monkeys. Science. 247, 336-8.

Colombo, M., Gross, C.G., 1994. Responses of inferior temporal cortex and hippocampal neurons during delayed matching to sample in monkeys (Macaca fascicularis). Behav Neurosci. 108, 443-55.

Colombo, M., Rodman, H.R., Gross, C.G., 1996. The effects of superior temporal cortex lesions on the processing and retention of auditory information in monkeys (Cebus apella). J Neurosci. 16, 4501-17.

Curtis, C.E., D'Esposito, M., 2004. The effects of prefrontal lesions on working memory performance and theory. Cogn Affect Behav Neurosci. 4, 528-539.

Cynx, J., 1993. Conspecific song perception in zebra finches (Taeniopygia guttata). J Comp Psychol. 107, 395-402.

Cynx, J., Hulse, S.H., Polyzois, S., 1986. A psychophysical measure of pitch discrimination loss resulting from a frequency range constraint in European starlings (Sturnus vulgaris). J Exp Psychol Anim Behav Process. 12, 394-402.

Cynx, J., Nottebohm, F., 1992. Role of gender, season, and familiarity in discrimination of conspecific song by zebra finches (Taeniopygia guttata). Proc Natl Acad Sci U S A. 89, 1368-71.

D'Amato, M.R., Colombo, M., 1988. On tonal pattern perception in monkeys (Cebus apella). Anim Learn Mem. 16, 417-424.

D'Amato, M.R., Salmon, D.P., 1982. Tune discrimination in monkeys (Cebus apella) and in rats. Anim Learn Behav. 10, 126-134.

D'Amato, M.R., Colombo, M., 1985. Auditory matching-to-sample in monkeys (Cebus apella). Anim Learn Behav. 13, 375-382.

Davis, M., 1989. Neural systems involved in fear-potentiated startle. Ann N Y Acad Sci. 563, 165-83.

deCharms, R.C., Merzenich, M.M., 1996. Primary cortical representation of sounds by the coordination of action-potential timing. Nature. 381, 610-3.

Dehaene-Lambertz, G., Dehaene, S., Anton, J.L., Campagne, A., Ciuciu, P., Dehaene, G.P., Denghien, I., Jobert, A., Lebihan, D., Sigman, M., Pallier, C., Poline, J.B., 2006. Functional segregation of cortical language areas by sentence repetition. Hum Brain Mapp. 27, 360-71.

Desimone, R., 1996. Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci U S A. 93, 13494-9.

Desimone, R., Albright, T.D., Gross, C.G., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. J Neurosci. 4, 2051-62.

Desimone, R., Gross, C.G., 1979. Visual areas in the temporal cortex of the macaque. Brain Res. 178, 363-80.

Dewson, III. J.H., Cowey, A., Weiskrantz, L., 1970. Disruptions of auditory sequence discrimination by unilateral and bilateral cortical ablations of superior temporal gyrus in the monkey. Exp Neurol. 28, 529-48.

Dewson, III. J.H., Pribram, K.H., Lynch, J.C., 1969. Effects of ablations of temporal cortex upon speech sound discrimination in the monkey. Exp Neurol. 24, 579-91.

Diamond, R., Carey, S., 1986. Why faces are and are not special: an effect of expertise. J. Exp. Psychol. Gen. 115, 107–117.

Dowling, W.J., 1978. Scale and contour: Two components of a theory of memory for melodies. Psychol Rev. 85, 341-354.

Dowling, W.J., 1991. Tonal strength and melody recognition after long and short delays. Percept Psychophys. 50, 305-313.

Dreisbach, L.E., Kramer, S.J., Cobos, S., Cowart, K., 2007. Racial and gender effects on pure-tone thresholds and distortion-product otoacoustic emissions (DPOAEs) in normal-hearing young adults. Int. J. Audiol. 46, 419-26.

Eals, M., Silverman, I., 1994. The hunter–gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. Ethol. Sociobiol. 15, 95-105.

Ecuyer-Dab, I., Robert, M., 2004. Have sex differences in spatial ability evolved from male competition for mating and female concern for survival? Cognition. 91, 221-57.

Ehret, G., 1987. Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. Nature. 325, 249-51.

Fecteau, S., Armony, J.L., Joanette, Y., Belin, P., 2004. Is voice processing species-specific in human auditory cortex? An fMRI study. Neuroimage. 23, 840-8.

Fitch, W.T., 2000. The evolution of speech: a comparative review. Trends Cogn Sci. 4, 258-267.

Fitch, W.T., Fritz, J.B., 2006. Rhesus macaques spontaneously perceive formants in conspecific vocalizations. J Acoust Soc Am. 120, 2132-41.

Frank, T., 1990. High-frequency hearing thresholds in young adults using a commercially available audiometer. Ear Hear. 11, 450-454.

Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2003. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. J Neurosci. 23, 5235-46.

Friedrich, A., Zentall, T., Weisman, R., 2007. Absolute pitch: frequency-range discriminations in pigeons (Columba livia): comparisons with zebra finches (Taeniopygia guttata) and humans (Homo sapiens). J Comp Psychol. 121, 95-105.

Fritz, J., Mishkin, M., Saunders, R.C., 2005. In search of an auditory engram. Proc Natl Acad Sci U S A. 102, 9359-64.

Fuster, J.M., Jervey, J.P., 1981. Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. Science. 212, 952-5.

Fuster, J.M., Jervey, J.P., 1982. Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. J Neurosci. 2, 361-75.

Geissler, D.B., Ehret, G., 2004. Auditory perception vs. recognition: representation of complex communication sounds in the mouse auditory cortical fields. Eur J Neurosci. 19, 1027-40.

Ghazanfar, A.A., Hauser, M.D., 2001. The auditory behaviour of primates: a neuroethological perspective. Curr Opin Neurobiol. 11, 712-20.

Ghazanfar, A.A., Turesson, H.K., Maier, J. X., van Dinther, R., Patterson, R.D., Logothetis, N.K., 2007. Vocal-tract resonances as indexical cues in rhesus monkeys. Curr Biol. 17, 425-30.

Gifford, G.W. 3rd., Hauser, M.D., Cohen, Y.E., 2003. Discrimination of functionally referential calls by laboratory-housed rhesus macaques: implications for neuroethological studies. Brain Behav Evol. 61, 213-24.

Gifford, G.W. 3rd., MacLean, K.A., Hauser, M.D., Cohen, Y.E., 2005. The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. J Cogn Neurosci. 17, 1471-82.

Gil-da-Costa, R., Braun, A., Lopes, M., Hauser, M.D., Carson, R.E., Herscovitch, P., Martin, A., 2004. 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. 101, 17516-21.

Goldman, P.S., Rosvold, H.E., Vest, B., Galkin, T.W., 1971. Analysis of the delayed-alternation deficit produced by dorsolateral prefrontal lesions in the rhesus monkey. J Comp Physiol Psychol. 77, 212-20.

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci. 10, 14-23.

Hackett, T.A., 2010. Information flow in the auditory cortical network. Hear Res. [Epub ahead of print]

Hackett, T.A., Stepniewska, I., Kaas, J.H., 1998. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol. 394, 475-95.

Hackett, T.A., Stepniewska, I., Kaas, J.H., 1999. Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Res. 817, 45-58.

Hallmo, P., Sundby, A., Mair, W.S., 1994. Extended high-frequency audiometry: Air- and bone-conduction thresholds, age, and gender variations. Scand. Audiol. 23, 165-170.

Harrington, I.A., Heffner, R.S., Heffner, H.E., 2001. An investigation of sensory deficits underlying the aphasia-like behavior of macaques with auditory cortex lesions. Neuroreport. 12, 1217-21.

Hauser, M.D., 1993. Right hemisphere dominance for the production of facial expression in monkeys. Science. 261, 475-7.

Hauser, M.D., 1998. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. Anim Behav. 55, 1647-58.

Hauser, M.D., 2007. When males call, females listen: sex differences in responsiveness to rhesus monkey, Macaca mulatta, copulation calls. Anim. Behav. 73, 1059-65.

Hauser, M.D., Marler, P., 1993. Food-associated calls in rhesus macaques (Macaca mulatta): I. Socioecological factors. Behav Ecol. 4, 194-205.

Hauser, M.D., McDermott, J., 2003. The evolution of the music faculty: a comparative perspective. Nat Neurosci. 6, 663-8.

Heffner, H.E., Heffner, R.S., 1984. Temporal lobe lesions and perception of species-specific vocalizations by macaques. Science. 226, 75-6.

Heffner, H.E., Heffner, R.S., 1986. Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques. J Neurophysiol. 56, 683-701.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat Rev Neurosci. 8, 393-402.

Hienz, R.D., Jones, A.M., Weerts, E.M., 2004. The discrimination of baboon grunt calls and human vowel sounds by baboons. J Acoust Soc Am. 116, 1692-7.

Höistad, M., Barbas, H., 2008. Sequence of information processing for emotions through pathways linking temporal and insular cortices with the amygdala. Neuroimage. 40, 1016-33.

Hubel, D.H., Wiesel, T.N., 1968. Receptive fields and functional architecture of monkey striate cortex. J Physiol. 195, 215-43.

Hulse, S.H., Cynx, J., Humpal, J., 1984. Absolute and relative pitch discrimination in serial pitch perception by birds. J Exp Psychol Gen. 113, 38-54.

Hurly, T.A., Ratcliffe, L., Weary, D.M., Weisman, R., 1992. White-throated sparrows (Zonotrichia albicollis) can perceive pitch change in conspecific song by using the frequency ratio independent of the frequency difference. J Comp Psychol. 106, 388-391.

Iversen, S.D., Mishkin, M., 1973. Comparison of superior temporal and inferior prefrontal lesions on auditory and non-auditory tasks in rhesus monkeys. Brain Res. 55, 355-67.

Jimura, K., Konishi, S., Miyashita, Y., 2009. Temporal pole activity during perception of sad faces, but not happy faces, correlates with neuroticism trait. Neurosci Lett. 453, 45-8.

Jonasson, Z., 2005. Meta-analysis of sex differences in rodent models of learning and memory: a review of behavioral and biological data. Neurosci Biobehav Rev. 28, 811-25.

Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A. 97, 11793-9.

Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos Trans R Soc Lond B Biol Sci. 361, 2109-28.

Keppel, G., 1982. Design and Analysis: A researcher's handbook, 2nd edn., Chapter 8. Prentice-Hall, Englewood Cliffs, NJ.

Kikuchi, Y., Horwitz, B., Mishkin, M., 2010. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. J Neurosci. 30, 13021-30.

Kim, J.W., Kim, J.J., Jeong, B.S., Ki, S.W., Im, D.M., Lee, S.J., Lee, H.S., 2005. Neural mechanism for judging the appropriateness of facial affect. Brain Res Cogn Brain Res. 25, 659-67.

Kimura, D., 1996. Sex, sexual orientation and sex hormones influence human cognitive function, Curr Opin Neurobiol. 6, 259–263.

Kimura, D., Clarke, P.G., 2002. Women's advantage on verbal memory is not restricted to concrete words. Psychol Rep. 91, 1137-42.

Kimura, D., Hampson, E., 1994. Cognitive pattern in men and women is influenced by fluctuations in sex hormones. Curr Dir Psychol Sci. 3, 57-61.

Kobatake, E., Tanaka, K., 1994. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. J Neurophysiol. 71, 856-67.

Kondo, H., Saleem, K.S., Price, J.L., 2003. Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. J Comp Neurol. 465, 499-523.

Kondo, H., Saleem, K.S., Price, J.L., 2005. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. J Comp Neurol. 493, 479-509.

Konishi, M., 1985. Birdsong: from behavior to neuron. Annu Rev Neurosci. 8, 125-70.

Kuhl, P.K., 1979. Models and mechanisms in speech perception: Species comparisons provide further contributions. Brain Behav Evol. 16, 374-408.

Kuhl, P.K., Miller, J.D., 1978. Speech perception by the chinchilla: identification function for synthetic VOT stimuli. J Acoust Soc Am. 63, 905-17.

Kupfer, K., Jurgens, U., Ploog, D., 1977. The effect of superior temporal lesions on the recognition of species-specific calls in the squirrel monkey. Exp Brain Res. 30, 75-87.

Kusmierek, P., Rauschecker, J.P., 2009. Functional specialization of medial auditory belt cortex in the alert rhesus monkey. J Neurophysiol. 102, 1606-22.

Lacreuse, A., Herndon, J.G., Killiany, R.J., Rosene, D.L., Moss, M.B., 1999. Spatial cognition in rhesus monkeys: male superiority declines with age. Horm Behav. 36, 70-6.

Lacreuse, A., Kim, C.B., Rosene, D.L., Killiany, R.J., Moss, M.B., Moore, T.L., Chennareddi, L., Herndon, J.G., 2005. Sex, age, and training modulate spatial memory in the rhesus monkey (Macaca mulatta). Behav Neurosci. 119, 118-26.

LeDoux, J.E., 1993. Emotional memory: in search of systems and synapses. Ann N Y Acad Sci. 702, 149-57.

LePrell, C.G., Niemiec, A.J., Moody, D.B., 2001. Macaque thresholds for detecting increases in intensity: effects of formant structure. Hear Res. 162, 29-42.

Levy, L.J., Astur, R.S., Frick, K.M., 2005. Men and women differ in object memory but not performance of a virtual radial maze. Behav Neurosci. 119, 853-62.

Levy, R., Goldman-Rakic, P.S., 2000. Segregation of working memory functions within the dorsolateral prefrontal cortex. Exp Brain Res. 133, 23-32.

Liu, Y., Murray, S.O., Jagadeesh, B., 2009. Time course and stimulus dependence of repetition-induced response suppression in inferotemporal cortex. J Neurophysiol. 101, 418-36.

Liu, X., Powell, D.K., Wang, H., Gold, B.T., Corbly, C.R., Joseph, J.E., 2007. Functional dissociation in frontal and striatal areas for processing of positive and negative reward information. J Neurosci. 27, 4587-97.

Lomber, S.G., Malhotra, S., 2008. Double dissociation of 'what' and 'where' processing in auditory cortex. Nat Neurosci. 11, 609-16.

Löppönen, H., Sorri, M.R.B., 1991. High-frequency air-conduction and electric bone-conduction audiometry. Age and sex variations. Scand. Audiol. 20, 181-189.

Luck, S.J., Vogel, E.K., 1997. The capacity of visual working memory for features and conjunctions. Nature. 390, 279-81.

MacDougall-Shackleton, S.A., Hulse, S.H., 1996. Concurrent absolute and relative pitch processing by European starlings (Sturnus vulgaris). J Comp Psychol. 110, 139-146.

Maeder, P.P., Meuli, R.A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J.P., Pittet, A., Clarke, S., 2001. Distinct pathways involved in sound recognition and localization: a human fMRI study. Neuroimage. 14, 802-16.

Markowitsch, H.J., Emmans, D., Irle, E., Streicher, M., Preilowski, B., 1985. Cortical and subcortical afferent connections of the primate's temporal pole: a study of rhesus monkeys, squirrel monkeys, and marmosets. J Comp Neurol. 242, 425-58.

Marler, P., 2004. Bird calls: their potential for behavioral neurobiology. Ann N Y Acad Sci. 1016, 31-44.

May, B., Moody, D.B., Stebbins, W.C., 1989. Categorical perception of conspecific communication sounds by Japanese macaques, Macaca fuscata. J Acoust Soc Am. 85, 837-847.

McMahon, D.B., Olson, C.R., 2007. Repetition suppression in monkey inferotemporal cortex: relation to behavioral priming. J Neurophysiol. 97, 3532-43.

Merzenich, M.M., Brugge, J.F., 1973. Representation of the cochlear partition of the superior temporal plane of the macaque monkey. Brain Res. 50, 275-96.

Meyers, E.M., Freedman, D.J., Kreiman, G., Miller, E.K., Poggio, T., 2008. Dynamic population coding of category information in inferior temporal and prefrontal cortex. J Neurophysiol. 100, 1407-19.

Miller, E.K., Desimone, R., 1994. Parallel neuronal mechanisms for short-term memory. Science. 263, 520-2.

Miller, E.K., Erickson, C.A., Desimone, R., 1996. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci. 16, 5154-67.

Miller, E.K., Li, L., Desimone, R., 1991. A neural mechanism for working and recognition memory in inferior temporal cortex. Science. 254, 1377-9.

Miller, E.K., Li, L., Desimone, R., 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. J Neurosci. 13, 1460-78.

Mishkin, M., Ungerleider, L.G., Macko, K.A., 1983. Object vision and spatial vision: two cortical pathways. Trends Neurosci. 6, 414-7.

Miyashita, Y., Chang, H.S., 1988. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature. 331, 68-70.

Moody, D.B., May, B., Cole, D.M., Stebbins, W.C., 1986. The role of frequency modulation in the perception of complex stimuli by primates. Exp Biol. 45, 219-32.

Moran, M.A., Mufson, E.J., Mesulam, M.M., 1987. Neural inputs into the temporopolar cortex of the rhesus monkey. J Comp Neurol. 256, 88-103.

Muhammad, R., Wallis, J.D., Miller, E.K., 2006. A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. J Cogn Neurosci. 18, 974-89.

Munoz-Lopez, M.M., Mohedano-Moriano, A., Insausti, R., 2010. Anatomical pathways for auditory memory in primates. Front Neuroanat. 4, 129.

Murray, E.A., Mishkin, M., 1998. Object recognition and location memory in monkeys with excitotoxic lesions of the amygdala and hippocampus. J Neurosci. 18, 6568-82.

Nagarajan, S.S., Cheung, S.W., Bedenbaugh, P., Beitel, R.E., Schreiner, C.E., Merzenich, M.M., 2002. Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. J Neurophysiol. 87, 1723-37.

Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Ito, K., Kojima, S., 2001. Neural substrates for recognition of familiar voices: a PET study. Neuropsychologia. 39, 1047-54.

Nakamura, K., Kubota, K., 1995. Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task. J Neurophysiol. 74, 162-7.

Nakamura, K., Kubota, K., 1996. The primate temporal pole: its putative role in object recognition and memory. Behav Brain Res. 77, 53-77.

Nakamura, K., Matsumoto, K., Mikami, A., Kubota, K., 1994. Visual response properties of single neurons in the temporal pole of behaving monkeys. J Neurophysiol. 71, 1206-21.

Nottebohm, F., Alvarez-Buylla, A., Cynx, J., Kirn, J., Ling, C.Y., Nottebohm, M., Suter, R., Tolles, A., Williams, H., 1990. Song learning in birds: the relation between perception and production. Philos Trans R Soc Lond B Biol Sci. 329, 115-24.

Olshausen, B.A., Field, D.J., 2004. Sparse coding of sensory inputs. Curr Opin Neurobiol. 14, 481-7.

Olson, I.R., Plotzker, A., Ezzyat, Y., 2007. The Enigmatic temporal pole: a review of findings on social and emotional processing. Brain. 130, 1718-31.

Osterhammel, D., Osterhammel, P., 1979. High-frequency audiometry. Age and sex variations. Scand Audiol. 8, 73-81.

Owen, A.M., Stern, C.E., Look, R.B., Tracey, I., Rosen, B.R., Petrides, M., 1998. Functional organization of spatial and nonspatial working memory processing within the human lateral frontal cortex. Proc Natl Acad Sci U S A. 95, 7721-7726.

Parr, L.A., Heintz, M., 2006. The perception of unfamiliar faces and houses by chimpanzees: influence of rotation angle. Perception. 35, 1473-83.

Parr, L.A., Heintz, M., 2008. Discrimination of faces and houses by rhesus monkeys: the role of stimulus expertise and rotation angle. Anim Cogn. 11, 467-74.

Passingham, R.E., 1975. Delayed matching after selective prefrontal lesions in monkeys. Brain Res. 92, 89-102.

Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci. 8, 976-87.

Pearson, J.D., Morrell, C.H., Gordon-Salant, S., Brant, L.J., Metter, E.J., Klein, L.L., Fozard, J.L., 1995. Gender differences in a longitudinal study of age-associated hearing loss. J Acoust Soc Am. 97, 1196-1205.

Pedreira, C., Mormann, F., Kraskov, A., Cerf, M., Fried, I., Koch, C., Quiroga, R.Q., 2010. Responses of human medial temporal lobe neurons are modulated by stimulus repetition. J Neurophysiol. 103, 97-107.

Persson, J., Habib, R., Nyberg, L., 2002. Decreased activity in inferotemporal cortex during explicit memory: dissociating priming, novelty detection, and recognition. Neuroreport. 13, 2181-5.

Petersen, M.R., Beecher, M.D., Zoloth, S.R., Green, S., Marler, P.R., Moody, D.B., Stebbins, W.C., 1984. Neural lateralization of vocalizations by Japanese macaques: communicative significance is more important than acoustic structure. Behav Neurosci. 98, 779-90.

Petkov, C.I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., Logothetis, N.K., 2008. A voice region in the monkey brain. Nat Neurosci. 11, 367-74.

Petkov, C.I., Logothetis, N.K., Obleser, J., 2009. Where are the human speech and voice regions, and do other animals have anything like them? Neuroscientist. 15, 419-29.

Plakke, B., Ng, C.W., Poremba, A., Neural correlates of auditory working and recognition memory in primate lateral prefrontal cortex. [Under review]

Poremba, A., Malloy, M., Saunders, R.C., Carson, R.E., Herscovitch, P., Mishkin, M., 2004. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature. 427, 448-51.

Poremba, A., Mishkin, M., 2007. Exploring the extent and function of higher-order auditory cortex in rhesus monkeys. Hear Res. 229, 14-23.

Poremba, A., Saunders, R.C., Crane, A.M., Cook, M., Sokoloff, L., Mishkin, M., 2003. Functional mapping of the primate auditory system. Science. 299, 568-72.

Postma, A., Izendoorn, R., De Haan, E.H., 1998. Sex differences in object location memory. Brain Cogn. 36, 334-45.

Prosen, C.A., Moody, D.B., Sommers, M.S., Stebbins, W.C., 1990. Frequency discrimination in the monkey. J Acoust Soc Am. 88, 2152-2158.

Quiroga, R.Q., Kreiman, G., Koch, C., Fried, I., 2008. Sparse but not 'grandmother-cell' coding in the medial temporal lobe. Trends Cogn Sci. 12, 87-91.

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., Fried, I., 2005. Invariant visual representation by single neurons in the human brain. Nature. 435, 1102-7.

Ralston, J.V., Herman, L.M., 1995. Perception and generalization of frequency contours by a bottlenose dolphin (Tursiops truncatus). J Comp Psychol. 109, 268-277.

Rämä, P., Courtney, S.M., 2005. Functional topography of working memory for face or voice identity. Neuroimage. 24, 224-34.

Rämä, P., Poremba, A., Sala, J.B., Yee, L., Malloy, M., Mishkin, M., Courtney, S.M., 2004. Dissociable functional cortical topographies for working memory maintenance of voice identity and location. Cereb Cortex. 14, 768-80.

Ranganath, C., Rainer, G., 2003. Neural mechanisms for detecting and remembering novel events. Nat Rev Neurosci. 4, 193-202.

Rauschecker, J.P., 1998. Cortical processing of complex sounds. Curr Opin Neurobiol. 8, 516-21.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci. 12, 718-24.

Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci U S A. 97, 11800-6.

Rauschecker, J.P., Tian, B., 2004. Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. J Neurophysiol. 91, 2578-89.

Rauschecker, J.P., Tian, B., Hauser, M., 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. Science. 268, 111-4.

Recanzone, G.H., 2000. Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. Hear Res. 150, 104-18.

Recanzone, G.H., Cohen, Y.E., 2010. Serial and parallel processing in the primate auditory cortex revisited. Behav Brain Res. 206, 1-7.

Recanzone, G.H., Guard, D.C., Phan, M.L., 2000. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. J Neurophysiol. 83, 2315-31.

Remedios, R., Logothetis, N.K., Kayser, C., 2009. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. J Neurosci. 29, 1034-45.

Rendall, D., 2003. Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. J Acoust Soc Am. 113, 3390-402.

Renier, L.A., Anurova, I., De Volder, A.G., Carlson, S., VanMeter, J., Rauschecker, J.P., 2009. Multisensory integration of sounds and vibrotactile stimuli in processing streams for "what" and "where". J Neurosci. 29, 10950-60.

Reser, D.H., Burman, K.J., Richardson, K.E., Spitzer, M.W., Rosa, M.G., 2009. Connections of the marmoset rostrotemporal auditory area: express pathways for analysis of affective content in hearing. Eur J Neurosci. 30, 578-92.

Rizk-Jackson, A.M., Acevedo, S.F., Inman, D., Howieson, D., Benice, T.S., Raber, J., 2006. Effects of sex on object recognition and spatial navigation in humans. Behav Brain Res. 173, 181-90.

Romanski, L.M., 2004. Domain specificity in the primate prefrontal cortex. Cogn Affect Behav Neurosci. 4, 421-9.

Romanski, L.M., Averbeck, B.B., Diltz, M., 2005. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophysiol. 93, 734-47.

Romanski, L.M., Bates, J.F., Goldman-Rakic, P.S., 1999a. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol. 403, 141-5.

Romanski, L.M., Goldman-Rakic, P.S., 2002. An auditory domain in primate prefrontal cortex. Nat Neurosci. 5, 15-6.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., Rauschecker, J.P., 1999b. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci. 2, 1131-6.

Royet, J.P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., & Gervais, R., 2000. Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: a positron emission tomography study. J Neurosci. 20, 7752-9.

Russ, B.E., Ackelson, A.L., Baker, A.E., Cohen, Y.E., 2008. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. J Neurophysiol. 99, 87-95.

Russ, B.E., Lee, Y.S., Cohen, Y.E., 2007. Neural and behavioral correlates of auditory Categorization. Hear Res. 229, 204–212.

Sadagopan, S., Wang, X., 2008. Level invariant representation of sounds by populations of neurons in primary auditory cortex. J Neurosci. 28, 3415-26.

Saleem, K.S., Kondo, H., Price, J.L., 2008. Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. J Comp Neurol. 506, 659-93.

Sawamura, H., Orban, G.A., Vogels, R., 2006. Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the FMRI adaptation paradigm. Neuron. 49, 307–318.

Schouten, M.E., Pols, L.C., 1989. Identification and discrimination of sweep formants. Percept Psychophys. 46, 235-244.

Seifert, L.S., 1997. Activating representations in permanent memory: different benefits for pictures and words. J Exp Psychol Learn Mem Cogn. 23, 1106-1121.

Seltzer, B., Pandya, D.N., 1978. Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. Brain Res. 149, 1-24.

Singh, N.C., Theunissen, F.E., 2003. Modulation spectra of natural sounds and ethological theories of auditory processing. J Acoust Soc Am. 114, 3394-411.

Sinnott, J.M., Adams, F.S., 1987. Differences in human and monkey sensitivity to acoustic cues underlying voicing contrasts. J Acoust Soc Am. 82, 1539-47.

Sinnott, J.M., Kreiter, N.A., 1991. Differential sensitivity to vowel continua in Old World monkeys (Macaca) and humans. J Acoust Soc Am. 89, 2421-9.

Sinnott, J.M., Owren, M.J., Petersen, M.R., 1987. Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo). J Comp Psychol. 101, 126-131.

Sinnott, J.M., Petersen, M.R., Hopp, S.L., 1985. Frequency and intensity discrimination in humans and monkeys. J Acoust Soc Am. 78, 1977-85.

Sinnott, J.M., Sachs, M.B., Hienz, R.D., 1980. Aspects of frequency discrimination in passerine birds and pigeons. J Comp Physiol Psychol. 94, 401-15.

Sommers, M.S., Moody, D.B., Prosen, C.A., Stebbins, W.C., 1992. Formant frequency discrimination by Japanese macaques (Macaca fuscata). J Acoust Soc Am. 91, 3499-510.

Stefanacci, L., Suzuki, W.A., Amaral, D.G., 1996. Organization of connections between the amygdaloid complex and the perirhinal and parahippocampal cortices in macaque monkeys. J Comp Neurol. 375, 552-82.

Sutcliffe, J.S., Marshall, K.M., Neill, J.C., 2007. Influence of gender on working and spatial memory in the novel object recognition task in the rat. Behav Brain Res. 177, 117-25.

Suzuki, W.A., Amaral, D.G., 1994. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. J Comp Neurol. 350, 497-533.

Suzuki, W.A., Miller, E.K., Desimone, R., 1997. Object and place memory in the macaque entorhinal cortex. J Neurophysiol. 78, 1062-81.

Tanaka, K., 1993. Neuronal mechanisms of object recognition. Science. 262, 685-8.

Tanaka, K., 1996. Inferotemporal cortex and object vision. Annu Rev Neurosci. 19, 109-39.

Tanaka, K., Saito, H., Fukada, Y., Moriya, M., 1991. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol. 66, 170-89.

Theunissen, F.E., Shaevitz, S.S., 2006. Auditory processing of vocal sounds in birds. Curr Opin Neurobiol. 16, 400-7.

Tian, B., Reser, D., Durham, A., Kustov, A., Rauschecker, J.P., 2001. Functional specialization in rhesus monkey auditory cortex. Science. 292, 290-3.

Tomonaga, M., 1994. How laboratory-raised Japanese monkeys (Macaca fuscata) perceive rotated photographs of monkeys: Evidence for an inversion effect in face perception. Primates. 35, 155-165.

Trehub, S.E., Thorpe, L.A., Morrongiello, B.A., 1987. Organizational processes in infants' perception of auditory patterns. Child Dev. 58, 741-9.

Voyer, D., Postma, A., Brake, B., Imperato-McGinley, J., 2007. Gender differences in object location memory: a meta-analysis. Psychon Bull Rev. 14, 23-38.

Ulanovsky, N., Moss, C.F., 2008. What the bat's voice tells the bat's brain. Proc Natl Acad Sci U S A. 105, 8491-8.

Ungerleider, L.G., Courtney, S.M., Haxby, J.V., 1998. A neural system for human visual working memory. Proc Natl Acad Sci U S A. 95, 883-90.

Wang, X., 2000. On cortical coding of vocal communication sounds in primates. Proc Natl Acad Sci U S A. 97, 11843-9.

Wang, X., 2007. Neural coding strategies in auditory cortex. Hear Res. 229, 81-93.

Wang, X., Lu, T., Snider, R.K., Liang, L., 2005. Sustained firing in auditory cortex evoked by preferred stimuli. Nature. 435, 341-6.

Wang, X., Merzenich, M.M., Beitel, R., Schreiner, C.E., 1995. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol. 74, 2685-706.

Warren, S.G., Juraska, J.M., 1997. Spatial and nonspatial learning across the rat estrous cycle. Behav Neurosci 111, 259-66.

Waters, R.S., Wilson, W.A., 1976. Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. Percept psychophys. 19, 285-289.

Weinberger, N.M., 1998. Physiological memory in primary auditory cortex: characteristics and mechanisms. Neurobiol Learn Mem. 70, 226-51.

Weiskrantz, L., Mishkin, M., 1958. Effects of temporal and frontal cortical lesions on auditory discrimination in monkeys. Brain. 81, 406-14.

Weisman, R., Njegovan, M., Ito, S., 1994. Frequency ratio discrimination by zebra finches (Taeniopygia guttata) and humans (Homo sapiens). J Comp Psychol. 108, 363-372.

Wetzel, W., Wagner, T., Ohl, F.W., Scheich, H., 1998. Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbils. Behav Brain Res. 91, 29-39.

Woloszyn, L., Sheinberg, D.L., 2009. Neural dynamics in inferior temporal cortex during a visual working memory task. J Neurosci. 29, 5494-507.

Wilson, F.A., Scalaidhe, S.P., Goldman-Rakic, P.S., 1993. Dissociation of object and spatial processing domains in primate prefrontal cortex. Science. 260, 1955-8.

Wright, A.A., 1998. Auditory list memory in rhesus monkeys. Psychol Sci. 9, 91-98.

Wright, A.A., 1999. Auditory list memory and interference processes in monkeys. J Exp Psychol Anim Behav Process. 25, 284-296.

Wright, A.A., Rivera, J.J., Hulse, S.H., Shyan, M., Neiworth, J.J., 2000. Music perception and octave generalization in rhesus monkeys. J Exp Psychol Gen. 129, 291-307.

Wright, A.A., Shyan, M.R., Jitsumori, M., 1990. Auditory same/different concept learning by monkeys. Learn Behav. 18, 287-294.

Yeterian, E.H., Pandya, D.N., 1989. Thalamic connections of the cortex of the superior temporal sulcus in the rhesus monkey. J Comp Neurol. 282, 80-97.

Yin, P., Fritz, J.B., Shamma, S.A., 2010. Do ferrets perceive relative pitch? J Acoust Soc Am. 127, 1673-80.

Zola, S.M., Squire, L.R., Teng, E., Stefanacci, L., Buffalo, E.A., Clark, R.E., 2000. Impaired recognition memory in monkeys after damage limited to the hippocampal region. J Neurosci. 20, 451-63.

Zoloth, S.R., Petersen, M.R., Beecher, M.D., Green, S., Marler, P., Moody, D.B., Stebbins, W., 1979. Species-specific perceptual processing of vocal sounds by monkeys. Science. 204, 870-3.