Empirical Constraints on the Evolutionary Origins of Music

by

Joshua H. McDermott

B.A. Brain and Cognitive Science
Harvard College, 1998

MPhil Computational Neuroscience
University College London, 2001

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Brain and Cognitive Sciences
at the
Massachusetts Institute of Technology

February 2007

Signature of Author: ______________________
Department of Brain and Cognitive Sciences
September 26, 2006

Faculty Advisor: ______________________
Edward H. Adelson
Professor of Brain and Cognitive Sciences
Thesis Supervisor

Accepted by: ______________________
Matt Wilson
Professor of Neuroscience
Chairman, Committee for Graduate Studies

Abstract

The origins and adaptive significance of music, long an elusive target, are now active topics of empirical study. I argue that empirical results can constrain discussions of the adaptive significance of music by testing whether various musical traits are innate, uniquely human, and specific to music. This thesis extends the body of empirical research along these lines, with a focus on comparative experiments in nonhuman animals. A series of studies in nonhuman primates explores whether they have preferences for sounds that might be related to music perception in humans. A second set of studies explores whether preferences for music can be instantiated in nonhuman animals by exposure to music. One study examines pet dogs, who receive extensive exposure to music courtesy of their owners. Another examines the effect of artificial music exposure on a colony of laboratory monkeys. Although there are a few potential homologies between the human response to music and that of nonhuman animals, the bulk of our results suggest dramatic differences between humans and other species. This leaves open the possibility of uniquely human music-specific capacities that might constitute an adaptation for music.

Table of Contents

Acknowledgments

I thank Ted Adelson for contributing to my initial scientific training and for support and mentorship during my time at MIT, Marc Hauser for welcoming me into his lab and collaborating with me on large portions of my PhD work, Nancy Kanwisher for endless support and guidance over the years, Jamshed Bharucha for helpful comments and constructive criticism of my thesis work, Ted Gibson for useful discussions and for serving on my oral committee, Tali Sharot for insightful comments on several of the manuscripts, John Canfield for technical support and assistance in setting up the dog experiment, and the many undergraduate research assistants who helped run the various experiments, each of whom is thanked in the relevant chapters.

Chapter 1: Introduction

portions excerpted from:

McDermott, J & Hauser, M. (2005) The origins of music: Innateness, development, and evolution. *Music Perception*, 23, 29-59.

From the perspective of cognitive science, music ranks among the most bizarre and fascinating features of human culture. Music is apparently universal, being found in every known human culture, past and present. It is incorporated into a vast array of cultural events, including weddings and funerals, religious services, dances and sporting events, as well as solitary listening sessions. It can make people feel happy or sad, so much so that music is central to modern advertising campaigns. And people throughout the world spend billions of dollars annually on the music and clubbing industries. Despite this central role in human culture, the origins and adaptive function of music remain virtually a complete mystery. Music stands in sharp contrast to most other enjoyable human behaviors (eating, sleeping, talking, sex) in that it yields no obvious benefits to those who partake of it. The evolutionary origins of music have thus puzzled scientists and philosophers alike since the time of Darwin (Darwin, 1871).

Theories about the evolution of music abound. Many have suggested that music might be a biological adaptation, with functions ranging from courtship to social cohesion in group activities like religion and war (I. Cross, 2001; Darwin, 1871; Hagen & Bryant, 2003; Huron, 2001; Merker, 2000; G. F. Miller, 2001). Still others have argued that music is not an adaptation but rather a side effect of properties of the auditory system that evolved for other purposes (Pinker, 1997). These hypotheses need not be mutually exclusive; it may well turn out that some aspects of music are the result of general purpose auditory mechanisms, and others the result of music-specific adaptations. In any case, at present there is relatively little evidence to distinguish the various hypotheses.

We suggest that rather than beginning with a debate about putative adaptive functions of music, a more reasonable goal is to establish whether any aspects of music are 1) innate, 2) unique to humans and 3) specific to music (McDermott & Hauser, 2005). Each of these three properties can help to constrain the evolutionary status of music-related traits. Innateness is a key issue because adaptations must involve changes to the genetic code. The issue of uniqueness is also critical because nonhuman animals do not make music. Traits that are present in both humans and nonhumans, and that are presumed to be homologous, must therefore not have evolved for the purpose of enabling us to produce and appreciate music. Specificity to music matters because otherwise some other function could have fueled the evolutionary process. Empirical results can thus push the debate forward by providing evidence for or against these three properties of a music-related trait. Given these goals, two main approaches are obviously relevant – developmental studies of infants, which attempt to assess the state of a human being at or shortly after birth, and comparative studies of nonhuman animals, which attempt to test whether traits are uniquely human. This thesis will focus on the latter approach, but we begin by reviewing relevant work in various areas of research as well. Because the logic behind each source of evidence is somewhat distinct, we will begin by detailing the inferential role of the different sorts of evidence we will discuss.

1. Kinds of Evidence

a. Developmental Evidence

Perhaps the most obvious way to study whether any aspects of music perception are innate is to study infants, who lack the cultural exposure that all adults have been subject to. Developmental psychology has been a particularly rich source of studies relevant to the origins of music, due in part to the development of powerful tools to probe infants' minds(see Trehub, 2003, for a review). Developmental studies can also be difficult to interpret, as infants never completely lack exposure to music, especially if one considers in utero experience during the third trimester of pregnancy when the fetus can hear.

Infants pose an experimental challenge because unlike an adult subject, they cannot verbally report on their experiences. Instead, developmental psychologists make use of the fact that changes that are salient to an infant attract its attention, which can be measured via nonverbal behavioral responses. Although the behavioral assays vary, the fundamental logic underlying the method is the same: exemplars from one category are repeatedly presented until the infant's response —sucking a non-nutritive pacifier for neonates, looking or orienting to a stimulus presentation for older infants — habituates, at which point exemplars from either the same or a different category are presented. In a classic setup, a sample of music is played repeatedly from a speaker. Once the orienting response to the music habituates, the experimenter conducts test trials, some of which introduce some change to the music sample, such as a change in key or a rearrangement of the notes. If the infant is sensitive to the change that is made, then they will tend to look longer at the speaker following the trials containing the change.

This kind of developmental approach has the virtue that it allows for tests of musical sensitivity well before infants have the capacity to speak, sing or act on the world. Nonetheless, the approach suffers from the fact that from the third trimester on, infants are exposed to an uncontrollable range of auditory experiences, some of which inevitably involve exposure to music (James, Spencer & Stepsis, 2002). It is thus difficult to assess to what extent musical competence reflects early exposure followed by rapid learning or tuning, as opposed to innate capacities. Broadly comparative studies, involving different cultures and different populations within cultures can help: convergence across these populations, in the face of significant differences in auditory experience, would provide significant evidence of an innate signature. Such cross-cultural developmental studies are understandably rare, however.

b. Comparative Evidence

Another way to limit musical exposure and its effects is to study animals, whose musical experience can be carefully controlled. There have been relatively few studies of music-related phenomena in other species (although see below for discussions of early work by Hulse, D'Amato and others), but we think the comparative approach is particularly powerful, as it can also provide constraints on evolutionary origins and adaptive specialization that are difficult to obtain in other ways.

Like a human infant, an animal cannot verbally report on its experiences, what it likes or doesn't like, what it considers the same, what it anticipates, and so on. In parallel with studies of human infants, however, animal studies have implemented a battery of tests to understand what animals think, perceive, and feel. Some of these tests are the same as those reported for infants, using the subject's spontaneous ability to orient or look longer at an unfamiliar or impossible event (Hauser & Carey, 1998). Other techniques involve training animals to detect or discriminate different classes of stimuli (e.g. Wegener, 1964). Once trained, the animals can be tested for generalization to new stimuli, the results of which can reveal the nature of their mental representations. For instance, such methods have been used to investigate whether transformations that preserve the identity of a melody in humans will also do so in animals (D'Amato, 1988; Wright, Rivera, Hulse, Shyan & Neiworth, 2000).Comparative studies are potentially useful tools for investigating the evolution of music, for at least two reasons (McDermott & Hauser, 2005). First, because nonhuman animals do not naturally produce music (as we define it; see below for discussions of animal song), any trait found in a nonhuman species cannot be part of an adaptation for music. If the trait in question is determined to be homologous to that found in humans, it must have evolved for some purpose other than that of making and perceiving music, only to be co-opted for use in music.

Second, because so much of the debate surrounding the evolution of music concerns the role of learning through exposure, it is useful to be able to precisely control an organism's musical experience. Although practical and ethical concerns preclude such an approach in humans, the music exposure of laboratory animals can be eliminated or carefully controlled. The animals can then be tested, using the methods described above, to see if they exhibit various perceptual effects found in humans. Comparative studies can thus provide insights into the evolution of music that are difficult to obtain with other methods.

c. Cross-Cultural Evidence

Other evidence comes from studies of music perception in different cultures (Nettl, 1956; 1983; Malm, 1996). Because different cultures have different musical traditions that in many cases developed independently of each other, common features provide evidence of innate constraints on what people are predisposed to perceptually discriminate, remember, and/or enjoy. As we shall see, these commonalities can either be features of the music itself, or of the patterns of perceptual judgments subjects from different cultures make. Similar insights can be gained from investigations of what music was like in ancient cultures. Again, given the large window of time separating ancient and modern cultures, similarities between musical styles from different periods might indicate that there are innate constraints on the music cultures are likely to produce. Here there is some risk that common features might have been simply passed down across the ages, and are not indications of anything built into the brain. Many features of music have, however, clearly undergone significant change over time. Those that have not most likely represent musical features that are stable given the brain's tendencies or constraints.

d. Neural Evidence

Genetic constraints on music might also be indicated by the existence of brain circuitry dedicated to music, i.e. circuitry that is used primarily during music perception or production. Such circuitry would be a candidate for an adaptation for music, just as the hypothesized functionally dedicated brain circuitry in other domains (Motion Perception: Newsome, Wurtz, Dursteler & Mikami, 1985; Face Recognition: Kanwisher, McDermott & Chun, 1997; Language: Caplan, 1995) are candidates for adaptations for those functions. Studies of patients with brain damage aim to show music-specific deficits –

patients with problems recognizing melodies, for instance, who have otherwise normal hearing and unimpaired cognitive function (Peretz & Coltheart, 2003). Such patients provide evidence that the damaged brain area is specialized for music perception, perhaps as part of a music-related adaptation. However, damage in such cases, which often results from stroke, is typically diffuse, making it hard to pinpoint specific regions as the source of the problem. A larger issue is that even if there is evidence that part of the brain functions specifically for music perception, it is difficult to rule out the possibility that the music-specific structures in question emerged through a lifetime of musical experience rather than being the product of innate constraints.

Having set out to discuss the origins of music, it might seem sensible to begin by defining what we mean by music. However, defining music is notoriously problematic given the diversity of musical phenomena that are found across the world (Nettl, 1983). Music is often said to involve combinations of tones, for instance, and yet pitch is a rather tangential component of many African musics, which rely more heavily on rhythm (Chernoff, 1979). In our view a definition of music is not particularly important at this stage as long as it is approximately clear what we refer to with the term. This might best be established ostensively over the course of the paper, but there are a few features of music that seem worth noting here at the outset. First, by music we denote structured sounds produced directly or indirectly by humans. These sounds often vary in pitch, timbre, and/or rhythm. Second, these sounds are often made to convey emotions and to produce enjoyment, though not always. Thirdly, they often have complex structure, especially of a hierarchical nature, though not always. It follows from the heterogeneity of music that any hypothetical innate constraints on music might only apply to some subset of musical phenomena, however they may be defined. That said, there are aspects of music that are likely to be universal or at least quite widespread across cultures, as we will discuss shortly. Understanding the origins of these musical features will be important even if there are musical phenomena to which they do not fully apply.

We next turn to a more detailed discussion of relevant findings in various fields of inquiry. We begin by discussing evidence for universal features of music, then turn to evidence for innate sensitivities to musical structure. From there we turn to experiments relevant to the origins of musical preferences and of the emotional responses to music. We conclude by discussing evidence for neural circuitry dedicated to music.

2. Universal Features of Music

a. Pitch

Although rhythm is arguably just as important, if not more so, to many cultures' music, pitch has received far more attention in the literature we will review. This is likely due to its importance in Western music, and the resultant theoretical ideas about how pitch functions in music. By comparison, there are fewer frameworks available to Western scholars through which to view and discuss rhythm, and perhaps for this reason it remains less well studied and documented. There are surely many revealing cross-cultural observations that could be made with regard to rhythmic properties of music, but given the current state of music research, we will confine ourselves predominantly to discussions of pitch.

In music, the relationships between pitches are generally more important than the absolute value of the pitches that are used. A melody will be recognized effortlessly even if it is transposed up or down by a fixed amount, a manipulation that alters the absolute pitch but preserves the relative pitch distances. As far as we know, relative pitch is fundamental to how music is perceived in every known culture, so much so that it is rarely cited as a universal. However, the centrality of relative pitch suggests a role for an innately specified auditory mechanism for encoding stimuli in terms of the distances between pitches. As we will see, the ability to hear relative pitch is nontrivial, and may not be shared by nonhuman animals.

Of particular importance are the relationships between pitches separated by an octave, which are generally heard as having the same pitch chroma. Every developed musical system known to Western scholars is thought to be based in part on the similarity relations the octave defines among pitches (Burns & Ward, 1982). The role of the octave in turn is thought to be partially due to the mechanisms for perceiving pitch (Terhardt, 1974), which are likely to be shared by all mammals.

Several other features of human music that seem to be universal, or nearly so, concern the structure of scales, i.e. the sets of pitches used in music. For instance, nearly every known musical culture appears to produce music from a discrete set of 5 to 7 pitches arranged within an octave range, such as the pentatonic and diatonic scales (Burns & Ward, 1982). Many have noted that the tendency to use a small set of discrete notes might be the product of well known constraints on short-term memory and categorization (G. A. Miller, 1956).

Most scales found in music around the world also share the property of having pitches separated by unequal steps, e.g. one and two semitones in the case of the diatonic scale or 2 and 3 semitones in the pentatonic scales common to many forms of indigenous music. Various explanations have been proposed for the ubiquitous presence of unequal interval scales. Most involve the fact that unequal intervals result in each note of the scale having a unique set of interval relations with the other notes of the scale (Balzano, 1980, 1982; Shepard, 1982). This makes it possible to assign different functions to different notes (e.g. the tonic) and to have a listener easily recognize which note serves each functional role in a given melody (a functional assignment which will change depending on the key). Thus, for music theoretic reasons, such unequal-step scales are perhaps more desirable, and it is possible that they have culturally evolved among many different societies for this reason. It is also possible that melodies whose notes are taken from unequal interval scales are for some reason encoded more easily by the auditory system, an idea that we will return to in a later section.

Most musical systems also feature intervals (note pairs) whose ratios approximate simple fractions. Although memory constraints are typically invoked to explain the five or seven pitches that are usually used in musical compositions, this number of discrete pitches, as well as perhaps their spacing, could also originate in a sensory or computational bias to have intervals that approximate simple integer ratios (W. J. Dowling & Harwood, 1986). Even musical systems that sound relatively foreign to the Western ear, such as those of Java and Thailand, are said to feature an interval that approximates a perfect fifth. Interestingly, although intervals with simple ratios (such as the fifth and the octave) often have structural importance in melodies, their occurrence is relatively rare, at least if one only considers the intervals between successive notes. In cultures all over the world, small intervals (one and two semitones) occur most often; the frequency of use drops exponentially with interval size above two semitones (W. J. Dowling & Harwood, 1986; Vos & Troost, 1989). Fifths and other intervals with simple ratios can be readily found in melodies, but they are usually reached via intermediate, smaller steps.

Thus despite the heterogeneity of music across the world, several common features are evident in the sets of pitches used in indigenous popular music. The focus on these aspects of pitch may reflect Western-centric biases, and their importance in music may vary from culture to culture, but their presence nonetheless suggests that music is shaped by constraints that are built into the brain. These common features will be further discussed below in the context of other methods of inquiry.

b. Lullabies

Lullabies — songs composed and performed for infants – are a particularly striking musical phenomenon found in cultures across the world, and appear to represent a true music universal. Lullabies are recognizable as such regardless of the culture (Sandra E. Trehub, Unyk, & Trainor, 1993), and even when verbal cues are obscured by low-pass filtering (Unyk, Trehub, Trainor, & Schellenberg, 1992). This suggests that there are at least some invariant musical features that characterize infant-directed music; this aspect of music directly parallels studies in language of infant-directed speech (Fernald, 1992). Lullabies are generally slow in tempo, are often characterized as simple and repetitive by adult listeners, and may feature more descending intervals than other melodies (Unyk et al., 1992). Both adults and children perform lullabies in a distinctive manner when singing to infants; listeners can pick out the version of a melody that was actually sung in the presence of an infant. Infant-directed singing tends to have a higher pitch and slower tempo than regular singing, and carries a particular timbre, jitter, and shimmer (Sandra E. Trehub, Hill, & Kamenetsky, 1997).

The characteristics of lullabies, as well as the particular acoustic properties that adults and children imbue them with when sung to infants, appear to be tailored to what infants like. When infants are played both lullabies and adult songs under similar conditions, adults who watch them on videotape judge the infants to be happier when played the lullabies than when played adult songs (S. Trehub, 2000). The fact that the preferred characteristics of lullabies are culturally universal suggests that infant preferences for lullabies are indeed innate. Further, because no other animal parent vocalizes to its offspring in anything resembling motherese or a lullaby, this style of musical expression also appears to be uniquely human. At this point the origin of lullabies and their particular features remain unknown, but their existence suggests that at least one major genre of music is predominantly innate in origin, and uniquely human.

c. Ancient instruments

Additional evidence for universal musical tendencies comes from archaeological discoveries of musical instruments and scores from thousands of years ago. If music were purely a cultural invention, one might expect ancient music to be dramatically different from modern music, given the huge cultural differences between then and now. Similarities between ancient and modern music provide a potential signature of innate constraints.

At present the earliest example of what may be a musical instrument is a bone "flute" that dates to approximately 50,000 years ago, during the middle Paleolithic (Kunej & Turk, 2000). Found in a Neanderthal campsite in Slovenia, the supposed flute was made from the femur of a bear cub, and has four visible holes. Fink (Anonymous, 1997) has noted that the distance between the second and third holes of the flute is twice that between the third and fourth holes, which is consistent with the whole and half-tones of the diatonic scale. Kunej & Turk (Kunej & Turk, 2000) constructed replicas of the fossilized flute, however, and found that although they could produce tones consistent with a diatonic scale, it was possible to produce a continuum of other tones depending on finger placement and other details of how the flute was played. There is also controversy surrounding whether this fossil was in fact used for music, as puncture holes are occasionally made in bones by carnivores in pursuit of the marrow inside, and there is no clear evidence that the holes in the fossil were made by hominids.

The earliest well-preserved musical instruments were recently found at a Neolithic site in China, and date to between 7000 BC and 5700 BC (J. Zhang, Harbottle, Wang, & Kong, 1999). These instruments are clearly flutes (some have as many as eight holes), and were made from crane bone. The best preserved of the flutes was played several times for the purposes of tonal analysis. As with the Neanderthal flute, the tones produced depend on how the instrument is played, but it was easy for a flute player to produce a diatonic scale. Although we are limited in the conclusions that can be drawn from known ancient musical instruments, their physical designs and apparent function are consistent with the notion that humans have long been predisposed to use particular sets of musical intervals.

The earliest known musical score is Sumerian, dating to approximately 1400 BC. The score was unearthed and decoded in the 1970s, and first performed by modern musicians in 1974. The scholars who decoded the piece are fairly confident that the notes largely conform to the diatonic scale (Kilmer, Crocker, & Brown, 1976). The score appears to represent notes via their interval distances from a tonic, and there is a conspicuous absence of tritone intervals. The recording made of the scholars' decoded score is reminiscent of a folk song or lullaby, and sounds more familiar than exotic. This again suggests that some central features of Western music, including the importance of a tonic note, and perhaps the prevalence of particular musical intervals, were present even before formal "Western music" existed.

The available cross-cultural and anthropological data thus are consistent with the notion that at least some features of music are universal, shared across cultures and historical eras. We now turn to studies suggesting that some aspects of sensitivity to musical structure are universal and arise in the absence of extensive exposure to music. Many of these studies are inspired by observations of apparent musical universals.

3. Sensitivity to Musical Structure

Another way to reveal innate constraints on music perception is to show that certain musical stimuli are represented or remembered more accurately than others, independent of experience. Often the structures that human subjects perceive most accurately are those that are prevalent in music across the globe, suggesting a common cause or perhaps a causal link. These sensitivity effects have the added virtue of providing measures that are well-suited to experiments in human infants and animals.

a. Developmental evidence

Many of the most interesting sensitivity effects come from studies of young infants with minimal musical experience. Much of the developmental work begins with prevalent features of Western and non-Western music (candidates for universals) and tests for sensitivity to them in infants. At the most basic level, infants as young as 8 months seem to perceive melodic pitch contours much as adults do, treating a transposed version of a melody as the same even though the tones composing the melody are different (Chang & Trehub, 1977; S. E. Trehub, Bull, & Thorpe, 1984). In contrast, if the tones are reordered, altering the melody, infants treat the tone sequence as new, directing their gaze towards the speaker through which it is played. Apparently, relative pitch changes are highly salient to infants, just as they are to adults. Infants are also capable of generalizing across tempo changes (Sandra E. Trehub & Thorpe, 1989), again demonstrating the ability to abstract melodic information from a tone sequence just as adults can. Thus some of the basic auditory perceptual abilities needed for music perception seem to be present in infants with minimal exposure to music. It remains to be seen whether these perceptual abilities are general purpose features of the mammalian auditory system, or whether they are unique to humans and perhaps evolved for music and/or speech perception; see below for discussion of related comparative studies.

Other candidate universals have also been the focus of much developmental work. We first turn to "natural" musical intervals; given the long history of interest in their possible universality and innateness, it is no surprise that they have been the subject of developmental research. Inspired no doubt by well-known Greek theories of aesthetics, Pythagoras first observed that pairs of vibrating strings whose lengths were related by simple integer ratios produced tones that sounded better together than did tones of strings with complex ratios. Centuries later, Helmholtz (Helmholtz & Ellis, 1954) famously proposed an explanation of consonance in terms of critical bands in the cochlea, claiming that dissonance is the result of "beating" between overtones of two simultaneously played sounds. Subsequent physiological investigations have shown that consonance and dissonance are indeed distinguished by these peripheral differences (Tramo, Cariani, Delgutte, & Braida, 2001). Further sensitivity to simple harmonic intervals, in which the two tones are played simultaneously, could result from the physical structure of natural sounds, whose overtones tend to be harmonic, and therefore related by simple ratios. Notably, however, simple intervals are still musically important when the notes are played in succession and peripheral interactions do not distinguish the different interval classes. Tritones (which have ratios of 32:45), for example, are rarely used in melodies (and were in fact banned from early Western music due to how difficult they were to sing), whereas simple intervals like the fifth (2:3) are more common, and often play critical roles in the structure of melodies. The reason for the "naturalness" of simple intervals in melodies is a matter of some debate, but the prevailing view is arguably that it is largely due to experience, tuned by the local culture (W. J. Dowling & Harwood, 1986; Schoenberg, 1984).

Trehub and colleagues have tested this view with a series of experiments exploring how human infants perceive musical intervals. In one early study, Trehub, Thorpe and Trainor (Sandra E. Trehub, Thorpe, & Trainor, 1990) compared short melodies containing simple intervals to "atonal" melodies that were not in any single key and had fewer simple melodic intervals. They found that infants were more sensitive to perturbations made to the typical Western melodies than they were to perturbations in "atonal" melodies. Such results suggest that infants are somehow better at encoding the structure of typical Western melodies, perhaps because they contain simple intervals. To isolate individual intervals, Schellenberg and Trehub (Schellenberg & Trehub, 1996) measured infants' sensitivity to changes made to a pair of tones when the tones were related either by simple (e.g. a perfect fifth or fourth) or complex ratios (e.g. a tritone). In one experiment the two notes of each interval were played simultaneously, while in another they were played one after the other. Critically, the notes composing each interval were pure tones. As a result, none of the stimuli, not even those in the simultaneous case, produced significant amounts of beating, which if present might have been used to detect the changes. Despite this, the authors found that infants much more readily detected changes made to simple intervals than to complex, both for simultaneously played and sequentially played tone pairs.

For the simultaneous case, the stimulus design precludes explanations in terms of beating, but the results might nonetheless be predicted if one supposes that the auditory system is attuned to harmonicity, for instance for the purpose of extracting pitch (Terhardt, 1974). The frequencies of the fifth and fourth are produced simultaneously by any harmonic complex tone - the second and third harmonics are related by a fifth, and the third and fourth harmonics by a fourth. In contrast the frequencies of a tritone are in practice not present in complex tones, being that they are related by a 32:45 ratio. Harmonic amplitudes generally drop off with increasing frequency, and due to the limited resolution of cochlear filters, only the first 8-12 harmonics of a complex are resolved to begin with. Thus one explanation of the result with simultaneous intervals is that any tendency of the auditory system to respond to harmonically related tones might produce responses to simple, and not complex, ratio intervals. These responses, if built in to the mammalian auditory system or acquired via exposure to harmonic sounds, could be used by infants and adults alike to detect changes to simple harmonic intervals, and might explain the superior performance compared to that for the tritone. They might also make simple intervals easier to remember, and could conceivably help to account for the prevalence of such intervals in human music.

Sensitivity to simple melodic intervals is more difficult to explain as the frequencies composing the intervals do not overlap in time, and thus presumably would not coactivate harmonicity detectors. It remains possible, however, that the effects are functions of exposure to tonal music. Comparative studies on this topic would be of great interest, as the exposure to different kinds of intervals could be completely controlled.

Another series of experiments was inspired by the apparent universality of scales with unequal intervals. Trehub and colleagues (Sandra E. Trehub, Schellenberg, & Kamenetsky, 1999) studied the perception of melodies composed of pitches taken from various kinds of scales, to see if scales similar to those used in indigenous musics would exhibit any perceptual advantages. They played stimuli to young infant and adult human subjects, and tested their ability to detect 1.5 semitone perturbations made to one of the notes of the melodies. In one set of conditions the pitches were drawn from the diatonic scale, in another from an unfamiliar scale with unequal intervals, and in another from an unfamiliar scale with equal intervals. The unfamiliar scales had 8 notes spanning an octave, just like the diatonic scale. Remarkably, the authors reported that the infant subjects were able to detect the perturbations made to the melodies taken from both the diatonic and unfamiliar unequal interval scale, but not to the melodies taken from the equal interval scale. Apparently there is something about unequal interval scales that makes melodies easier to perceive and/or remember. The adult subjects showed a different pattern of results. They were able to detect the changes made to melodies whose pitches came from the diatonic scale, but not the changes made to melodies taken from either of the unfamiliar scales. Evidently the exposure to music that occurs during human development renders adults insensitive to unfamiliar musical structures, paralleling the case for language acquisition. The effect seen in infants nonetheless requires explanation, as it is hard to see how it could be the product of incidental exposure, an explanation to which the interval results are more vulnerable.

As discussed earlier, the standard explanations for unequal interval scales are music theoretic in nature, involving the assignment of functional roles to different pitches, which is easier for unequal than equal interval scales (Balzano, 1980, 1982; Shepard, 1982). These explanations suppose that unequal interval scales have arisen in many different cultures because they enable certain properties of music, properties that are by hypothesis desirable to the cultures in question. However, the results of Trehub, Schellenberg and Kamenetsky (Sandra E. Trehub et al., 1999) indicate that unfamiliar unequal scales are encoded more accurately than equal interval scales, suggesting an alternative reason for their prevalence. Apparently, melodies from equal interval scales are harder to remember. It is unclear what might cause this effect, but it clearly merits further study. The effect could be an incidental side effect of some pre-existing property of the auditory system, in which case one might expect to find it in a nonhuman animal. Alternatively, if uniquely human it would be a candidate for a music-specific adaptation, which could conceivably be driven in part by the music theoretic considerations discussed above.

The studies we have discussed thus far concern sensitivity to musical structure that can be found in the absence of extensive musical experience. Although infants display an impressive array of such sensitivities, many aspects of music perception seem to require more time or exposure to develop. Several other divergent results between adults and infants support this idea. Lynch and colleagues found that American infants were equally sensitive to perturbations in Western and Javanese melodies, whereas American adults were better at detecting changes to Western melodies (Lynch, Eilers, Oller, & Urbano, 1990). This again suggests that just as is the case with language, infants are sensitive to many different types of musical structures, and lose their sensitivity to some of them with exposure to a particular kind or genre of music. Lynch and Eilers (Lynch & Eilers, 1992) found evidence that this process of acculturation can begin to have effects by a year of age, and possibly much earlier.

Several other studies have examined the development of the tonal hierarchy – the system of expectations that endows the different notes of a scale with different degrees of "stability", i.e. appropriateness (Krumhansl, 1990). For instance, in Western popular music the tonic note within a key is the most stable, in that it occurs most frequently, often with longer durations than other notes, and is expected to occur at the end of a piece. These systems of expectations, which normal listeners acquire through incidental exposure to music, are critical to the perception of tension and resolution within a piece of music (Lerdahl, 2001; Lerdahl & Jackendoff, 1983). Tonal hierarchies are culture-specific in that different cultures use different scales (sets of pitches/intervals chosen within an octave), but have been demonstrated in Western and nonwestern cultures alike (Castellano, Bharucha, & Krumhansl, 1984; Kessler, Hansen, & Shepard, 1984). The formation of tonal hierarchies likely involves the acquisition of culture-specific musical parameters, perhaps modulating innate principles as is thought to occur in language acquisition (Chomsky, 1986). Listeners probably monitor statistical regularities from musical pieces (most obviously, the number of occurrences of various notes, and their duration) that provide cues to the structure of the hierarchy. To investigate the timecourse of this acquisition, Krumhansl and Keil (Krumhansl & Keil, 1982) made a detailed assessment of tonal expectations in children of elementary school age. They found that by first grade, children hear the difference between in-key and out-of-key notes, and consider the in-key notes to be more appropriate when played in melodies. The tonal hierarchy becomes increasingly elaborated as children age; older children distinguish between notes of the tonic triad and other notes within a key just as adults do. However, even fifth and sixth graders do not evidence the full hierarchy expressed in adults. It is unclear to what extent the gradual onset is due to the maturation of the brain as opposed to the gradual accumulation of musical exposure, but the culture-specificity of the tonal hierarchy (Castellano et al., 1984) suggests that brain maturation is not the only contributing factor. Further to these findings, Trainor and Trehub (Trainor & Trehub, 1992) have found that while adults are much better at detecting changes to melodies when the changes violate key structure, 8-month old infants are just as good at detecting in-key as out-of-key changes. This again suggests that at least some aspects of diatonic key structure are learned from exposure to music and/or depend on the maturation of the brain. Trainor and Trehub (Trainor & Trehub, 1994) also found that sensitivity to implied harmony is absent in 5 year olds but present in 7 year olds, suggesting that it may be learned over time. The exposure to music that occurs after infancy thus clearly has substantial effects, and the mechanisms that allow for learning from this exposure deserve further study.

Although infants clearly have not learned as much from their limited exposure to music as adults have from a lifetime of listening, it is nonetheless difficult to account for the effects of the exposure that occurs both in the womb and in the first few months of life. A skeptic could always argue that this exposure could endow infants with the sensitivities that are measured in some of these experiments, particularly given the myriad examples of rapid learning in human infants. In utero recordings in sheep reveal that environmental sounds are rather faithfully transmitted from the environment through the uterine wall (Lecanuet, 1996), and recent studies of human infants before and after birth suggest that musical stimuli played prior to birth can be learned by the baby and recalled after birth (James, Spencer, & Stepsis, 2002). Thus any music in the environment of a pregnant mother could conceivably have some effect on the developing fetus. Even brief experience following birth could be sufficient for rapid learning of musical structure (although it might be argued that a music-specific learning mechanism might be involved). Many results from the developmental literature are thus suggestive but inconclusive because it is impossible to control for the amount of exposure to music.

b. Comparative evidence

Animals represent a complementary experimental population to human infants and adults, since their musical exposure can be rigorously controlled in an experimental setting. Evidence of music-related perceptual abilities in animals is additionally important because such abilities cannot be attributed to music-specific adaptations, if such exist. This unique inferential role of comparative data relies on the claim that nonhuman animals do not normally make or experience music, which might seem at odds with singing behavior in animals ranging from birds to gibbons and whales. We therefore begin with a discussion of animal song before proceeding to investigations of music perception in animals.

For the purposes of this paper we define music on the basis of what humans do, as that is the phenomenon we are interested in explaining. Although the boundaries of what counts as music are nebulous and ill-defined, there are several key features of interest to virtually everything that we would categorize as music. First, like language, music consists of combinations of sounds into a hierarchical structure that allows for massive variation. Second, although many might describe music as conveying a message, it is not referentially precise in the way that language is, and its medium of expression is primarily emotional. Perhaps for this reason, music is commonly produced and listened to for enjoyment, rather than communicative purposes. In some instances music is used to convey a mood to a group of people, as in wartime, sporting events, carnivals and so forth, but individuals often listen to music on their own (in Western cultures, at least, this is clearly the dominant mode of listening), in which case there is not even this most rudimentary of communicative functions. Finally, anyone can perceive and enjoy music without training (although the lifetime of exposure to music that each person in a culture receives clearly has a profound influence on their comprehension and enjoyment of music).

Birds (Catchpole & Slater, 1995), gibbons (Geissmann, 2000), and whales (Payne, 1999), among others, produce song-like structures, and these might seem plausible candidates for homologies of human music. Closer inspection reveals a host of key differences, and we think there is good reason to think that human and animal songs are neither homologous nor homoplasic, and thus have little to do with each other (see Fitch, 2006, for an alternative perspective). The fact that some animal songs sound musical to our ear is likely a coincidence, as they function as communication signals to the animals that produce them, and are typically produced only under highly restricted contexts. As Darwin pointed out, when animals sing, they do so almost exclusively in the context of courtship or territorial defense. If one were to eliminate song from the vocal repertoire of animals that sing, one would effectively cut out one of the major sources of communication, critical for survival and reproduction. Although animal songs may in some sense alter the emotions of animal listeners, no animal sings for the pure enjoyment of others, or for its own enjoyment, at least not as far as we know. When individuals sing, the structure of song is typically quite stereotyped, even though some species alter song from season to season. In most singing species, except for those that duet (e.g., neotropical wrens, gibbons), only males sing and show unique neural adaptations for song.

On a structural level, there are admittedly some parallels with human music, more so in some species than in others. Song in birds and whales seems to be generated by a rule-based system for stringing together notes into phrases, and phrases into larger themes. Within a given species, there are innate constraints on the kinds of notes used, and even on some of the ways in which they are sequenced. However, there is also a degree of variation within these constraints, although arguably much less than with human music. There is further evidence of dialects, showing that some aspects of song are learned in both songbirds and whales. One could perhaps argue that the mechanisms for producing animal songs might have been passed down to form human musical abilities, except that numerous other species more closely related to humans (e.g., chimpanzees, gorillas, orangutans) lack song. It thus seems unlikely that any resemblance between some elements of human and animal song is due to a homology. One interesting possibility is that animal song and human music were shaped by common perceptual constraints, for instance on what sorts of acoustic structures are easy to encode or remember (Hauser & McDermott, 2003). In general, though, we do not regard animal song as the most productive avenue for comparative research.

Investigations of the perception of human musical structure in animals are potentially more relevant to the evolution of human musical abilities, as the perceptual systems of animals and humans are better candidates for homologies than are the systems for producing songs. Studies of music perception in animals are few and far between, but we hope this thesis will help to inspire more of them. Most studies have involved training animals to associate some musical stimulus with a reward followed by tests of generalization.

Studies in both birds and monkeys have used such methods to study the perception of consonance and dissonance. These studies tested whether the perceptual distinction between consonance and dissonance is apparent to nonhuman animals, without regard to whether one is preferred over the other. Izumi (Izumi, 2000) trained Japanese monkeys to discriminate changes from octave intervals (consonant) to major sevenths (dissonant), and then tested for generalization to other consonant and dissonant intervals; complex tones were used to generate the stimuli. The three monkeys used in the study acquired this discrimination rather quickly (between 4 and 16 sessions of approximately 100 trials), and were then run in transfer test sessions in which different consonant and dissonant intervals were used. Izumi found that the animals reliably detected changes from consonant intervals to dissonant ones, but not the reverse. This suggests that the animals had learned to respond to the general class of dissonant stimuli, even though their prior training had been almost exclusively with one particular dissonant interval. The results thus support the notion that the beating present in dissonant stimuli is readily apparent to animals as well as humans. The speed with which the animals acquired the discrimination is further consistent with this conclusion, as is physiological evidence from macaques that beating remains salient at the level of the auditory cortex (Fishman et al., 2001). The results of the Izumi study are also at least superficially similar to the asymmetries observed in human adults and infants: their Japanese macaque subjects detect a change from a consonant interval to a dissonant one, but not the reverse. In this case, though, the effect is likely an artifact of the training procedure, in that the monkeys were trained to detect changes from consonance to dissonance, but not the reverse (as noted by the authors).

There is also some evidence that birds can be trained to discriminate consonant and dissonant chords (Hulse, Bernard, & Braaten, 1995), again consistent with the notion that beating is likely present in their cochlea as it is in ours. As with monkeys, there is no reason to suppose that birds perceive consonant chords as pleasant or less aversive than dissonant chords, but the timbral distinctions appear to be readily apparent to both species.

Hulse, Cynx and colleagues have also used operant methods in songbirds to study how they represent melodies. When exposed to a melody, humans generally extract and remember the sequence of relative pitch changes from note to note. This sequence of pitch changes, often termed the melodic contour, identifies a melody independent of the absolute pitch range in which it is played. The fact that we can easily recognize melodies and speech intonation patterns across different keys and speakers illustrates the importance of the melodic contour, and as noted above, even very young infants seem to hear and remember the sequence of relative pitch changes produced by a series of notes (S. E. Trehub et al., 1984). The extraction of relative pitch changes is something that might be a basic built-in capacity of the auditory system, shared by nonhuman animals, but it might also have evolved in humans to enable speech and music perception.

Relative pitch has some importance in birdsong recognition, at least in some species in which songs are defined by specific frequency ratios between the "notes" composing the song (Hurly, Weisman, Ratcliffe, & Johnsrude, 1991; Weary, Weisman, Lemon, & Chin, 1991; Weisman, Ratcliffe, Johnsrude, & Hurly, 1990). Typically, individuals sing at fixed frequencies that vary from bird to bird, but the ratios between frequencies is fixed for a particular species. Artificially altering the pitch intervals of a conspecific's song produces a less aggressive response, suggesting that the pitch intervals play a role in song recognition (Hurly, Ratcliffe, Weary, & Weisman, 1992; Shackleton, Ratcliffe, & Weary, 1992). There is thus some reason to expect that birds might be able to represent musical stimuli in terms of relative pitch, although the possibility remains that song recognition relies on modular acoustic analysis that might not apply to arbitrary stimuli, or that relative pitch perception is only present in those species that use it as a song cue.

To test whether birds extract relative pitch from melodies as humans do, starlings were trained to respond to falling but not rising melodies, a task they can master given enough training (Hulse, Cynx, & Humpal, 1984). The rising and falling melodies used were composed of pure tone notes and typically spanned an octave range of frequencies. After mastering the discrimination, the birds were able to generalize to novel rising and falling melodies whose notes were taken from the same frequency range. However, the pattern of responses suggested that they had memorized the absolute frequency of many of the notes of the training exemplars, and were relying on this rather than the relative pitch changes across notes. More strikingly, the birds were unable to generalize the discrimination to novel melodies transposed up or down by an octave, and whose pitches, as a result, fell outside the range in which they had been trained (Cynx, Hulse, & Polyzois, 1986; Hulse & Cynx, 1985; MacDougall Shackleton & Hulse, 1996). This result has been replicated several times in various species of birds, including some non-songbirds (Cynx, 1995). The failure to generalize occurs even when the birds are trained on melodies in a high and low range of pitches, separated by an octave, and then tested on the intermediate octave (Hulse & Cynx, 1985). Moreover, when absolute pitch is fully removed as a cue to discrimination, birds seem unable to acquire the discrimination at all (Page, Hulse, & Cynx, 1989). Apparently birds do not readily perceive relative pitch. This would appear to be at odds with the behavior of adult and infant humans, who readily recognize melodies across large transpositions, especially octave transpositions.

One caveat is that most of the studies on this topic were conducted with rising and falling melodies whose notes were separated by constant intervals (usually 2 semitones, forming whole tone scales). Such melodies are quite atypical of those used in indigenous human music, and would be considered atonal (in reference to the fact that none of the notes of such a scale has a unique set of interval relationships with the other notes). Humans recognize transpositions of novel atonal melodies less readily than transpositions of other sorts of melodies (Cuddy, Cohen, & Mewhort, 1981; W. Dowling, Kwak, & Andrews, 1995). Although the reason for this effect is unclear, as a result such stimuli arguably do not provide the strongest test of melody perception. That said, one would think that with the thousands of exposures that the birds have in these studies, humans would learn the pitch contour in great detail and have no trouble recognizing transpositions. There thus seems to be a legitimate species difference.

Birds have also been trained to discriminate complex classes of musical stimuli. Porter and Neuringer (Porter & Neuringer, 1984) trained pigeons to discriminate music by Bach from that of Stravinsky, by rewarding responses to one of the two classes of stimuli. They found that birds who were rewarded for responding to Bach and not Stravinsky generalized to Buxtehude and Scarlatti (two other classical composers), while those trained to respond to Stravinsky generalized to Carter and Piston (two other modern composers). Similarly, Watanabe and Sato (Watanabe & Sato, 1999) trained Java sparrows to discriminate Bach from Schoenberg. They found that the birds trained to respond to Bach generalized to Vivaldi, while those trained to respond to Schoenberg generalized to Carter. It is unclear what acoustic features of the musical pieces the birds in these studies were using to make their discrimination, but the results suggest that they can acquire sensitivity to some of the acoustic dimensions along which musical genres differ. It would be interesting to test birds on more controlled stimuli that differ along specific dimensions (conformity to a particular key, for instance, which is one way in which classical and "difficult" modern music differ), to get a better idea of what musical features they can learn.

Auditory operant procedures have also been employed in nonhuman primates, so far with mixed results. Interestingly, it is generally held that it is much harder to train nonhuman primates in operant conditioning paradigms with auditory stimuli than with visual stimuli, for reasons that are not well understood (D'Amato, 1988; Wegener, 1964). In contrast to songbirds and humans, nonhuman primates show weak evidence of vocal learning for their own, species-specific calls (Egnor & Hauser, 2004), which may be related to the difficulty they have with general auditory tasks. No such modality difference exists in humans, to our knowledge. Perhaps for this reason there are only a handful of primate studies involving musical structures.

D'Amato and colleagues (reviewed in (D'Amato, 1988)) conducted a series of studies in which they trained capuchin monkeys to respond to one of two melodies to get a food reward. They then presented the animals with octave transpositions of the training stimuli to test whether they had learned the melodic contour. As was the case with the birds tested in the studies described above, the capuchin monkeys performed at chance levels when transferred to these octave-transposed stimuli, demonstrating that they had not extracted a representation of the melodic contour. It is worth noting that in all cases, the animals were trained on pairs of melodies that could be differentiated on the basis of simple and often local cues, such as the absolute frequency of one or two of the notes. Further experiments by D'Amato and colleagues confirmed that the monkeys were indeed attending to these local cues rather than the global pattern of the melodies. It would be of interest to see whether nonhuman primates can learn to discriminate melodies when local cues are eliminated, thereby forcing them to learn something about the global pattern. Nonetheless, the monkeys' behavior is quite different from what one would expect from a human in a similar task. The monkeys seem to readily learn the absolute frequencies of the notes of a melodic stimulus, whereas humans would surely find it easier to remember the melodic contour. Moreover, in an additional experiment D'Amato and colleagues repeated the Hulse and Cynx (Hulse & Cynx, 1985) experiment in which an animal is trained to discriminate rising and falling melodies in both a high and a low range of pitches, and is then tested in the intermediate octave. Like the birds, the monkeys show no generalization to melodies with the same contour drawn from this intermediate octave.

As with the bird experiments on melodic contour, D'Amato and colleagues used melodic stimuli whose intervals differ notably from typical indigenous human music, in some case being taken from whole-tone scales. It would again be of interest to test the animals with typical tonal melodies drawn from a diatonic scale. However, the experiments suggest that like birds, monkeys represent melodies differently than humans do –they do not appear to represent a melody's contour in discrimination tasks as long as other means are available to do the tasks. One recent study by Brosch and colleagues (Brosch, Selezneva, Bucks, & Scheich, 2004) demonstrates that macaque monkeys can be trained, with great effort, to discriminate the direction of pitch changes when all other cues to a discrimination task are eliminated. The results are consistent with those of D'Amato in that the animals initially adopted various strategies to rely on absolute pitch rather than relative pitch changes, only learning the relative pitch discrimination when absolute pitch was removed as a cue. Moreover, the monkeys were never forced to discriminate pitch changes below half an octave in magnitude. As mentioned before, the most common melodic intervals in human music are one and two semitones, and it would be of interest to test for generalization to pitch changes this small, which normal, untrained humans readily hear.

Ohl and colleagues (Ohl, Scheich, & Freeman, 2001; Wetzel, Wagner, Ohl, & Scheich, 1998) trained gerbils to discriminate rising from falling FM sweeps. They found that the gerbils could learn to respond to rising FM sweeps independent of the absolute frequency range covered by the sweep. The sweeps used were quite large (typically an octave) and fast (250 ms or less in duration), which is rather far from the one or two semitone jumps found most commonly in human music. Moreover, they found that performance declined in test sessions where the sweep range was reduced to half an octave, which is still far greater than typical musical intervals. FM sensitive neurons in the auditory cortex of various species have been well documented (Mendelson, Schreiner, Sutter, & Grasse, 1993; Biao Tian & Rauschecker, 1994; B. Tian & Rauschecker, 2004). Unfortunately, such neurons are also typically tested with quite large FM sweeps, as the interest is mainly in their selectivity for direction and speed. As such it is unclear whether they are relevant to the representation of musical stimuli. Experiments with smaller sweeps and with stimuli consisting of successive discrete notes would help to clarify their role. Interestingly, a lesion study by Ohl and colleagues (Wetzel, Ohl, Wagner, & Scheich, 1998) found that lesions to the right hemisphere greatly impaired the discrimination of rising and falling FM sweeps, whereas similar lesions to the left hemisphere had no significant effect. As will be discussed below, this has an intriguing parallel in the human literature (Johnsrude, Penhune, & Zatorre, 2000) that suggests that the human mechanisms for relative pitch may have evolved from related mechanisms in nonhumans, even if these mechanisms have much poorer resolution in nonhumans.

Taken together, these studies suggest that animals can learn to discriminate coarse pitch changes independent of absolute frequency if they are trained on enough transposed versions of the stimuli. Although further studies testing fine pitch changes (on the order of typical musical intervals) would be useful, on the basis of these studies it does not seem that animals represent these pitch changes anywhere near as readily as humans do.

They require extensive training to extract them, and do not generalize in a way that suggests any sort of primacy for the melodic contour as a form of representation. Rather, they seem to most naturally encode musical stimuli in terms of either absolute pitch or the absolute frequency content. In contrast, human infants with no training per se appear to readily perceive and encode the relative pitch changes in melodies, suggesting that it is a representation mode that comes naturally to humans.

The comparative results reviewed thus far largely underscore the notion that animals perceive musical structures quite differently from humans. The one exception to this in the comparative literature is a recent study on rhesus monkeys (Wright, Rivera, Hulse, Shyan, & Neiworth, 2000). In contrast to other studies of melody perception in monkeys and birds, which have used go/no-go tasks in which responses to specific classes of stimuli were rewarded, Wright and colleagues trained two monkeys to make same/different judgments on successively presented nonmusical sounds, and then substituted short melodies as stimuli. As with other studies, they were interested in whether the monkeys would identify melodies as the same even if the second melody was transposed upwards or downwards so that the tones in the two instances were physically different. Their results differ markedly from the other studies in birds and monkeys in showing that like humans, monkeys display octave generalization – they tend to identify two melodies as the same if they are transposed by one or even two octaves. Octave generalization was not found for individual pitches, suggesting that it was a function of the pitch contour of the melodies. Moreover, the octave was found to have a privileged status. No generalization was obtained if the melodies were transposed by 0.5 or 1.5 octaves, leaving the notes physically closer to the originals but changing the chroma of the notes and the key of the melody. Most intriguingly, octave generalization occurred only for melodies taken from the diatonic scale. When the monkeys were tested on "atonal" melodies whose notes were chosen randomly from the full 12 tones of the chromatic scale, they responded as though the melodies an octave apart sounded different.

There are thus two key results to the study: the octave specificity of transposition recognition in tonal melodies, and the failure to recognize octave transpositions of atonal melodies. The octave specificity effect is unlikely to be the mere product of octave equivalence of the pitches composing the melodies, as no effect was found for individual pitches. It thus seems likely that the animals were recognizing the preserved melodic contour rather than transpositions of individual pitches. The failure to recognize tritone transpositions as the same could indicate that the animals were nonetheless using pitch chroma as an additional cue, or that they were somehow encoding tonal melodies relative to their key (both chroma and key change for tritone, but not octave transpositions). As for the failure to recognize atonal transpositions, one possibility is that the monkeys had difficulty remembering atonal melodies. Alternatively, the atonal melodies could have been remembered but not represented in a manner that permitted them to be matched to transposed versions (e.g. in terms of absolute rather than relative pitch). Either way, the results suggest two key conclusions: first, that the rhesus monkeys used in the study reliably encoded the relative pitch changes in the tonal melodies, and second, that tonal melodies have special status even in nonhuman primates.

Clearly, the results are quite different from what would be expected on the basis of the other studies of melody perception in birds and nonhuman primates. There are several differences in Wright's protocol that may have been key to the animals' apparent ability to extract the global features of the melodies. First, Wright's monkeys were trained to perform a same-different judgment with arbitrary stimuli, whereas all the other studies we have discussed trained animals to associate a particular stimulus or class of stimuli with reward. Wright's monkeys were thus tested on many different melodic stimuli, each of which they heard a few times at most. In contrast, the animal subjects of all the other studies reviewed here heard the rewarded melodic stimuli many times more, often completing thousands of trials in the course of learning the discrimination. Wright's animals completed many trials during the course of training as well, but with unrelated stimuli, as they were learning to make same/different judgments. This difference is substantial and could be important – as mentioned earlier, humans seem to retain fairly accurate information about the absolute pitch of highly overlearned melodies (Levitin, 1994; Schellenberg & Trehub, 2003). Of course, humans are also better at recognizing transpositions of highly overlearned melodies, so it would be surprising if exposing an animal to the same melody thousands of times in a row somehow rendered the melodic contour inaccessible. But the difference in exposure could conceivably play a role in the differing results.

The differing paradigms also place different demands on memory that could be important. To complete a trial, Wright's animals presumably had to store the first of two melody stimuli in short term memory, and then make a comparison between that stored representation and the second stimulus. In contrast, the animal subjects in the other studies reviewed here presumably acquired a representation of the rewarded stimuli in long term memory, and then compared novel stimuli to that stored representation. Thus differences between short and long-term memory could also be relevant to the pattern of results.

Another potentially important difference between the Wright et al. study and the other comparative studies of melody perception is that Wright made use of "tonal" melodies, drawn from the diatonic scale, whereas all other studies have used "atonal" melodies, the tones of which were drawn from whole-tone or other scales atypical of human music. Indeed, Wright found that tonality was the critical variable determining whether his monkeys could recognize octave transpositions. This parallels results in humans, who are generally worse at recognizing transpositions of atonal compared to tonal melodies (Cuddy et al., 1981). This result in adult humans is readily explained by the greater exposure to tonal than atonal music, but the Wright result suggests there might be a biological basis for some of this effect. At any rate this difference in stimuli between the Wright et. al study and those that preceded it could be important.

It is also possible that the observed patterns of results are due to species differences in auditory perception, with the Old World monkey macaques having evolved different perceptual abilities from the New World monkey capuchins and birds. However, the Brosch et al. (Brosch et al., 2004) study used macaques as subjects, and found, as

D'Amato had with capuchins, that they had great difficulty learning to recognize pitch changes. A species difference thus seems unlikely to account for the divergent results.

An additional caveat is in order as well, in that the monkeys' behavior is not completely consistent with human behavior in similar tasks. Key distance effects, exhibited in Wright's monkeys by the smaller number of "same" responses to tritone-transposed tonal melodies than to octave-transposed tonal melodies, are generally weak in humans, and only found in rather specific experimental circumstances (Bartlett & Dowling, 1980; Cuddy et al., 1981; Cuddy, Cohen, & Miller, 1979; W. Dowling, 1991; van Egmond, Povel, & Maris, 1996). Adult humans can generally recognize a transposed tonal melody as the same, regardless of the key it is played in (Attneave & Olson, 1971; W. Dowling & Bartlett, 1981); subjects are only somewhat worse for unfamiliar melodies transposed to "far" keys (Trainor & Trehub, 1993). The monkeys in the Wright et al. study, in contrast, almost never categorized two melodies as the same if they were played in different keys, and the magnitude of this effect is surprising given human perception. This difference could be related to the fact that Wright's monkeys had been trained to match exact replications of the same sound rather than transposed melodies, but the results are nonetheless surprising given how humans hear melodic stimuli.

Although fully reconciling the Wright results with previous work in humans and animals will, in our view, require further research, the study is significant because it is the first suggestion that animals are naturally sensitive to some signature features of human musical structure. It also raises two significant points with respect to the role of comparative data in illuminating the psychological design features of music. First, assuming the effects cannot be attributed to incidental exposure the monkeys may have had to music, the Wright study provides evidence that there are innate constraints on music perception, since the monkeys certainly did not acquire their melodic sensitivity through cultural exposure. Second, because monkeys — especially the rhesus monkeys investigated — do not produce or experience music on their own, the fact that they apparently possess musical sensitivity suggests that at least some aspects of music perception are determined by pre-existing structures in the auditory nervous system. The monkeys clearly did not evolve musical sensitivity for the purpose of listening to or producing music, which means that their sensitivity must be the byproduct of a mechanism evolved for some other purpose. This study thus provides an intriguing demonstration of the power of comparative data.

In summary, at present there are relatively few studies of musical sensitivity effects in nonhuman animals, and those that exist do not paint a completely consistent picture. Most studies suggest that animals have trouble representing the pitch contour of a melody, a representation that is key to human music perception. The Wright et al. study suggests that at least one species of monkey can recognize transpositions and therefore extract the melodic contour, but only for "tonal" melodies taken from the diatonic scale. It remains to be seen how crucial a role tonality plays in other species and paradigms.

4. Musical Preferences

The studies reviewed in the previous sections all describe sensitivity to musical structure present either across cultures or in the absence of extensive experience, suggesting that such sensitivities may be built into the way the auditory system works. Sensitivity to musical structure does not, however, explain why we like it. Clearly, many aspects of musical preferences are learned, as exhibited by the liking most humans take to the music of their own culture. Nonetheless certain elementary preferences might be innate, and could be evidenced by the prevalence of some musical features across cultures. Other explanations of such features are, however, equally viable (e.g., prevalent features might be those that make the resulting music easier to remember) and the innate presence of such preferences thus requires direct experimental tests.

Debate over whether particular musical structures are inherently pleasing dates back at least to the time of the Greeks and the interest in consonant and dissonant intervals. Although there are differences in how consonant and dissonant pairs of tones excite the peripheral auditory system, as mentioned earlier (Helmholtz & Ellis, 1954; Tramo et al., 2001), this does not explain the most important feature of the phenomenon – that, in isolation, consonance sounds good and dissonance bad. Functional imaging studies suggest that consonant and dissonant musical stimuli activate some of the same brain regions that are found to be active for other pleasant and unpleasant stimuli (Blood, Zatorre, Bermudez, & Evans, 1999), but do not explain the origins of the pleasant and unpleasant nature of the stimuli. The aesthetic responses to consonance and dissonance could themselves be acquired through cultural exposure. Perhaps surprisingly, there is relatively little cross-cultural data on the perception of consonance and dissonance. In one study, Butler and Daston (Butler & Daston, 1968) found that consonance judgments were largely similar across American and Japanese subjects. Maher (Maher, 1976) compared judgments from Indian and Canadian subjects, finding the Indian subjects to be more tolerant of dissonant intervals. To our knowledge these are the only two studies that have compared consonance perception across different cultures. This is clearly an area that would benefit from more research, as it will be important to determine to what extent the preference for consonance that is widespread among Western listeners is universal (and therefore probably innate).

Infant studies provide another way to look at innateness. Schellenberg and Trehub's experiments with infants suggest that consonant musical intervals seem to be more discriminable than dissonant ones even in the absence of extensive experience with such sounds (Schellenberg & Trehub, 1996). But do the infants hear the different intervals as pleasant and aversive just as adults do? Zentner and Kagan (Zentner & Kagan, 1996, 1998) asked just this question, playing 4-month old infants melodies composed of consonant or dissonant intervals (major thirds and minor seconds, respectively) and recording their facial and orienting reactions. Infants spent more time looking at the music source and made fewer movements for consonant melodies than to dissonant melodies, suggesting that they preferred the consonant melodies. The infant subjects also showed signs of distress (as judged by observers blind to the condition) when listening to the dissonant versions. Similar results were obtained by Trainor and Heinmiller (Trainor & Heinmiller, 1998) and earlier by Crowder and colleagues (Crowder, Reznick, & Rosenkrantz, 1991). Trainor and colleagues also replicated their results in 2-month old infants (Trainor, Tsang, & Cheung, 2002). The authors suggest that the preference for consonant melodies is innate, emerging independent of experience.

As with some of the sensitivity effects, it is conceivable that the preferences found in infants are due to early exposure. It is well-known that adults and infants tend to prefer stimuli to which they have had prior exposure, all other things being equal (Zajonc, 1968). So if the infants had heard consonant musical stimuli more than dissonant ones, which seems likely given their relative prevalence in music and given the nature of the music typically played to infants and children in Western cultures, this could conceivably have produced the observed preferences. Even if the experiential account is wrong, and the preference is innate, it is unclear whether the mechanism is part of a music-specific adaptation, or some other, domain-general mechanism. To address these issues, in this thesis we examine whether similar preferences can be found in nonhuman primates both before and after extensive exposure to music.

Consonance and dissonance in harmonic intervals is only one (and one of the most basic) of the aesthetic contrasts that are ubiquitous in music. Also of interest is the perception of melodic intervals. Generally speaking, the intervals that are consonant and dissonant when the component tones are played simultaneously also sound more and less natural, respectively, when played sequentially, even though there is no obvious physiological correlate to distinguish them, as the notes are separated in time. The origins of basic phenomena such as this as well as more complicated aspects of what make melodies sound good and bad remain largely unstudied. The prevailing view is arguably that the aesthetic judgments for melodies are largely a function of experience, tuned by the local culture (W. J. Dowling & Harwood, 1986). This hypothesis has yet to be tested, and will clearly require a richer cross-cultural sample.

In the one study to our knowledge that has tested for musical preferences in nonhuman animals, Watanabe and Nemoto (Watanabe & Nemoto, 1998) recently reported experiments on java sparrows, using a paradigm similar to that used in Chapters 2-5 of this thesis. The birds were placed in a room with three perches, rigged such that the perch they chose to rest on determined which of two kinds of music, or silence, they heard. The authors measured the amount of time the birds spent on each perch, with the assumption that the time on a perch would be related to the relative preferences for the associated auditory stimulus. In the first experiment, one perch triggered a piece by Bach, one triggered silence, and the other triggered a piece by Schoenberg. The authors found that two of their four subjects spent more time on the Bach perch than on the Schoenberg perch (the other two showed no difference). These subjects also spent more time listening to Bach than to silence. The effect replicated in the same two subjects in a second experiment with different pieces by Bach and Schoenberg. In a third experiment, the same two birds also spent more time listening to a piece by Vivaldi to one by Carter. Evidently there is some property of music by Bach and Vivaldi that causes some Java sparrows to prefer it to the modern classical music of Carter and Schoenberg. One possibility is that there is some resemblance between the songs the birds produce themselves, and some kinds of classical music. We have speculated (Hauser & McDermott, 2003) that animal vocalizations and human music might be shaped by similar constraints on auditory perception or memory, and it is possible that some features of tonal music ("natural" intervals, for instance) might be found in animal vocalizations for this reason. Perhaps some similarity for this or other reasons drives the reported preferences, although Java sparrow songs to our ears bear little resemblance to any sort of human music. Additional experiments with more controlled manipulations would help clarify what underlies the effects. It would of great interest, for instance, to know whether the birds that prefer Bach to Schoenberg would also prefer consonant intervals to dissonant ones.

Additional developmental experiments on preferences also suggest themselves. It would be particularly interesting to check for preferences for tonal over atonal music in infants (again, by tonal we refer to melodies whose notes come from the diatonic scale, and by atonal to melodies whose notes come from chromatic or whole-tone scales, and thus are not in any particular key). Although infants are reported to be more sensitive to changes to tonal melodies than to atonal ones (Sandra E. Trehub et al., 1990), it is unclear if this would translate to a preference for one over the other. In general, the relationship between the sensitivity differences often seen in infants and the musical preferences that are the most salient effect in adults merits further exploration (see below for further discussion).

5. Emotional Responses to Music

Music produces strong emotional responses in those who listen to it. At least among members of a particular culture, certain pieces of music sound happy, others sad, others contemplative, and so forth. This is remarkable given the often abstract, non-representational nature of music. How is it that a simple sequence of tones can evoke a particular emotion? In our view there are three key questions. First, what are the acoustic cues to emotion in music, second, are the responses to these cues innate, and third, why do these cues signal particular emotions? With respect to the last question, it is of particular interest to know whether associations between acoustic cues and emotion derive from nonmusical contexts, or whether they are specific to music.

In Western music, happy and sad emotions are often conveyed with fast tempos and major keys, and slow tempos and minor keys, respectively. At present, the origins of these cues remain unclear. One obvious approach would be to test whether major and minor keys have similar effects on individuals from foreign cultures, who have little to no prior exposure to Western music. Unfortunately, individuals with little exposure to Western music are in practice hard to come by, as Western music has by now permeated most of the globe. Balkwill and Thompson (Balkwill & Thompson, 1999) therefore adopted the opposite approach. They took excerpts of North Indian ragas performed with different emotional connotations, and played them to Westerners, to see if the Westerners would perceive the intended emotion. More often than not their Western subjects perceived the intended emotion, suggesting that at least some of the cues to emotion are shared across cultures. Tempo may be primarily responsible for their results, but other variables, such as melodic and rhythmic complexity as well as pitch range, also seem to be implicated.

Developmental research has also addressed these questions. In Western music, one of the primary acoustic cues to emotion is the sort of scale from which the notes of a piece are drawn. All other things being equal, pieces in major keys typically sound happy, while those in minor keys sound sad (Hevner, 1935). The major/minor distinction has been of interest to researchers for some time, and several studies suggest that children below the age of 6 years do not readily associate major and minor keys with a mood (Gerardi & Gerken, 1995; Gregory, Worrall, & Sarge, 1996; Kastner & Crowder, 1990). To separate the contributions of tempo and mode to emotional judgments in music, Peretz and colleagues manipulated the two cues independently in a recent developmental study (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001). Children of different ages were played excerpts of classical music, some happy, some sad. Happy selections were played at the same fast tempo and were written in a major key; sad selections were played at the same slow tempo and written in a minor key. To test the role of mode and tempo in perceived affect, each selection was shifted to the other tempo in one set of trials, transposed to the other mode in another set of trials, and shifted in both tempo and mode in yet another. Subjects were asked to judge whether a given stimulus sounded happy or sad. Peretz and colleagues report that the judgments of children 6-8 years old resembled those of adults in being affected both by tempo and mode changes. Five year olds, however, although responding to tempo changes, did not associate changes from major to minor keys with changes in affective content. 3- and 4-year-old children were at chance in all conditions.

The results are consistent with the idea that emotional sensitivity to mode may depend more on learning than emotional sensitivity to tempo, and might be taken to suggest that the emotional connotations of major and minor keys are not intrinsic to the key, but rather are arbitrarily acquired. As the authors note, it is also possible that the classical excerpts that were used were too complicated for the young children to comprehend. It is well-known that childhood songs ("Old MacDonald" etc.) and lullabies tend to be among the most repetitive of songs (Unyk et al., 1992), and perhaps this is because children have trouble representing or remembering musical structures that are more complex. But supposing the emotional connotations of music are acquired through learning, how might this occur? Simple associative learning is a candidate. If enough examples of happy and sad events co-occur with major and minor-key music, for instance through weddings, funerals, movies, plays, television etc., one might eventually come to automatically hear minor keys as sad and major keys as happy. Even if emotional connotations are learned rather than innate, there is still something about music that demands explanation, which is the ease with which it acquires emotional associations. Such associations are arguably more potent in music than in any other art form, and one can only speculate as to why.

Although future developmental work with simpler pieces of music will certainly help to further illuminate the role of learning in music and emotion, cross cultural studies would also be of great value. It would obviously be of interest to know to what extent the contributions of tempo and mode are culturally invariant. The predominance of major and minor modes is specific to Western music, and it remains to be seen whether other cultures will have the same emotional associations that Westerners do.

6. Dedicated Brain Mechanisms for Music?

Given speculations that certain aspects of music are innate, neuroscientists have naturally been interested in whether there is dedicated neural circuitry for music perception (Peretz & Zatorre, 2003). Neuropsychology is perhaps most relevant to this issue, as it is only by removing or inactivating part of the brain that one can show that it is necessary for a particular capacity, such as music perception. However, the recent advent of functional imaging techniques has provided another tool with which to investigate these issues, and in time will presumably yield a wealth of data about the brain networks active during music perception and production. Music perception involves many different kinds of processes, including basic perceptual analysis of pitch and rhythm information, the extraction of music-specific structures such as the tonal hierarchy in Western music, the interpretation of this structure in terms of emotions and meaning, and the interaction of these representations with memory. Neuropsychology and neuroimaging studies have targeted many of these various levels of processing, and we will discuss them in turn. We will predominantly focus on studies of individuals who have not had extensive musical training. There is an extensive literature on the effects of musical training on the brain (Gaser & Schlaug, 2003; Stewart et al., 2003), but for the most part it does not bear on the issues that are central to this paper, as training is not needed for normal human listeners to develop music comprehension skills. Studies of untrained listeners are therefore more relevant to understanding the evolutionary origins of these skills, the development of which merely requires exposure to music rather than training.

a. Pitch Perception

Music is typically described in terms of pitch and rhythm. A myriad of perceptual processes are no doubt involved in both, but thus far more attention has been devoted to pitch (though see, for example, Liégeois-Chauvel, Peretz, Babai, Laguitton & Chauvel, 1998, for a patient with meter perception deficits, Sakai et al., 1999, for a neuroimaging study on meter perception, and Alcock, Wade, Anslow & Passingham, 2000, for dissociations between impairments in melody and rhythm in the singing of brain-damaged patients) (Alcock, Wade, Anslow, & Passingham, 2000; Liegeois Chauval, Peretz, Babai, Laguitton, & Chauvel, 1998; Sakai et al., 1999). Although the neural code for pitch remains a controversial issue in auditory neuroscience, non-primary regions of auditory cortex appear to be involved (Bendor & Wang, 2005; Patterson, Uppenkamp, Johnsrude, & Griffiths, 2002; Penagos, Melcher, & Oxenham, 2004; Tramo, Shah, & Braida, 2002), and the right auditory cortex seems to be particularly important (Zatorre, 1988), with lesions therein leading to pitch perception deficits. In melodies, the relative pitch between notes is arguably more important than the absolute pitch values. It is conventional to distinguish between contour information (whether the pitch goes up or down from one note to the next) and interval information (the precise amount by which the pitch changes, which differentiates a fifth from a tritone, for instance). Both cues are generally used to discriminate melodies in normal adults.

Several studies suggest that the right temporal lobe is critical to the perception of the melodic contour. Zatorre (Zatorre, 1985) found that melody discrimination was particularly impaired by damage to right anterolateral temporal lobe regions, and although the task used did not specifically isolate melodic contour cues, the results are consistent with a role for right temporal regions in extracting the melodic contour. Johnsrude, Penhune and Zatorre (Johnsrude et al., 2000) compared pitch discrimination and pitch direction discrimination in normal controls and patients with temporal lobe excisions. Subjects were presented with two tones in succession, and either had to judge whether the two tones were different, or whether the pitch increased or decreased from the first tone to the second. The authors found that thresholds for pitch discrimination and pitch direction discrimination were comparable in normal controls and in patients with left hemisphere excisions, but that pitch direction thresholds were markedly worse in patients with right hemisphere excisions. Thresholds in such patients were approximately two semitones on average (compared to well under a semitone for normals), meaning that without the right auditory cortex, the most common note-to-note pitch changes in melodies would be imperceptible. Consistent with these results, Brechmann and Scheich (Brechmann & Scheich, 2005) found in an fMRI study that a pitch direction discrimination task activated the right auditory cortex more than the left, whereas a duration judgment task had the opposite effect. There are thus several strands of evidence suggesting that in humans, the right auditory cortex is important for pitch perception and in particular the detection of the pitch changes that make up the melodic contour.

Based on these and other imaging and anatomical data, Zatorre, Belin and Penhune (Zatorre, Belin, & Penhune, 2002) have proposed that the left and right auditory cortices serve complementary functions, resulting from the need to simultaneously optimize resolution in the temporal and frequency domains. High resolution is not possible in both domains at once, and they propose that the left auditory cortex has been optimized for temporal resolution and the right for spectral resolution. These constraints may underlie the role of the right auditory cortex in pitch perception.

If the mechanisms for pitch perception were damaged via a brain lesion or developmental disorder, music perception would be expected to be impaired. Congenital amusia, colloquially known as tone-deafness, appears to be an example of this. There are numerous anecdotal reports of tone-deaf individuals (Theodore Roosevelt and Che Guevera are alleged to have been examples), but a series of studies by Peretz and colleagues are the first rigorous investigation of the phenomena. Subjects were recruited with newspaper ads seeking "musically impaired" individuals, and then subjected to a battery of tests (Ayotte, Peretz, & Hyde, 2002; Peretz et al., 2002). They were found to have normal IQ, working memory capacity, and peripheral hearing abilities, but marked impairments in tasks of music perception. The most common deficit in such subjects is an inability to discriminate or recognize melodies, and this seems to be due to severe deficits in detecting pitch changes. Thresholds for detecting pitch changes are on the order of several semitones, at least an order of magnitude higher than those for normal subjects, and comparable to the deficits seen in patients with right hemisphere excisions (Johnsrude et al., 2000). Most of the pitch steps in typical melodies are thus below threshold for tone-deaf individuals, and it is no surprise that they are unable to recognize melodies. Recognition of environmental sounds is unimpaired, however, as is recognition of song lyrics. Rhythm perception is impaired in some but not all cases; pitch impairment is what most commonly characterizes "tone-deafness". Similar results have recently been reported by a separate group of investigators as well (Foxton, Dean, Gee, Peretz, & Griffiths, 2004).

Can congenital amusia and the results of the lesion studies discussed earlier be used to infer the presence of music-dedicated architecture? Clearly, most amusic individuals have basic perceptual deficits that are not necessarily specific to music. Indeed, in one experiment Peretz and colleagues showed that if linguistic information is removed from spoken sentences, amusic patients are impaired at detecting intonation changes in speech just as they are at detecting pitch changes in melodies (Ayotte et al., 2002; see also Patel, Peretz, Tramo & Labrecque, 1998). However, the possibility remains that the early cortical mechanisms that seem to be abnormal in those with congenital amusia, and damaged in certain lesion patients, evolved as part of an adaptation for music and/or speech perception. Both music and speech perception necessitate or at least benefit from the fine-grained perception of pitch changes – music via melodies, and speech via intonation patterns (Pierrehumbert, 1979). Pitch variations in speech are admittedly often much larger (on the order of 7 semitones for pitch accents) than those in music (which are typically 1 or 2 semitones), but many unaccented pitch changes in speech are closer to the musical norms.

The hypothesis that the fine-grained perception of pitch changes might be the product of a uniquely human mechanism receives some intriguing support from the comparative literature reviewed in Section 3c of this chapter. Nearly all studies of nonhuman animals have found that they have great difficulty extracting pitch changes, and as a result cannot generally recognize transpositions of melodies; the Wright et al. (2000) study is the one exception. In contrast, even young human infants seem to extract the melodic contour from melodies without training (S. E. Trehub et al., 1984). It is thus possible that humans have mechanisms for perceiving pitch changes that are unique among the primates, and that might have evolved to assist in speech and/or music perception. One wrinkle in the story is that the right hemisphere specialization in humans may have an analogue in non-human animals, as mentioned earlier. Right hemisphere lesions in gerbils greatly impair the discrimination of upwards and downwards FM sweeps, whereas left hemisphere lesions have no such effect (Wetzel, Ohl et al., 1998). The FM sweeps used in these studies were an octave in extent, much larger than the intervals used in music and those that have been used in experiments in humans, but the presence of a similar asymmetry in gerbils is striking. At present it is unclear whether this pattern of results would be found generally in most mammals, but it is possible that right hemisphere mechanisms present in our non-human ancestors were adapted in humans to enable the perception of fine-grained pitch changes.

In addition to perceiving the pitch change directions that contribute to the melodic contour, humans also encode the precise pitch intervals between notes. Being able to hear that a particular interval is five semitones rather than six, for instance, is probably critical to the perception of key and to the elaboration of the tonal hierarchy. Less is known about the mechanisms for extracting intervals, but evidence from neuropsychology suggests that they are distinct from the mechanisms for processing the melodic contour, perhaps relying on left hemisphere mechanisms (Peretz, 1990). It remains to be seen whether the ability to individuate different intervals is uniquely human.

b. Higher-level Musical Structure

Once basic descriptions of pitch and rhythm have been extracted from musical stimuli, musical structure can be analyzed. The mechanisms that perform this structural analysis have been the subject of both neuropsychological and neuroimaging studies. Numerous examples exist of patients with brain damage, often from strokes, who experience deficits in recognizing music, which could conceivably be due to damage to music-specific mechanisms. However, most of these patients have also had other deficits as well, particularly in language (Basso, 1993), which precludes such an explanation. Over the last decade, a number of studies have demonstrated patients with brain-damage induced music deficits who seem markedly free of linguistic or other deficits. Some such cases can be explained in terms of basic perceptual deficits in detecting pitch direction or pitch intervals (Liegeois Chauval et al., 1998; Peretz, 1990) rather than higher level music-specific impairments. However, several patients have been studied who have music deficits despite having pitch perception that is mostly normal.

One of the most interesting cases is that of G.L., who appears to have a selective deficit in knowledge of tonal organization (Peretz, 1993). As we have discussed, normal human listeners of conventional Western music encode pitches relative to a structure known as the tonal hierarchy. Different pitches within a key serve different functions, with the tonic being the most important. The mechanism that stores this structural information and links it to the pitches in a given piece of music appears to be disrupted in G.L., who has lesions in the left temporal and right frontal lobes. For most of his life G.L. was an avid music listener, but in the aftermath of the lesions complained of being unable to recognize or enjoy music. When tested in the laboratory, G.L. was found to exhibit normal pitch and melodic contour discrimination. Interval information was disrupted but still partially available as measured by melody discrimination tasks. However, his knowledge of tonality seems to have been completely destroyed by his brain lesions. G.L. was found to be insensitive to scale violations ("sour" notes, to which untrained Western listeners are acutely sensitive), gave abnormal key profiles using the probe tone paradigm introduced by Krumhansl and colleagues, did not exhibit any preferences for melodies ending on the tonic (unlike normal control subjects), and lacked the normal preference for tonal over atonal music. There thus appears to be neural circuitry that represents tonal knowledge in normal Western listeners, which has unfortunately been damaged in G.L. This circuitry is apparently distinct from that which extracts and represents the more elementary representations of melodies in terms of pitch changes and intervals.

Janata and colleagues (Janata et al., 2002) reported neuroimaging evidence for what may be related neural representations of tonal knowledge. Their results suggest the presence of a map of the perceptual space of key laid out in the prefrontal cortex of musically trained adult listeners. A small patch of cortex is active when subjects listen to music in the key of C; a different patch is active for the key of G, and so forth. Adjacent keys on the circle of fifths seem to be represented by adjacent bits of cortex; providing additional support for the hypothesis that the neural map actually represents the psychological relations between keys. The map is not fixed, but rather reorganizes itself from scan to scan, and is instantiated in regions of cortex that appear to serve other functions depending on the task. Nonetheless their results provide the first suggestion from functional imaging of where our representations of high-level musical structure might be represented in the brain, and the prefrontal regions they describe may correspond to the brain regions damaged in patient G.L.

Given that there are brain regions representing the detailed knowledge of music that is acquired by nearly all members of a culture, it is of obvious interest to establish whether the neural circuits involved are programmed into the brain from birth, and whether they serve other functions as well. In particular, many have suggested that music might rely on linguistic computational mechanisms. Perhaps surprisingly given the widespread interest in their potential relationship, relatively few studies have directly compared speech and music perception. Patel (Patel, 2003) has proposed that music and language may rely on shared computational resources tapped by syntax in language and tonality in music. Consistent with this idea, a few imaging and electrophysiology studies report activations for music-related tasks in auditory association cortex that appear to overlap with regions thought to play a role in language, such as Broca's area (Koelsch et al., 2002; Levitin & Menon, 2003; Maess, Koelsch, Gunter, & Friederici, 2001; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Unfortunately there are few studies that have directly compared music and language with matched tasks in the same subjects. Using such studies to rigorously test the specificity of the brain circuitry involved in music will clearly be an important direction for future research.

c. Memory

Another example of deficits from brain damage that appear to be specific to music can be found in C.N., a patient with bilateral temporal lobe damage, now deceased, who appears to have had music agnosia – a memory deficit specific to music. Initially this patient had pronounced perceptual deficits in music as well (Peretz, Kolinsky, Tramo, & Labrecque, 1994), but over time these abilities mostly recovered while the memory deficits remained (Peretz, 1996). Patient C.N. was unable to recognize even highly familiar melodies, despite knowing the titles and lyrics. Nor could she hum back a familiar tune when prompted with a title. Many of the song titles evoked specific memories, but the musical information that had once accompanied them appeared mostly lost or inaccessible. Only when played a melody and forced to choose between two titles could C.N. successfully identify the name of the song. C.N. was also impaired at memorizing novel melodies, even with multiple repeated exposures, and did not even show implicit effects of repeated exposure to melodies (such as preferences for previously heard tunes). In all other respects her memory appears to be normal. The pattern of results suggests that there is a particular region of memory that is dedicated to music, which happens to have been damaged in C.N.

The case of C.N. is noteworthy for our purposes as it is difficult to explain the music-specific memory deficit by appealing to deficits of a more general capacity. But does this bolster the case for innate mechanisms for processing or storing music? It is certainly intriguing that memory for music may be segregated from other types of memories, even other types of sounds. It remains possible, however, that such a segregated organization arises naturally from the demands of storing different types of stimuli, with similar stimuli being stored nearby for reasons of retrieval efficiency. In general it is difficult to ascertain what effect C.N.'s lifetime of musical experience might have had on the organization of her memory prior to the brain damage.

In sum, studies of the brain have begun to characterize the various stages involved in the complex perceptual and cognitive processes that occur when we listen to music. It appears that there are distinct mechanisms for extracting the perceptual representations of melodies, for representing knowledge of musical structure, and for encoding music in memory. At this point it remains unclear whether any music-specific mechanisms might be hard-wired into the brain and thus candidates for the product of natural selection. Together with the comparative work on relative pitch perception, the characteristics of congenital amusia patients are consistent with a uniquely human mechanism for fine-grained relative pitch perception, but without more empirical support this possibility remains highly speculative. The mechanisms that encode tonal knowledge and possibly other aspects of musical structure are also candidates for uniquely human adaptations, but it remains unclear to what extent these mechanisms function exclusively for music perception. They could simply be co-opted from mechanisms that evolved for other functions. Future research will hopefully clarify this issue.

7. Discussion

This chapter was intended to sketch out a framework for studying the evolutionary origins and adaptive significance of music, and to review the available evidence from a variety of relevant fields of inquiry. We contend that evolutionary theories of music will be facilitated if we can identify what, if any innate constraints on music are present in the brain at birth, and then determine which of these are unique to humans and specific to music. Evidence from many areas can converge to suggest innateness, and then comparative studies of other animals can establish the uniquely human, and possibly domain-specific aspects of music.

What have we learned? In our view there is suggestive evidence that, at least to some extent, the structure of music is constrained by innate features of the brain. Most obviously, music is defined in part by pitch changes. These have perceptual prominence even in young infants, suggesting that the auditory system is set up to represent stimuli in terms of such pitch changes. Perhaps unsurprisingly, a case can also be made for the biological basis of the octave, which seems predisposed to have a central role in music; it is both prevalent in every known musical system. There is also some evidence that simple ratios are predisposed to have important roles in music. They are common in the music of many cultures, modern and ancient, and have a unique perceptual status in both human infants and adults. Evidence from infant, animal, and cross-cultural studies also suggests that "tonal" melodies (composed with notes from the diatonic scale) are processed differently by the brain than atonal ones (composed with notes from equal interval scales), again at least in part independent of experience. Certain elementary musical preferences also appear to be potentially innate, as they are found in very young infants. Lullabies provide another example of innate constraints on an aspect of music, as they are apparently universal, engineered with consistent acoustic features across cultures, and unique to humans. Finally, there is some preliminary evidence that there are culturally invariant cues to emotion in music. Evidence from developmental psychology, however, is consistent with a prominent role for learning, and could indicate that the major/minor mapping is the arbitrary product of cultural exposure.

One central question is whether the prevalence and special perceptual status of these aspects of music are the result of a uniquely human adaptation or are rather a byproduct of auditory mechanisms that evolved for other purposes. It is generally accepted, for instance, that the importance of the octave derives at least in part from mechanisms for estimating the pitch of complex tones (Terhardt, 1974), which we most likely inherited from nonhuman ancestors or at least, evolved independently (Bendor & Wang, 2005; Chung & Colavita, 1976; Cynx & Shapiro, 1986; Heffner & Whitfield, 1976; Tomlinson & Schwarz, 1988). With regard to relative pitch, "natural" intervals, "tonality", and preferences, evidence for uniqueness is less clear, due in part to the relatively thin comparative database. The available data are consistent with the idea that the primacy and fine-grained resolution of relative pitch in humans might be unique among other animals, and as such would be a candidate for a relatively recent adaptation. Preferences for consonance over dissonance in animals will be examined in this thesis, but up to now have not been studied. Little is known about the status of different kinds of musical intervals or the various features of tonal music in animals. At this point, therefore, additional studies are needed before we can speculate about the evolutionary and developmental origins of such characteristics of music.

Any innate biases for music must somehow be instantiated in the brain, but at present there is little evidence for neural circuitry dedicated to music. It might well be the case that any music-specific adaptation is not anatomically punctate, in which case it might be hard to detect with functional imaging or neuropsychology. In our view, the two best candidates at this point are the mechanisms for detecting pitch changes and for representing the tonal hierarchy. Many of the other perceptual biases we have reviewed may result from general-purpose features of the auditory system. For instance, as we have discussed, it might be the case that the special perceptual status of simple intervals also derives from mechanisms for inferring pitch. Pitch perception involves analyzing the relationships between the harmonics of a complex sound signal, the frequencies of which are often related by simple ratios. Due to the demands of such analysis, the neural representations of such frequencies could perhaps develop connectivity that causes simple intervals to be processed differently than complex ones. Such connectivity could also conceivably emerge from mere exposure to harmonic sounds. Given that connections between neurons that fire at the same time tend to be strengthened over time (Sejnowski, 1999), harmonic sounds might tend to strengthen the connections between neurons coding frequencies related by simple ratios.

In studying the origins of music we would ultimately like to explain the prevalence of certain features of music as well as the aesthetic and emotional responses that are arguably its most important and salient feature. However, much of the work we have reviewed describes something quite different – perceptual sensitivity differences in human infants, adults, and nonhumans for certain musical and nonmusical structures. Sensitivity differences are no doubt a popular target of research in part because sensitivity is often easier to measure than aesthetic or emotional responses. One might nonetheless wonder as to their significance, given that they are not obviously an important part of musical experience. Sensitivity differences are important because they presumably are due to neural circuits tuned to musical structures, and as such indicate how the brain might be shaped for music. Notably, the documented sensitivity differences occur for musical structures that are prevalent in music and that adults prefer to hear in the context of music (e.g. simple ratios, tonal melodies). Although it seems unlikely that this is merely a coincidence, it is not obvious how differences in sensitivity might be causally related to preferences for some structures over others or to the prevalence of these structures in popular music. It is possible that the observed sensitivity differences could somehow result from innate generative rules for music (Lerdahl & Jackendoff, 1983). This kind of explanation would also account for the prevalence of the structures we are sensitive to, as the same rules that give rise to the sensitivity would also in part determine what people would be capable of producing. Another possibility is that sensitivity differences derive from certain signals being better encoded and remembered. In this case the prevalence of certain features of music could result from these features being transmitted with greater likelihood from person to person. At this point, these suggestions are speculative, with little supporting evidence one way or the other. We think, however, that they represent an important avenue for future research.

In addition to suggesting that certain features of music are the result of innate biases, the available evidence suggests that other aspects are learned through cultural exposure. The sensitivity that adults have to key structure (Krumhansl, 1990; Trainor & Trehub, 1992) and culturally specific scale structure is not present in infants (Lynch et al., 1990), suggesting that it is acquired through experience, perhaps subject to certain constraints. It is unclear whether there are music-specific learning mechanisms involved in this process, and experiments on the effects of musical exposure on nonhuman primates could help to clarify both the evolution and development of this capacity.

There is some evidence from young infants that certain basic musical preferences are innate, namely the preference for consonance over dissonance. However, it remains difficult to rule out a role for the exposure even young infants have had since birth. To clarify the role of exposure, some of the experiments to be described will attempt to assess the effect of comparable exposure on nonhuman animals. It would also be useful to test infants and adults from nonwestern cultures to see if they have similar preferences, at least with respect to the consonant/dissonant distinction.

One point is clear: much remains to be studied. A number of topics have scarcely been touched at all, and we have therefore shied away from them in this review. The enjoyment of rhythm and dancing is also apparently universal – people everywhere like to move to music, and dancing is as ubiquitous as is the performance and enjoyment of music itself. At present, very little is known about these most mysterious features of music. Human adults are often spontaneously entrained to musical rhythms, moving in time with the beat. As with other aspects of music perception, we would like to determine whether this response to music is learned or innate, and whether it is uniquely human. The entrainment that occurs during dancing could simply be learned from examples, but entrainment also occurs outside the realm of dance, as evidenced by head nodding and foot tapping, which apparently are often executed unconsciously. Studies in young infants could in principle help to address how, and to what extent, experience shapes this capacity. However, Trehub (personal communication) reports that very young infants, although clearly engaged by music, do not move much at all in response to it. Indeed, this may be the main reason that parents sing to their infants. Although infants start to move to rhythmic music towards the end of their first year, and often move somewhat more to music with fast tempos, their movements are not synchronized to the music. Synchronized movement to music does not emerge until at least several years later. The capacity to become entrained by rhythms is likely limited to some extent by the development of motor coordination, but in any case behavioral studies in human infants seem unlikely to resolve the role of learning in this behavior. Comparative studies of these issues would clearly also be of interest, as the role of learning and musical exposure could be controlled. While there are a few examples of animals maintaining a musical rhythm after having been trained to do so (e.g. the Thai elephant orchestra), we know of no demonstrations of animals spontaneously becoming entrained to rhythms as people do. There are numerous well-known examples of elaborate dances in animal courtship rituals, but as with animal song, these have a very specific and narrow function. We think animal dances are unlikely to be related to dance in humans, but it might nonetheless be profitable to examine whether movements in such dances are synchronized to song.

We have also avoided discussion of the literature on perfect or absolute pitch (Zatorre, 2003), mainly because it is a rare phenomenon that is of questionable relevance to music perception in the vast majority of listeners. It has been suggested, however, that all humans are born with absolute pitch, and that most people learn to hear relative pitch as the result of exposure to speech and music, in which the absolute pitch varies and what matters is the relative pitch of the different syllables or tones. Consistent with this notion, some evidence suggests that infants rely more on absolute pitch than do adults (Saffran, 2003; Saffran & Griepentrog, 2001), although infants seem to readily hear the relative pitch sequences that define a melody's contour (S. E. Trehub et al., 1984). It is important to note, though, that any hypothetical enhanced absolute pitch perception in infants would be quite different from that in adults with perfect pitch, in which verbal category labels are key. Normal humans, though lacking the large number of fixed pitch categories found in those possessing perfect pitch, nonetheless have the usual ability to make limited absolute judgments along a perceptual dimension (G. A. Miller, 1956), and it is unclear that infants can do any better. They may simply place more emphasis on crude absolute pitch information than they do on relative pitch. Complicating matters are recent suggestions that normal humans may have much better memory for absolute pitch than has been traditionally thought, at least for familiar melodies (Levitin, 1994; Schellenberg & Trehub, 2003). As noted earlier, the capacity for fine-grained relative pitch perception is key to music perception and might be unique to humans; further experiments clarifying the relationship between and development of absolute and relative pitch would be most useful.

One of the most interesting aspects of music, particularly from the standpoint of cognitive science, is the modulation of tension that occurs within a piece (Lerdahl, 2001; Lerdahl & Jackendoff, 1983; Narmour, 1990). This is in part a function of cognitive representations of tonality (Krumhansl, 1990, 1996; Smith & Cuddy, 2003), which appear to be learned, perhaps subject to innate constraints. But the very fact that tension and resolution are associated with music is itself interesting and, we think, worthy of investigation. We suspect that certain acoustic cues to tension (increases in volume or tempo, for instance) may well be innate, and could be present in animals (we explore this issue in Chapter 3). The origins of the association between tension and various other cues, such as those that involve tonality, are less clear. Given the importance of introducing and resolving tension in music, this will certainly be a worthwhile focus of future research, but we have largely avoided its discussion due to the paucity of work on the topic.

In general, our coverage of music perception in this review has been highly Western-centric. This is mainly a reflection of the state of music perception research, most of which occurs in the west and is conducted by researchers who are most familiar with Western music, using subjects who have generally had extensive exposure to Western music. It is important to note, though, that many features of Western music are not as central in other cultures (harmony being perhaps the most notable), and that many features of music in other cultures are not as central in Western music (complex rhythms, for instance). Rhythm is arguably more important than melody and harmony in many cultures' music, and we have almost entirely avoided its discussion. This is again mainly because there is less empirical work addressing whether any aspects of rhythm perception are innate (though see (Hannon & Trehub, 2005a; Phillips-Silver & Trainor, 2005)). It is also worth noting that the function of music in Western culture is not entirely representative of how music is used in other cultures. In many indigenous cultures music is often more closely tied to particular rituals, and indeed one of the most obvious universal properties of music is its association with religion (Nettl, 1983), which occurs in every culture so far as we know. Although we consider one of the most interesting features of music to be the fact that it is often made purely for enjoyment's sake, it is unclear how often this is actually the case in less developed cultures. The differences between Western music and those of the many other cultures around the globe should thus not be overlooked.

References

Alcock, K. J., Wade, D., Anslow, P., & Passingham, R. E. (2000). Pitch and timing abilities in adult left-hemisphere-dysphasic and right-hemisphere-damaged subjects. *Brain and Language, 75*(1), 47-65.

Anonymous. (1997). Early music. *Science, 276*, 205.

Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology, 84*(2), 147-166.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain, 125*(Pt 2), 238-251.

Balch, W. R., & Lewis, B. S. (1999). Music-dependent memory: The role of tempo changes and mood mediation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1354-1363.

Balkwill, L. L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception, 17*(1), 43-64.

Balzano, G. J. (1980). The group-theoretic description of 12-fold and microtonal pitch systems. *Computer Music Journal, 4*(4), 66-84.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind and brain: the neuropsychology of music* (pp. 321–351). New York: Plenum.

Bartlett, J. C., & Dowling, W. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance, 6*(3), 501-515.

Baru, A. V. (1975). Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration, and number of formants) in dog. In G. Fant & M. A. A. Tatham (Eds.), *Auditory Analysis and Perception of Speech* (pp. 91-101). New York: Academic Press.

Basso, A. (1993). Amusia. In H. Spinler & F. Boller (Eds.), *Handbook of neuropsychology* (Vol. 8, pp. 391-410). Amsterdam New York: Elsevier.

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature, 426*, 1161-1165.

Bickerton, D. (1975). *Dynamics of a creole system.* Cambridge: Cambridge University Press.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics, 58*(125-141).

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2*(4), 382-387.

Boulez, P. (1971). *Boulez on music today.* Cambridge, Mass.,: Harvard University Press.

Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cerebral Cortex, 15*(5), 578-587.

Brosch, M., Selezneva, E., Bucks, C., & Scheich, H. (2004). Macaque monkeys discriminate pitch relationships. *Cognition, 91*(3), 259-272.

Burns, E. M., & Ward, W. D. (1982). Intervals, scales, and tuning. In D. Diana (Ed.), *The psychology of music* (pp. 241-269). San Diego: Academic Press.

Butler, J. W., & Daston, P. G. (1968). Musical consonance as musical preference: A cross-cultural study. *Journal of General Psychology, 79*, 129-142.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113*(3), 394-412.

Catchpole, C. K., & Slater, P. J. B. (1995). *Bird song: Biological Themes and Variations*. Cambridge: Cambridge University Press.

Chang, H. w., & Trehub, S. E. (1977). Auditory processing of relational information by young infants. *Journal of Experimental Child Psychology, 24*(2), 324-331.

Chomsky, N. (1986). *Knowledge of language : its nature, origins, and use*. New York: Praeger.

Chung, D. Y., & Colavita, F. B. (1976). Periodicity pitch perception and its upper frequency limit in cats. *Perception and Psychophysics, 20*(6), 433-437.

Cleveland, J., & Snowdon, C. T. (1981). The complex vocal repertoire of the adult cotton-top tamarin, Saguinus oedipus oedipus. *Zeitschrift fur Tierpsychologie, 58*, 231-270.

Cleveland, J., & Snowdon, C. T. (1982). The complex vocal repertoire of the adult cotton-top tamarin, Saguinas oedipus oedipus. *Zeitschrift fur Tierpsychologie, 58*, 231-270.

Cross, H. A., Holcomb, C. G., & Matter, W. W. (1967). Imprinting or exposure learning in rats given early auditory stimulation. *Psychonomic Science, 7*, 233-234.

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Science, 930*, 28-42.

Crowder, R. G., Reznick, J., & Rosenkrantz, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society, 29*(3), 187-188.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance, 7*(4), 869-883.

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology, 33*(3), 148-157.

Curtiss, S. (1977). *Genie: A psycholinguistic study of a modern day "wild child"*. New York: Academic Press.

Cynx, J. (1995). Similarities in absolute and relative pitch perception in songbirds (starling and zebra finch) and a nonsongbird (pigeon). *Journal of Comparative Psychology, 109*(3), 261-267.

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). *Journal of Experimental Psychology: Animal Behavior Processes, 12*(4), 394-402.

Cynx, J., & Shapiro, M. (1986). Perception of missing fundamental by a species of songbird (Sturnus vulgaris). *Journal of Comparative Psychology, 100*(4), 356-360.

D'Amato, M. R. (1988). A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune. *Music Perception, 5*(4), 453-480.

Dahlhaus, C., & Gjerdingen, R. O. (1990). *Studies in the Origin of Harmonic Tonality.* Princeton, NJ: Princeton University Press.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode. *Cognition, 80,* B1-B10.

Darwin, C. (1871). *The descent of man and selection in relation to sex.* London: John Murray.

Dowling, W. (1991). Tonal strength and melody recognition after long and short delays. *Perception and Psychophysics, 50*(4), 305-313.

Dowling, W., & Bartlett, J. C. (1981). The importance of interval information in long-term memory for melodies. *Psychomusicology, 1*(1), 30-49.

Dowling, W., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception and Psychophysics, 57*(2), 136-149.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition.* Orlando: Academic Press.

Egnor, S. E., & Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in Neurosciences, 27*(11), 649-654.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papousek, U. Jurgens & M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262-282). Cambridge: Cambridge University Press.

Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., et al. (2001). Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology, 86*(6), 2761-2788.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, in press.*

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*(4), 801-810.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*(27), 9240-9245.

Geissmann, T. (2000). Gibbon songs and human music from an evolutionary perspective. In B. Merker & N. L. Wallin (Eds.), *The origins of music.* (pp. 103-123). Cambridge, MA, US: The MIT Press.

Gerardi, G. M., & Gerken, L. (1995). The development of affective responses to modality and melodic contour. *Music Perception, 12*(3), 279-290.

Gregory, A. H., Worrall, L., & Sarge, A. (1996). The development of emotional responses to music in young children. *Motivation and Emotion, 20*(4), 341-348.

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature, 14*(1), 21-51.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48-55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, 102,* 12639-12643.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science, 298*, 1636-1639.

Hauser, M. D., & McDermott, J. (2003). The evolution of the music faculty: A comparative perspective. *Nature Neuroscience, 6*(7), 663-668.

Heffner, H. E. (1983). Hearing in large and small dogs: Absolute thresholds and size of the tympanic membrane. *Behavioral Neuroscience, 97*, 310-318.

Heffner, H. E., & Whitfield, I. C. (1976). Perception of the missing fundamental by cats. *Journal of the Acoustical Society of America, 59*, 915-919.

Helmholtz, H. v., & Ellis, A. J. (1954). *On the sensations of tone as a physiological basis for the theory of music* (2d English , ed.). New York,: Dover Publications.

Hess, E. H. (1973). *Imprinting: Early experience and the developmental psychobiology of attachment.* New York: Van Nostrand Rheinhold.

Hevner, K. (1935). The affective character of the major and minor modes in music. *American Journal of Psychology, 47*, 103-118.

Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology, 49*, 621-630.

Hubel, D. H., & Weisel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology, 206*, 419-436.

Hulse, S. H., Bernard, D. J., & Braaten, R. F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (Sturnus vulgaris). *Journal of Experimental Psychology: General, 124*(4), 409-423.

Hulse, S. H., & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (Mimus, Molothrus, and Sturnus). *Journal of Comparative Psychology, 99*(2), 176-196.

Hulse, S. H., Cynx, J., & Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *Journal of Experimental Psychology: General, 113*(1), 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. *Journal of Comparative Psychology, 106*(4), 388-391.

Hurly, T. A., Weisman, R. G., Ratcliffe, L., & Johnsrude, I. S. (1991). Absolute and relative pitch production in the song of the white-throated sparrow. *Bioacoustics, 3*, 81-91.

Huron, D. (1991). Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Perception, 9*(2), 135-154.

Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception, 11*(3), 289-305.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Science, 930*, 43-61.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities: Re-examination of the "Mozart effect". *Music Perception, 20*, 151-171.

Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America, 108*(6), 3073-3078.

James, D. K., Spencer, C. J., & Stepsis, B. W. (2002). Fetal learning: a prospective randomized controlled study. *Ultrasound in Obstetrics and Gynecology, 20*(5), 431-438.

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science, 298*(5601), 2167-2170.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second-language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-90.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123*(1), 155-163.

Kastner, M. P., & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception, 8*(2), 189-201.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the West. *Music Perception, 2*(2), 131-165.

Kilmer, A. D., Crocker, R. L., & Brown, R. R. (1976). Sounds from silence. On *recent discoveries in ancient Near Eastern music*. Berkeley, CA: Bit Enki Records.

Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience, 16*, 1412-1425.

Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage, 17*(2), 956-966.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension and musical ideas. *Music Perception, 13*(3), 401-432.

Krumhansl, C. L., & Keil, F. C. (1982). Acquisition of the hierarchy of tonal functions in music. *Memory and Cognition, 10*(3), 243-251.

Kunej, D., & Turk, I. (2000). New perspectives on the beginnings of music: Archaeological and musicological analysis of a middle paleolithic bone "flute". In B. Merker & N. L. Wallin (Eds.), *The origins of music*. (pp. 235-268). Cambridge, MA, US: The MIT Press.

Lecanuet, J. P. (1996). Prenatal auditory experience. In I. Deliège & J. A. Sloboda (Eds.), *Musical beginnings : origins and development of musical competence* (pp. 3-34). Oxford: Oxford University Press.

Lerdahl, F. (2001). *Tonal pitch space*. New York: Oxford University Press.

Lerdahl, F., & Jackendoff, R. S. (1983). *A generative theory of tonal music*. Cambridge, Mass.: MIT Press.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception and Psychophysics, 56*(4), 414-423.

Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *Neuroimage, 18*(1), 74-82.

Liegeois Chauval, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121*(10), 1853-1867.

Lynch, M. P., & Eilers, R. E. (1992). A study of perceptual development for musical tuning. *Perception and Psychophysics, 52*(6), 599-608.

Lynch, M. P., Eilers, R. E., Oller, D., & Urbano, R. C. (1990). Innateness, experience, and music perception. *Psychological Science, 1*(4), 272-276.

MacDougall Shackleton, S. A., & Hulse, S. H. (1996). Concurrent absolute and relative pitch processing by European starlings (Sturnus vulgaris). *Journal of Comparative Psychology, 110*(2), 139-146.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*(5), 540-545.

Maher, T. F. (1976). "Need for resolution" ratings for harmonic musical intervals: A comparison between Indians and Canadians. *Journal of Cross Cultural Psychology, 7*(3), 259-276.

Marler, P. (1955). Characteristics of some animal calls. *Nature, 176*, 6-7.

Marler, P. (1970). A comparative approach to vocal learning: Song learning in white-crowned sparrows. *Journal of Comparative Physiological Psychology, 71*, 1-25.

McDermott, J. (2006). Can music exposure induce preferences for consonance? Evidence from pet dogs. *Submitted.*

McDermott, J., & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition, 94*(2), B11-B21.

McDermott, J., & Hauser, M. D. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23*(1), 29-59.

McDermott, J., & Hauser, M. D. (2006). Nonhuman primates prefer slow tempos but dislike music overall. *Cognition, In press.*

Mendelson, J. R., Schreiner, C. E., Sutter, M. L., & Grasse, K. L. (1993). Functional topography of cat primary auditory cortex: responses to frequency-modulated sweeps. *Experimental Brain Research, 94*(1), 65-87.

Merker, B. (2000). Sychronous chorusing and human origins. In B. Merker & N. L. Wallin (Eds.), *The origins of music.* (pp. 315-327). Cambridge, MA, US: The MIT Press.

Meyer, M. (1903). Experimental studies in the psychology of music. *American Journal of Psychology, 14*, 456-478.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352.

Miller, G. F. (2001). *The mating mind : how sexual choice shaped the evolution of human nature* (1st Anchor Books ed.). New York: Anchor Books.

Nakahara, H., Zhang, L. I., & Merzenich, M. M. (2004). Specialization of primary auditory cortex processing by sound exposure in the "critical period". *Proceedings of the National Academy of Sciences, 101*(18), 7170-7174.

Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development, 27*(4), 455-464.

Narmour, E. (1990). *The analysis and cognition of basic melodic structures : the implication-realization model.* Chicago: University of Chicago Press.

Nettl, B. (1983). *The study of ethnomusicology : twenty-nine issues and concepts.* Urbana: University of Illinois Press.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11-28.

Ohl, F. W., Scheich, H., & Freeman, W. J. (2001). Change in the pattern of ongoing cortical activity with auditory category learning. *Nature, 412*(6848), 733-736.

Page, S. C., Hulse, S. H., & Cynx, J. (1989). Relative pitch perception in the European starling (Sturnus vulgaris): Further evidence for an elusive phenomenon. *Journal of Experimental Psychology: Animal Behavior Processes, 15*(2), 137-146.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*(7), 674-681.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717-733.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36*(4), 767-776.

Payne, K. (1999). The progressively changing songs of humpback whales: a window on the creative process in a wild animal. In B. Merker & N. L. Wallin (Eds.), *The origins of music.* (pp. 135-150). Cambridge, MA, US: The MIT Press.

Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience, 24*(30), 6810-6815.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113 (Pt 4)*, 1185-1205.

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology, 10*(1), 21-56.

Peretz, I. (1996). Can we lose memory for music? A case of music agnosia in a nonmusician. *Journal of Cognitive Neuroscience, 8*(6), 481-496.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron, 33*(2), 185-191.

Peretz, I., Kolinsky, R., Tramo, M., & Labrecque, R. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117*(6), 1283-1301.

Peretz, I., & Zatorre, R. J. (Eds.). (2003). *The cognitive neuroscience of music.* Oxford New York: Oxford University Press.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat in music: Movement influences rhythm perception in infants. *Science, 308*, 1430.

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. *Journal of the Acoustical Society of America, 66*, 363-369.

Pinker, S. (1997). *How the mind works* (1st ed.). New York: Norton.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548-560.

Porter, D., & Neuringer, A. (1984). Music discriminations by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 10*(2), 138-148.

Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science, 6*(1), 35-43.

Saffran, J. R., & Griepentrog, G. J. (2001). Absolute pitch in infant auditory learning: Evidence for developmental reorganization. *Developmental Psychology, 37*(1), 74-85.

Saffran, J. R., Loman, M. M., & Robertson, R. R. (2000). Infant memory for musical experiences. *Cognition, 77*(1), B15-23.

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *Journal of Neuroscience, 19*(22), 10074-10081.

Schellenberg, E., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science, 7*(5), 272-277.

Schellenberg, E., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14*(3), 262-266.

Schoenberg, A. (1984). *Style and idea.* Berkeley: University of California Press.

Seiden, H. R. (1958). *Auditory acuity of the marmoset monkey (Hapale jacchus).* Unpublished Unpublished Doctoral Dissertation, Princeton University.

Sejnowski, T. J. (1999). The book of Hebb. *Neuron, 24*(4), 773-776.

Serpell, J. e. (1995). *The Domestic Dog, its evolution, behavior and interactions with people.* Cambridge, UK: Cambridge University Press.

Shackleton, S. A., Ratcliffe, L., & Weary, D. M. (1992). Relative frequency parameters and song recognition in black-capped chickadees. *Condor, 94*, 782-785.

Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *The psychology of music* (pp. 343-390). San Diego: Academic Press.

Smith, N. A., & Cuddy, L. L. (2003). Perceptions of musical dimensions in Beethoven's Waldstein sonata: An application of Tonal Pitch Space theory. *Musicae Scientiae, 7*(1), 7-34.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Becoming a pianist. An fMRI study of musical literacy acquisition. *Annals of the New York Academy of Science, 999*, 204-208.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55*, 1061-1069.

Tian, B., & Rauschecker, J. P. (1994). Processing of frequency-modulated sounds in the cat's anterior auditory field. *Journal of Neurophysiology, 71*(5), 1959-1975.

Tian, B., & Rauschecker, J. P. (2004). Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. *Journal of Neurophysiology, 92*(5), 2993-3013.

Tomlinson, R., & Schwarz, D. W. (1988). Perception of the missing fundamental in nonhuman primates. *Journal of the Acoustical Society of America, 84*(2), 560-565.

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21*(1), 77-88.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to Western musical structure. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 394-402.

Trainor, L. J., & Trehub, S. E. (1993). Musical context effects in infants and adults: Key distance. *Journal of Experimental Psychology: Human Perception and Performance, 19*(3), 615-626.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics, 56*(2), 125-132.

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception, 20*(2), 187-194.

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. *Annals of the New York Academy of Science, 930*, 92-116.

Tramo, M. J., Shah, G. D., & Braida, L. D. (2002). Functional role of auditory cortex in frequency processing and pitch perception. *Journal of Neurophysiology, 87*(1), 122-139.

Trehub, S. (2000). Human processing predispositions and musical universals. In B. Merker & N. L. Wallin (Eds.), *The origins of music*. (pp. 427-448). Cambridge, MA, US: The MIT Press.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6*(7), 669-673.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development, 55*(3), 821-830.

Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology, 51*(4), 385-396.

Trehub, S. E., Schellenberg, E., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 965-975.

Trehub, S. E., & Thorpe, L. A. (1989). Infants' perception of rhythm: Categorization of auditory sequences by temporal structure. *Canadian Journal of Psychology, 43*(2), 217-229.

Trehub, S. E., Thorpe, L. A., & Trainor, L. J. (1990). Infants' perception of good and bad melodies. *Psychomusicology, 9*(1), 5-19.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behavior and Development, 16*(2), 193-211.

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music, 20*(1), 15-28.

van Egmond, R., Povel, D. J., & Maris, E. (1996). The influence of height and key on the perceptual similarity of transposed melodies. *Perception and Psychophysics, 58*(8), 1252-1259.

Vos, P., & Troost, J. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception, 6*, 383-396.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the priveleged status of speech for young infants. *Developmental Science, 7*(3), 270-276.

Watanabe, S., & Nemoto, M. (1998). Reinforcing property of music in Java sparrows (Padda oryzivora). *Behavioural Processes, 43*(2), 211-218.

Watanabe, S., & Sato, K. (1999). Discriminative stimulus properties of music in Java sparrows. *Behavioural Processes, 47*(1), 53-57.

Weary, D. M., Weisman, R. G., Lemon, R. E., & Chin, T. (1991). Use of the relative frequency of notes by veeries in song recognition and production. *Auk, 108*(4), 977-981.

Wegener, J. G. (1964). Auditory discrimination behavior of normal monkeys. *Journal of Auditory Research, 4*(2), 81-106.

Weisman, R., Ratcliffe, L., Johnsrude, I. S., & Hurly, T. A. (1990). Absolute and relative pitch production of the black-capped chickadee. *Condor, 92*, 118-124.

Wetzel, W., Ohl, F. W., Wagner, T., & Scheich, H. (1998). Right auditory cortex lesion in Mongolian gerbils impairs discrimination of rising and falling frequency-modulated tones. *Neuroscience Letters, 252*(2), 115-118.

Wetzel, W., Wagner, T., Ohl, F. W., & Scheich, H. (1998). Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbils. *Behavioural Brain Research, 91*(1-2), 29-39.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129*(3), 291-307.

Yin, S. (2002). A new perspective on barking in dogs (Canis familiaris). *Journal of Comparative Psychology, 116*(2), 189-193.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1-27.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia, 23*(1), 31-41.

Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of the Acoustical Society of America, 84*(2), 566-572.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function. *Nature Neuroscience, 6*(7), 692-695.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6*(1), 37-46.

Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature, 383*(6595), 29.

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development, 21*(3), 483-492.

Zhang, J., Harbottle, G., Wang, C., & Kong, Z. (1999). Oldest playable musical instruments found at Jiahu early Neolithic site in China. *Nature, 401*, 366-368.

Zhang, L. I., Bao, S., & Merzenich, M. M. (2001). Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nature Neuroscience, 4*, 1123-1130.

Chapter 2:

Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate

Josh McDermott
Department of Brain and Cognitive Sciences, MIT

Marc Hauser
Department of Psychology, Harvard University

Cognition, 94, B11-B21 (2004)

Abstract

Humans find some sounds more pleasing than others; such preferences may underlie our enjoyment of music. To gain insight into the evolutionary origins of such preferences, we explored whether they are present in other animals. We designed a novel method to measure the spontaneous sound preferences of cotton-top tamarins, a species that has been extensively tested for other perceptual abilities. Animals were placed in a V-shaped maze, and their position within the maze controlled their auditory environment. One sound was played when they were in one branch of the maze, and a different sound for the opposite branch; no food was delivered during testing. We used the proportion of time spent in each branch as a measure of preference. The first two experiments were designed as tests of our method. In Experiment 1, we used loud and soft white noise as stimuli; all animals spent most of their time on the side with soft noise. In Experiment 2, tamarins spent more time on the side playing species-specific feeding chirps than on the side playing species-specific distress calls. Together, these two experiments suggest that the method is effective, providing a spontaneous measure of preference. In Experiment 3, however, subjects showed no preference for consonant over dissonant intervals. Finally, tamarins showed no preference in Experiment 4 for a screeching sound (comparable to fingernails on a blackboard) over amplitude-matched white noise. In contrast, humans showed clear preferences for the consonant intervals of Experiment 3 and the white noise of Experiment 4 using the same stimuli and a similar method. We conclude that tamarins' preferences differ qualitatively from those of humans. The preferences that support our capacity for music may, therefore, be unique among the primates, and could be music-specific adaptations.

1. Introduction

Music is among the defining features of human culture, playing a central role in every society known to Western scholars. However, from the standpoint of evolution, music is also one of the most mysterious of human behaviors, as it serves no obvious function that might have driven its evolution. Evolutionary theorists since the time of Darwin have speculated about the adaptive function of music and its evolutionary origins (Darwin, 1871), with little consensus or empirical support. Recently, however, work on infants and animals (reviewed in Trehub, 2003; Hauser and McDermott, 2003) has begun to illustrate how empirical evidence might shape theories of music's evolution. In particular, because animals can be tested in the absence of any exposure to music, parallel perceptual abilities in nonhuman animals can help establish whether aspects of our music faculty are innate and therefore candidate products of natural selection. Moreover, as nonhuman animals do not themselves make music, any perceptual effect found in a nonhuman animal cannot be part of an adaptation for music. Music-related experiments on animals are thus poised to play an important role in the debate about the origins of music.

One of the striking and mysterious features of how we experience music and other forms of art is the aesthetic response we often have to what we experience. As is the case for most aspects of music, the function and origins of aesthetic responses are unclear. As a first step in investigating these issues, we studied preferences for relatively simple sounds that lack the complex temporal structure of extended passages of music. Perhaps the best-known example of such acoustic preferences involves harmonic musical intervals. Some combinations of notes tend to sound good, at least to Western listeners, and are termed consonant; others sound bad and are termed dissonant (Malmberg, 1918; Terhardt, 1984; Dowling and Harwood, 1986; Krumhansl, 1990). Pythagoras was the first to note that consonance tends to be generated by pairs of tones whose fundamental frequencies are related by simple integer ratios. Helmholtz later proposed the widely accepted notion that peripheral auditory effects (namely, beating) distinguish consonance and dissonance, and both neurophysiological (Fishman et al., 2001; Tramo et al., 2001) and behavioral (Izumi, 2000; Hulse et al., 1995) studies in birds, cats, and primates suggest that these peripheral differences are shared across mammals and birds. These peripheral effects account for the discriminability of consonant and dissonant intervals, but shed little light on the preferences between the two classes of stimuli that are arguably the main reason for their importance in music. Where do such preferences come from? Are they acquired through exposure to music, which perhaps contains more consonant intervals than dissonant ones? Are they part of an adaptation to music? Or might they be a byproduct of some general feature of the auditory system? Experiments in closely related animals, especially primates, can help to clarify these issues.

2. Experiment 1: Intensity

To test whether humans share any acoustic preferences with other primates, we first developed a method to measure such preferences in a well-studied nonhuman primate: the cotton-top tamarin (see Watanabe and Nemoto (1998) for a related method developed for use in birds). In Experiment 1 we compared a low amplitude white noise signal to a high amplitude white noise signal. We expected the animals to find the high amplitude signal unpleasant.

2.1 Method

2.1.1 Participants

We tested 6 adult cotton-top tamarins (*Saguinus oedipus*), 3 males and 3 females.

2.1.2 Apparatus and procedure

We tested subjects in a V-shaped maze (Figure 1). A concealed speaker (Advent Powered Partners) was located at the end of each branch of the maze. Each speaker and branch was paired with a different sound. Subjects were initially moved from their home room cage to the test room, and then placed at the entrance to the maze. The experimenter then left the room and raised the door to the maze by means of a pulley system, thereby allowing the tamarin to enter. When the subject moved into a branch for the first time, an experimenter started the playback. The stimulus for a particular side played continuously as long as the animal was on that side, and switched as soon as they switched sides. The animal's position in the maze thus determined which sound they heard. Testing continued for 5 minutes, during which subjects were videotaped. No food reward was given.

2.1.3 Stimuli and design

The amplitudes of the two white noise signals were 60 and 90 dB respectively when measured with a sound meter at the center point of the maze. The stimuli were randomly assigned to the sides of the apparatus for each animal. After two sessions of this condition, each separated by a full day, the sound-side pairing was reversed, and the animals were run for two more sessions, again on separate days. A bias to spend more time in one branch than the other was taken as evidence for a preference for one sound over the other.

The experimenters and trained assistants coded the video recordings with the sound turned off and without knowledge of the side assignment. The video displayed the time of recording down to a second's resolution. To code an experimental session the coder noted each time at which the animal moved from one side of the apparatus to the other. From these times the length of each excursion to one side or the other could be computed, and these were then added to yield the total amount of time spent on each side during an experimental session. Inter-observer reliability was high; over ten sessions coded by two observers, the correlation coefficient for the switch times noted by two different coders was .99.

2.2 Results & Discussion

Figure 2 plots the time spent on each side of the maze, averaged across the 6 monkeys, in each of 4 successive experimental sessions. Even in the first session there is a pronounced tendency to spend more time on the side playing the low amplitude white noise, a tendency that increases during the second session. After the first two sessions the sound-side assignments were swapped for each animal, and on average the animals spent equal amounts of time on each side, suggesting that they had learned an association between one side and the low amplitude noise. By the next session they regained the tendency to spend more time on the side with lower amplitude noise. Across all four sessions the animals averaged 70% of the time on the soft side, which was highly significant (t[23]=5.5, p<.00001). In a second experiment we modified the noise amplitudes so that there was only a 10 dB difference between the two sides (75 and 85 dB respectively). All six animals again spent more time on the side with the lower amplitude noise (68%, SE = 4.5%) over 2 sessions (t[11]=4.09, p<.001).

These results, together with those of Watanabe and Nemoto (1998) suggest that our method provides one way to assess spontaneous acoustic preferences in animals, especially for stimuli other than their species-specific vocalizations (for a related technique used to study such vocalizations, see Ryan, 1980; Gerhardt, 1987; Wilczynski *et al.*, 1995; Miller *et al.*, 2001).

3. Experiment 2: Distress Calls

To provide a second verification of our method and extend its ecological validity, we ran the tamarins on an experiment contrasting two species-specific vocalizations - screams given during distress and chirps made during feeding. Given the negative associations of the screams and the positive associations of the food-related chirps, we predicted that the tamarins would spend more time on the side with food chirps than on the side with distress screams.

3.1 Method

3.1.1 Participants

We tested 5 of the 6 cotton-top tamarins used in Experiment 1.

3.1.2 Apparatus and procedure

The apparatus and procedure of Experiment 1 were used again.

3.1.3 Stimuli and design

The distress calls were screams produced by animals being held by our veterinary staff during routine checkups. The food chirps were produced by individuals while eating food or just as food was presented. Files were created with six exemplars of either the screams or the chirps separated by brief periods of silence (a variable period between 1 and 1.5 seconds). The vocalizations were recorded from three different animals who were not run in the experiment. Two screams and two chirps from each of the three animals were used for the stimuli. The six screams or chirps looped continuously during playback.

This experiment was run approximately 3 months after the conclusion of Experiments 1, 2 and 4. Each animal was run in 3-4 sessions with a particular side assignment (determined at random), followed by 2-4 sessions with the side assignment reversed. The video recordings were coded as in Experiment 1.

3.2 Results & Discussion

Over the course of several sessions (N = 41 total across the 5 animals), subjects showed a statistically significant (t[40]=2.53; p<.01; Figure 3) preference for the side with food chirps over the side with screams.

This result provides further evidence that our method is appropriate for assaying spontaneous sound preferences in tamarins, and shows that these animals can have preferences for a range of stimuli, be they species-specific or artificial.

4. Experiment 3: Consonance

As our primary interests are centered on the origins of musical preferences, we began by testing tamarins for preferences for consonant stimuli over dissonant stimuli. Although humans, at least in Western cultures, tend to show a preference for consonant sounds, we ran adult humans on an analogue of the tamarin experiment to ascertain whether the method would translate to another species known to show the preference.

4.1 Method

4.1.1 Participants

The participants were the 5 cotton-top tamarins used in Experiment 2, along with 5 Harvard undergraduates (18-21 years old; 1 male, 4 female). All 5 human subjects had some degree of musical training, ranging from one to many years of music lessons.

4.1.2 Apparatus and procedure

The apparatus and procedure of Experiment 1 were used again with the tamarin subjects. The human subjects were placed in a room divided in half by a stripe taped to the floor. The front wall of the room concealed two speakers, one on each side of the dividing line. Each speaker played a particular sound when the subject was in the corresponding half of the room, thereby mimicking the tamarin setup.

The human subjects were told only that they had to stay within the confines of the room for the designated period of 5 minutes. No other instructions were given. All the human subjects were naïve as to the purpose of the experiment.

4.1.3 Stimuli and design

The consonant stimulus consisted of a sequence of two-note chords chosen randomly from the set of the octave, the fifth, and the fourth, subject to the constraint that no particular interval repeated more than once in a row. The dissonant stimulus was a similarly constructed sequence of minor seconds, tritones, and minor ninths. Each note composing the interval was a synthesized complex tone with ten harmonics. The bass note of each interval was middle C. Each interval was 1.5 seconds in duration and was ramped on and off over 100 msec. There was no gap between successive intervals, and the sequence of intervals played continuously as long as the subject was on the corresponding side, switching when they switched sides. The consonant and dissonant stimuli had equal amplitudes, which were set such that the sound level measured at the center of the apparatus/room was 80 dB.

As in Experiment 1, subjects were left in the apparatus for 5 minutes, during which they were free to move within its confines. The human subjects were run in a single session, while the tamarins were run repeatedly, up to 10 sessions in a row in some cases, as we wanted to maximize the chances of revealing an effect. The video recordings were coded as in Experiment 1.

4.2 Results & Discussion

One of the human subjects stood in the same place for the entire experiment, and his data were thrown out. The average results for the other four human subjects are plotted in Figure 4a. Human subjects spent most of their time on the consonant side of the room (t[3]=10.26; p<.001); this pattern was consistent across subjects (Mann Whitney test, U = 2.31, p < 0.02). In contrast, the tamarins showed no preference, spending approximately equal amounts of time on each side of the maze (Figure 4b; t[30]=0.47; p=0.32). The lack of preference is not due to habituation to our test apparatus, as all five animals showed robust preferences for low over high amplitude white noise when tested again at the conclusion of the experiment (1 session per animal, mean of 70% of time spent on low amplitude side). We conclude that under these particular test conditions, tamarins do not show a spontaneous preference for consonance over dissonance, differing notably from human adults tested with a similar paradigm.

5. Experiment 4: Screeching

As a second test of whether tamarins might have acoustic preferences based on something other than amplitude or behavioral relevance, we attempted to generate two nonmusical stimuli with similar amplitudes that were expected to produce a large preference in humans. We began by generating a stimulus that is highly aversive to most humans – the sound of fingernails on a blackboard (Halpern et al., 1986). The relationship between the responses that humans have to this stimulus and to musical stimuli is unclear, but it seemed conceivable that nonhuman animals might respond aversively to such a stimulus despite the lack of preference for consonance over dissonance.

5.1 Method

5.1.1 Participants

We tested 5 new adult cotton-top tamarins and 4 of the 5 Harvard undergraduates used in Experiment 3.

5.1.2 Apparatus and procedure

The apparatus and procedure of Experiment 3 were used again.

5.1.3 Stimuli and design

To facilitate stimulus generation, we used a variant of the fingernails-on-a-blackboard sound produced by scraping a three-pronged metal garden tool down a pane of glass (Halpern et al., 1986). Informal tests showed that stimuli produced the desired response in humans, suggesting it would produce a pronounced preference for a suitable comparison stimulus. The acoustic structure of our screeches was similar to that previously reported; there were typically several prominent harmonics overlaid with broadband noise. The experimental stimulus consisted of several concatenated recordings of individual screeches. As a comparison stimulus we generated white noise with the amplitude envelope of the screech stimulus. Both files looped continuously as long as a subject remained on the corresponding side of the apparatus. Subjects were again left to move freely in the apparatus for 5 minutes, during which they were videotaped. The amplitude of both stimuli was set to 80 dB as measured at the center point of the apparatus. The video recordings were coded as in Experiment 1.

5.2 Results & Discussion

Figure 5a shows the proportion of time humans spent on each side of the test room. As expected, there was a pronounced preference for the white noise ($t[3]=2.94$, $p<.05$); a Mann-Whitney test revealed that all subjects followed this pattern ($U = 2.31$, $p < 0.02$). Because the tamarins used in this experiment had not been run in Experiments 1 and 2, we first ran all 5 tamarins in a replication of Experiment 1. All of the tamarins spent more time on the side of the maze with the low amplitude noise, and this tendency reversed itself when the side assignments were reversed, as expected (66% of time on soft side; SE=2.17%; $t[37]=7.24$; $p<.00001$). When tested on the screech and control stimuli, however, the tamarins showed no evidence of a preference. We ran the tamarins for several consecutive sessions (N = 37 sessions) to see if a preference would emerge over time. As shown in Figure 5b, there was no preference ($t[36]=0.89$; $p=0.15$). In contrast with humans, who show a pronounced preference for white noise over the screeching sound, tamarins do not exhibit a preference.

6. Conclusions

Preferences for consonance over dissonance are widespread in human adults (Dowling and Harwood, 1986) and have also been demonstrated in human infants (Zentner and Kagan, 1996; 1998; Trainor and Heinmiller, 1998). Our results suggest that although such preferences may be innate in humans, they likely have evolved after the divergence point with our primate cousins. It is of course possible that another primate species, more closely related to humans (e.g., chimpanzees), might exhibit more similar acoustic preferences, or that tamarins tested with a different procedure would show a preference. It is also worth noting that Watanabe and Nemoto (1998) recently found that certain Java sparrows showed preferences for some types of music over others. This preference could conceivably be related to the singing behavior of this species, and it would be interesting to test them with the consonant and dissonant stimuli that we used. Given the present results, however, we conclude that if humans and nonhuman primates share acoustic preferences for sounds, this capacity evolved more recently than the divergence with New World monkeys such as the cotton-top tamarin (i.e., some 40 million years ago). This conclusion stands in contrast to the many perceptual mechanisms shared between humans and tamarins (and other species as well), particularly with respect to speech perception (Ramus et al., 2000; Miller et al., 2001, Newport et al., in press) and presumably also to the discriminability of consonance and dissonance. This contrast raises the possibility that some of the acoustic preferences observed in humans evolved as a specific adaptation for music.

Acknowledgements

We thank Altay Guvench and Matt Kamen for help building the apparatus, and Altay Guvench, Matt Kamen, Fernando Vera, Adam Pearson, Tory Wobber, Matthew Sussman, and Alex Rosati for help running the experiments. MDH acknowledges the NIMH and the McDonnell foundation for funding this project.

References

Dowling, J. and Harwood, D.L. *Music Cognition.* San Diego: Academic Press, 1986.

Fishman, Y.I. *et al.* Consonance and dissonance of musical chords: neuronal in auditory cortex of monkeys and humans. *J. Neurophysiol.* 86, 271-278 (2001).

Gerhardt, H.C. (1987). Evolutionary and neurobiological implications of selective phonotaxis in the green treefrog (*Hyla cinerea*). — Anim. Behav. 35, p. 1479-1489.

Halpern, D.L., Blake, R., & Hillenbrand, J. (1986). Psychoacoustics of a chilling sound. *Perception & Psychophysics,* 39: 77-80.

Hauser, M.D. & McDermott, J. The evolution of the music faculty: comparative perspectives. *Nat. Neurosci.* 6, 663-668 (2003).

Hulse, S.H., Bernard, D.J. & Braaten, R.F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (sturnus vulgaris). *Journal of Experimental Psychology: General.* 124, 409-423.

Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America,* 108, 3073-3078.

Krumhansl, C.L. *Cognitive Foundations of Musical Pitch* (Oxford Univ. Press, New York, 1990).

Malmberg, C.F. (1918) The perception of consonance and dissonance. *Psychological Monographs,* 25: 93-133.

Miller, C.T.; Miller, J.;Gil de Costa, R.; Hauser, M.D. (2001) Selective phonotaxis by cotton-top tamarins. *Behaviour,* 138, 811-826.

Miller, C.T., Dibble, E.; Hauser, M.D. (2001) Amodal completion of acoustic signals by a nonhuman primate. *Nature Neuroscience,* 4, 783-784.

Newport E.L., Hauser, M.D., Spaepen, G., & Aslin, R.N. (in press) Learning at a distance: II. Statistical learning of non-adjacent dependencies in a nonhuman primate. *Cognitive Psychology.*

Ramus, F.; Hauser, M.D.; Miller, C.T. ; Morris, D. Mehler, J. (2000) Language discrimination by human newborns and cotton-top tamarin monkeys. *Science,* 288, 349-351.

Ryan, M.J. (1980). Female mate choice in a Neotropical frog. *Science* 209, p. 523-525.

Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. Music Perception 1, 276-295.

Trainor, L.J. & Heinmiller, B.M. (1998). The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, 21, 77-88.

Tramo MJ, Cariani PA, Delgutte B, Braida LD. (2001). Neurobiological foundations for the theory of harmony in western tonal music. *Ann N Y Acad Sci.* 930:92-116.

Trehub, S.E. (2003) The developmental origins of musicality. *Nature Neuroscience* 6, 669-673.

Watanabe, S., Nemoto, M. (1998). Reinforcing properties of music in Java sparrows. *Behavioral Processes* 43, 211-218.

Wilczynski, W., Rand, A.S. & Ryan, M.J. (1995). The processing of spectral cues by the call analysis system of the tungara frog, *Physalaemus pustulosus. — Animal Behavior* 49, 911-929.

Zentner, M. R. & Kagan, J. (1996). Perception of music by infants. *Nature* 383, 29.

Zentner, M.R. & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development*, 21, 483-492.

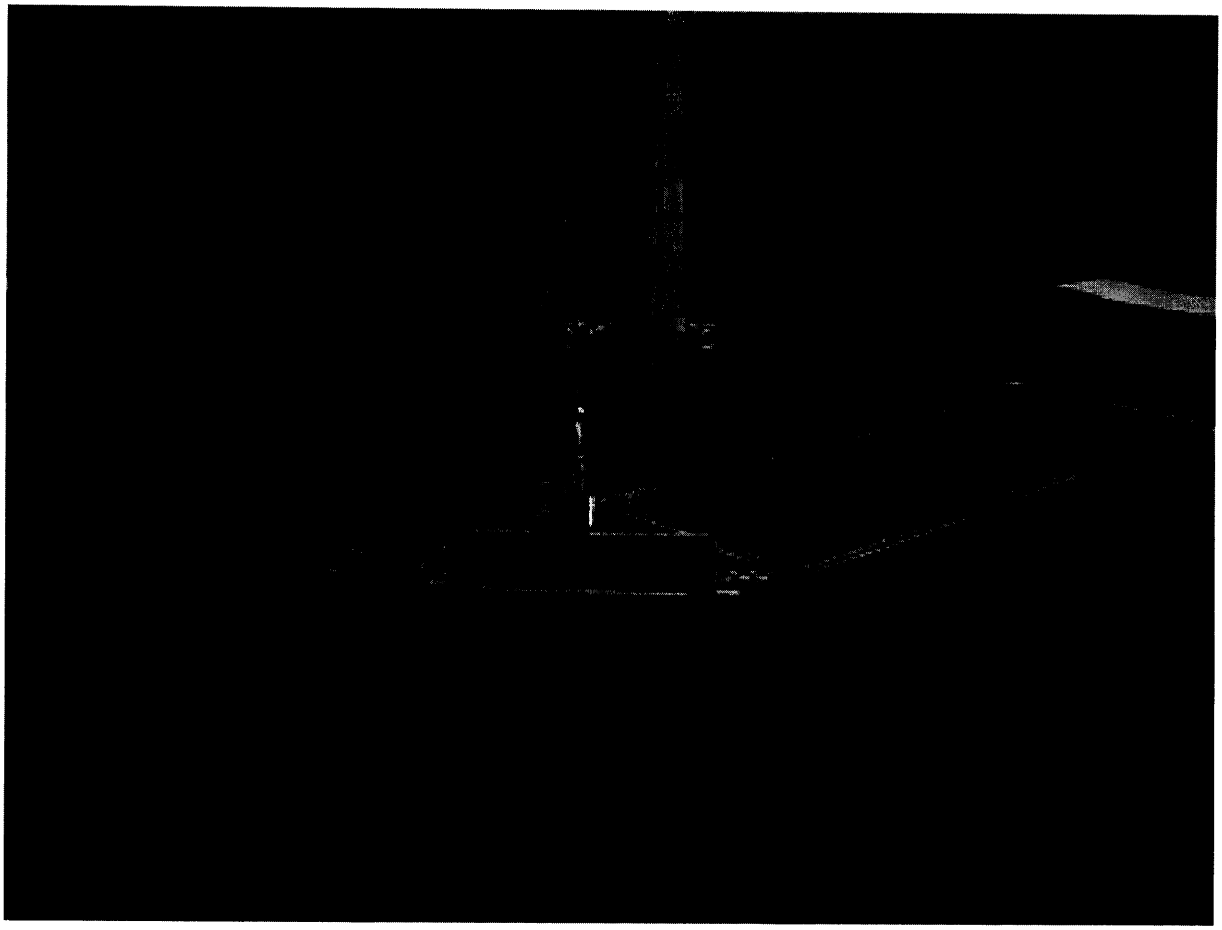

Figure 1. Photo of the apparatus used in the tamarin experiments. The maze was elevated off the floor. There was a concealed speaker at the end of each branch of the maze.

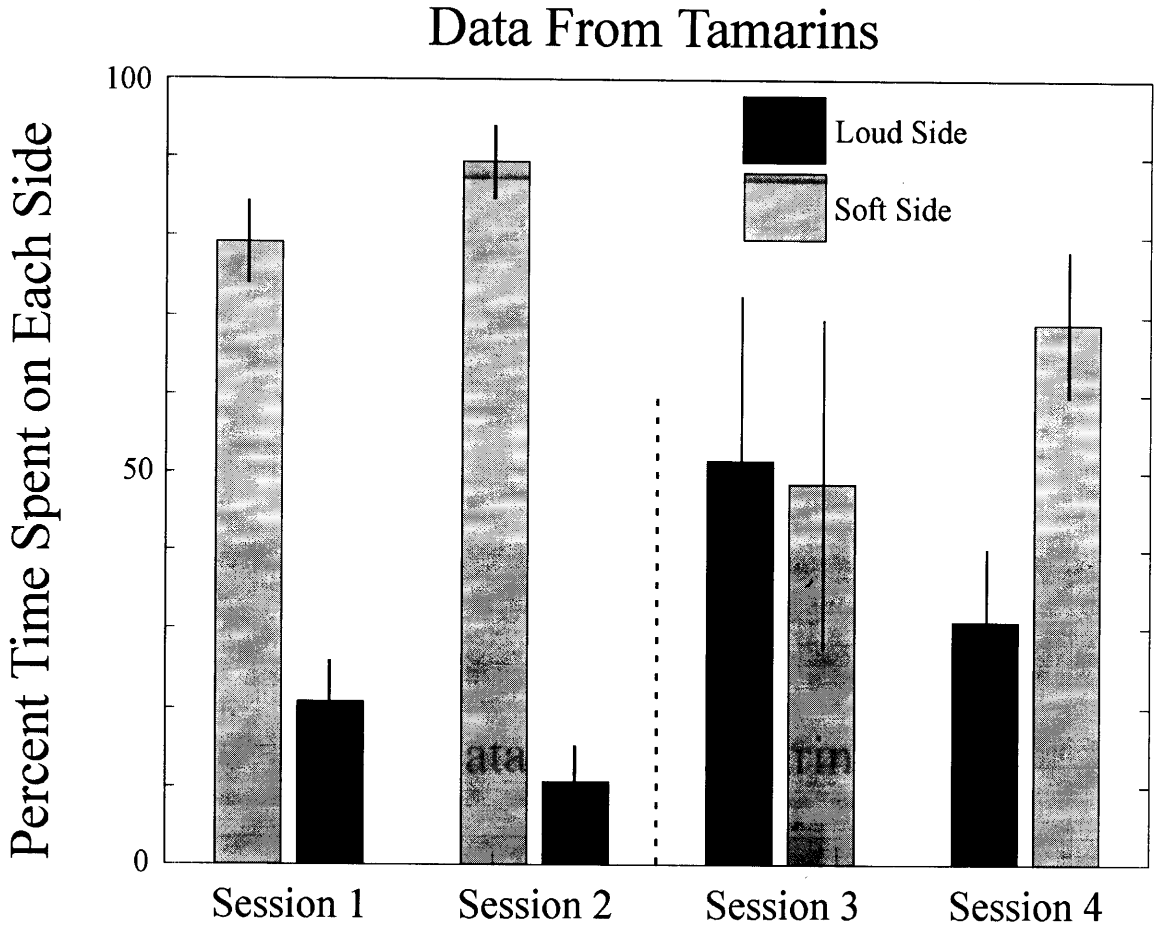

Figure 2. Results of Experiment 1, comparing high and low amplitude white noise. Each bar plots the average data from 6 subjects, as a proportion of the total time spent in the apparatus. Error bars here and elsewhere denote standard errors. The dashed line denotes reversal of the side assignment that occurred after the second session.

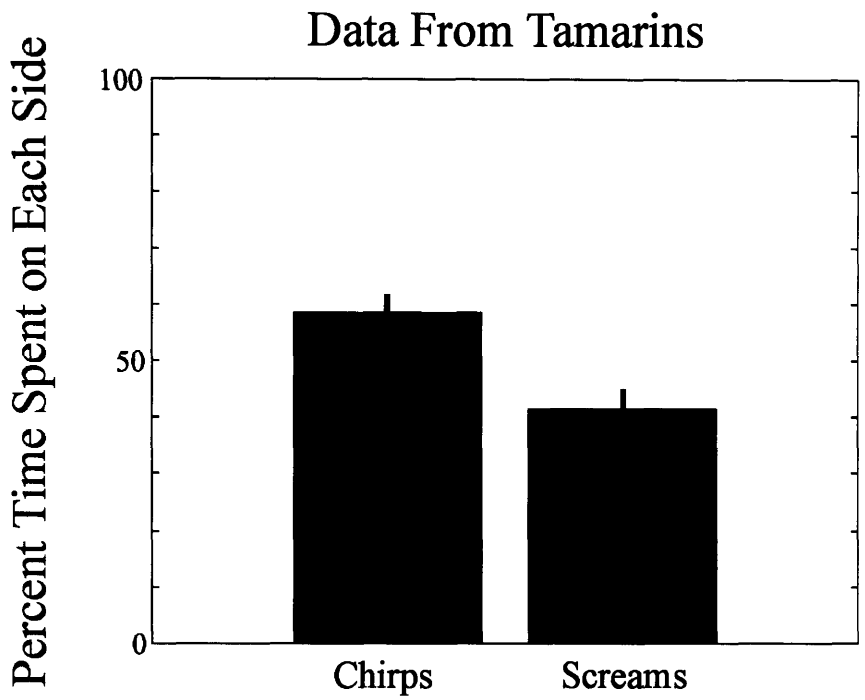

Figure 3. Results from Experiment 2, comparing food chirps with distress screams.

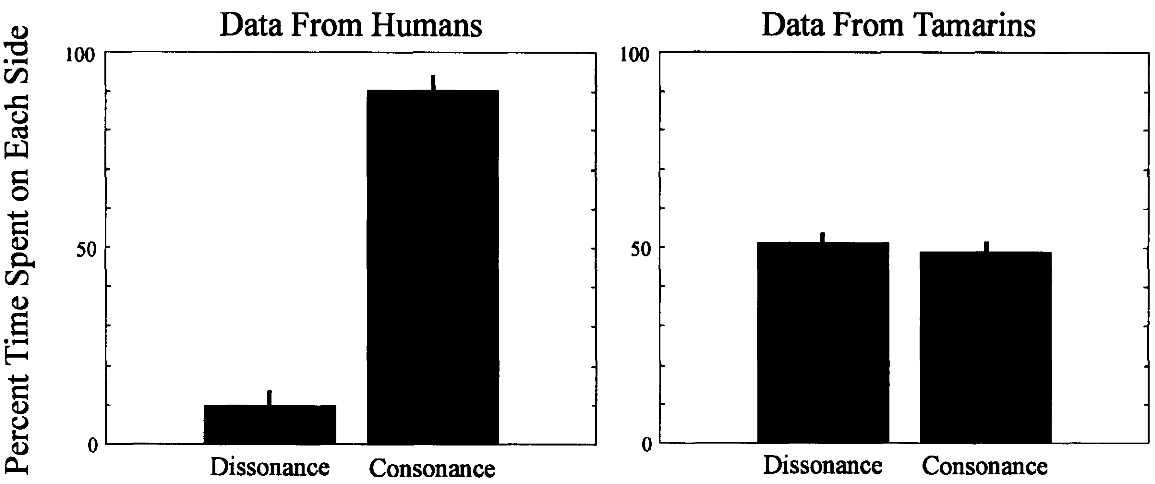

Figure 4. Results of Experiment 3, comparing consonant and dissonant musical intervals. (a) Results for human subjects. (b) Results for tamarin subjects.

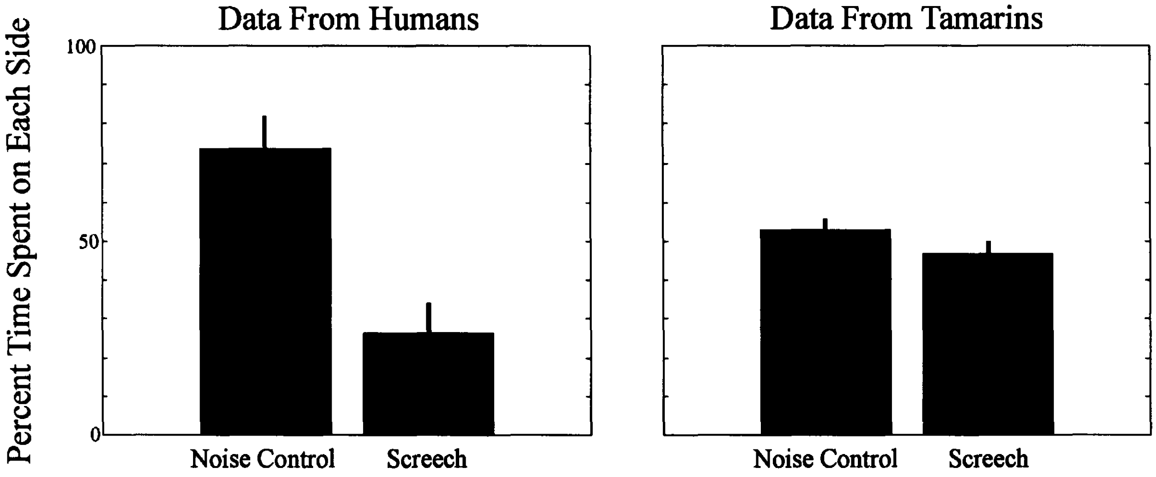

Figure 5. Results of Experiment 4, comparing a screeching sound with an amplitude matched noise control stimulus. (a) Results for human subjects. (b) Results for tamarin subjects.

Chapter 3:

Nonhuman primates prefer slow tempos but dislike music overall

Josh McDermott
Department of Brain and Cognitive Sciences, MIT

Marc Hauser
Departments of Psychology, Organismic & Evolutionary Biology, and
Biological Anthropology, Harvard University

To appear in *Cognition*

Abstract

Human adults generally find fast tempos more arousing than slow tempos, with tempo frequently manipulated in music to alter tension and emotion. We used a previously published method (McDermott & Hauser, 2004, Cognition) to test cotton-top tamarins and common marmosets, two new-World primates, for their spontaneous responses to stimuli that varied systematically with respect to tempo. Across several experiments, we found that both tamarins and marmosets preferred slow tempos to fast. It is possible that the observed preferences were due to arousal, and that this effect is homologous to the human response to tempo. In other respects, however, these two monkey species showed striking differences compared to humans. Specifically, when presented with a choice between slow tempo musical stimuli, including lullabies, and silence, tamarins and marmosets preferred silence whereas humans, when similarly tested, preferred music. Thus despite the possibility of homologous mechanisms for tempo perception in human and nonhuman primates, there appear to be motivational ties to music that are uniquely human.

keywords: Music; Evolution; Monkeys; Tempo; Preference

Introduction

The origins of music are a long standing puzzle (Darwin, 1871). Empirical work in developmental (Trainor & Heinmiller, 1998; Trehub, 2003; Zentner & Kagan, 1996) and comparative (D'Amato, 1988; Hulse et al., 1984; Izumi, 2000; Trainor & Heinmiller, 1998; Sandra E. Trehub, 2003; Watanabe & Nemoto, 1998; Watanabe & Sato, 1999; Wright et al., 2000; Zentner & Kagan, 1996) psychology can, however, constrain evolutionary theories by testing whether musically relevant traits observed among humans are innate and uniquely human (McDermott & Hauser, 2005). Traits that are innate and unique to humans are candidates for music-related adaptations; those that are shared by nonhuman animals are more likely to have evolved for other, domain-general functions.

To probe the uniqueness of our aesthetic response to music, we introduced a method for measuring spontaneous acoustic preferences in nonhuman animals (McDermott & Hauser, 2004); Watanabe and Nemoto (1998) had previously used a similar method to ask related questions in songbirds. We used this method to test cotton-top tamarins for several preferences that are pronounced and widespread in humans. The method revealed expected preferences in tamarins for soft over loud white noise, and for positive- over negatively-valenced conspecific vocalizations, suggesting its appropriateness as a measure of preference. When subsequently tested for other acoustic preferences found in humans, however, tamarins were indifferent. Specifically, tamarins failed to demonstrate a preference for consonant over dissonant musical intervals. In contrast, human adults showed the expected preference for consonance when measured with an analogous method. We concluded from these results that some of the essential preferences that underlie music perception in humans appear to be missing from some species of nonhuman primates, raising the possibility that such preferences are unique to our species.

In this paper we use the same method to probe nonhuman primate responses to tempo. Tempo, which we operationally define as the rate of acoustic events in a stimulus, is routinely manipulated in music to induce and resolve tension (Krumhansl, 1996) and to modulate affect (Hevner, 1937). Fast tempos are more arousing than slow, all other things being equal (Balch & Lewis, 1999; Husain, Thompson, & Schellenberg, 2002). So far as we know this is true across cultures, suggesting the effect of tempo may have an innate basis. Lullabies, for instance, tend to have slow tempos irrespective of their country of origin (Unyk et al., 1992), presumably in part because the slow tempo causes them to have a pacifying effect on their target listeners. This innate predisposition could be uniquely human, perhaps part of an adaptation for music. It is also possible that the predisposition is not specific to music, and could conceivably have been inherited from our nonhuman ancestors. To help distinguish these possibilities, we conducted experiments on cotton-top tamarins, as well as a close relative, the common marmoset.

1. General Methods

1.1 Apparatus and Procedure

The apparatus [a V-shaped maze, Fig. 1] and procedure were identical to that of McDermott and Hauser (2004). We placed concealed speakers (PolkAudio Atrium 45p) at the end of each branch of the maze; each speaker and branch was paired with a different sound. Subjects were initially placed at the entrance to the maze. The experimenter then left the room and raised the door to the maze by means of a pulley, allowing the animal to enter. We placed small pieces of food in both branches of the maze, equidistant from the entrance, to entice them to enter; the animals always ate both pieces of food. When the subject moved into a branch for the first time, the experimenter started the playback. The stimulus for a particular side played continuously as long as the animal was on that side, and switched as soon as it switched sides. The animal's position in the maze thus determined which sound it heard. Testing continued for 5 minutes. After a certain number of sessions with one sound on the left side and another on the right, generally separated by a full day, the sound-side pairing was reversed, and the animals were run for the same number of additional sessions. Data before and after the reversal were combined to distinguish stimulus preferences from side biases. The number of sessions for a single animal in a single experiment was usually 6 or 8. It was always fixed in advance of the experiment.

The computer program that turned the sounds on and off generated a list of the times at which one sound was started and the other was stopped. These times were almost identical to those extracted from video recordings of the sessions by blind observers ($r= .99$). The total time that a subject spent on each side was computed from these switch times. A statistically significant bias to spend more time in one branch than the other was taken as evidence for a preference for one sound over the other. Throughout, we test for statistical significance using two-tailed t-tests, with significance set at $p < 0.05$.

1.2 Method Validation

To validate the method, we first ran all participating animals in a control experiment using 60 and 90 dB white noise. If the method is appropriate for measuring preferences in our subjects, they should spend less time on the side playing loud white noise, which we presume to be relatively aversive. These experiments were described for the tamarin subjects in a previous publication (McDermott & Hauser, 2004); we repeated the same experiment in the six marmosets used in the present study. Marmosets averaged 63.3% (SE=1.58%) of the time on the soft side, which was statistically significant across the six animals ($t(41)=-7.62$, $p<0.0001$), and for each animal individually (individuals spent between 56.5% and 68.8% on the side of the soft noise).

1.3 Participants

Both tamarin and marmoset subjects were born and raised in the lab, and had no exposure to music prior to our studies. We attempted to run as many animals as possible in each experiment, but due to issues of pregnancy, health and other lab constraints, this number varied somewhat from experiment to experiment. In a few experiments, the marmosets were unavailable to run. As it turned out, however, the marmosets and tamarins exhibited similar results in every experiment in which they were both run. Every animal run in every experiment showed a significant preference for soft over loud noise, demonstrating the efficacy of the method in every subject we used. Care was taken to ensure that the total number of sessions was constant across experiments wherever possible.

1.4 Stimuli

Apart from the control experiments with white noise, the average intensity of each stimulus was always set at 62 dB. Pilot experiments at other stimulus levels suggested the specific amplitudes were not critical as long as they were matched across stimuli.

2. Experiment 1: Lullabies

We started with an extremely crude and coarse-grained contrast between two pieces of instrumental music: a Russian lullaby played on a flute, and an excerpt of German electronic techno. The two stimuli were intentionally chosen to be as different as possible, matching in average intensity while differing along almost every other dimension imaginable: spectral composition, attack rate, tempo etc. This large stimulus contrast was motivated primarily by our prior failure to find music-related preferences in tamarins (McDermott & Hauser, 2004). We wanted to see if any variables at all, apart from amplitude or behavioral relevance, might contribute to preferences in our nonhuman primate subjects. To rule out a possible role for differences in amplitude fluctuations for each stimulus, we filtered the techno excerpt to have the same amplitude envelope as the lullaby. To ensure that this amplitude modulation was not critical, we also ran a second experiment with the lullaby and the original techno excerpt as stimuli. To avoid any effects that human voices might have, we chose stimuli that were purely instrumental.

2.1 Method

In Experiment 1a we tested four adult cotton-top tamarins, 2 males and 2 females, as well as 4 common marmosets, 3 males and 1 female. In Experiment 1b we tested five adult cotton-top tamarins, one male in addition to the original 4; the marmosets were unavailable for this experiment.

The lullaby was a Russian folk song played on a flute. The techno was an excerpt from the track "Nobody Gets Out Alive" by Alec Empire (Digital Hardcore Recordings, 2000). In Experiment 1b, the techno was filtered so as to have the same amplitude envelope as the lullaby.

The number of total sessions (before and after the reversal) per subject was 8 for the tamarins in Experiment 1a, who were run first. As this effect was quite robust, we subsequently collected only 6 sessions per subject for the marmosets in Experiment 1a as well as for the tamarins in Experiment 1b. The number of sessions before and after the side reversal was always the same, as was the case in all the experiments.

2.2 Results and Discussion

Both tamarins and marmosets showed a robust preference for the lullaby over the techno (tamarins: $t(31)=4.47$, $p<0.0001$; marmosets: $t(23)=7.27$, $p<0.00001$; Fig. 2) in Experiment 1a. This preference was not an artifact of the amplitude modulation given to the techno, as it persisted in Experiment 1b with the original techno excerpt and the same lullaby stimulus (the animals spent an average of 62.79% (SE=1.7%) on the lullaby side; t(29)=6.45, p<0.00001).

The two stimuli were equally loud to our ears, but it seemed conceivable that spectral differences between the stimuli might produce a loudness difference for the monkeys, whose loudness perception is not well characterized. We therefore ran a second control experiment with the tamarins, using 61 and 64 dB white noise stimuli. The difference in loudness produced by this 3 dB level difference is small but clearly perceptible to humans, and the experiment thus served as a test of whether small differences in loudness could account for our results. The tamarins showed no evidence of a preference for the lower amplitude noise (t(31)= 0.92, p=.36; 8 sessions per animal for the four animals from Experiment 1a). Thus, it seems unlikely that a small loudness difference can account for the observed lullaby preference. Apparently some other difference between the two stimuli caused the monkeys to prefer one over the other.

3. Experiment 2: Click Trains

Any of the various differences between the stimuli of Experiment 1 could have conceivably caused the measured preference. One such difference was in the event rate, or tempo. We measured the average event rate of the lullaby and techno to be 65.26 and 369.23 beats per minute, respectively. To isolate the role of tempo, we used click trains, which were identical apart from the rate of repetition.

3.1 Method

We tested the 4 tamarins from Experiment 1a and 5 marmosets, four of whom were used in Experiment 1a. The clicks were impulses 0.12 ms in duration. In Experiment 2a, one stimulus contained 60 clicks per minute, evenly spaced, and the other 400 clicks per minute. These tempos were approximately matched to those of the stimuli of Experiment 1, and represent the approximate upper and lower limits of tempos found in music[1]. To ensure that the results were not limited to such vast tempo differences, Experiment 2b presented click trains at 80 and 160 clicks per minute to 5 tamarins (including the four from Experiment 1). In both experiments each animal was run for 8 total sessions.

---

[1] It is rare to hear of tempos in music above 160 beats per minute, but this has more to do with limits on how fast humans move to music than with limits on how fast events occur in music. Humans typically entrain to rhythms at relatively slow rates, usually to metrical pulses that occur at some fraction of the event rate. This rate of metrical pulses is usually cited as the tempo even though the actual event rate of the piece in question might be considerably faster than that. It is common for the rate of events in a musical stimulus to range far upwards of 180 per minute, e.g. in drum rolls that often signal musical climaxes or in percussive dance music. It is this event rate that we are concerned with here rather than the rate of metrical pulses.

3.2 Results and Discussion

In Experiment 2a, both tamarins (t(31)=3.54, p=.001) and marmosets (t(39)=3.33, p=.002) preferred 60 clicks per minute to 400 clicks per minute (Fig. 3). This pattern of results was replicated with the smaller tempo difference (80 vs. 160 clicks per minute) of Experiment 2b (t(39)=2.9212, p=.006). When forced to choose, both species preferred stimuli in which events occurred at a slow rate. This provides a parsimonious explanation for the lullaby preference of Experiment 1.

4. Experiment 3: Attack Velocity

The stimuli presented in Experiment 1 also differed in the average attack velocity of the sounds composing the stimuli. The notes of the lullaby were gently blown and thus their onset was gradual; in contrast, the techno featured many brief percussive sounds, whose onset was rapid. It seemed possible that the rapid onsets were arousing to the animals and thus avoided. To investigate the role of the attack velocity we conducted an experiment in which this variable was isolated. We presented subjects with a choice between a pure tone with either a square wave or a triangle wave amplitude envelope; these stimuli differed maximally in their attack velocities but were equated for acoustic energy.

4.1 Method

We tested 2 marmosets and 4 tamarins out of the pool used in Experiment 1a, with 8 sessions per animal. The stimuli were produced by applying a 1 Hz envelope to an 800 Hz pure tone. The two envelopes used were a square wave and a triangle wave of the same amplitude, whose integrals were thus the same.

4.2 Results and Discussion

To improve power, we pooled the results across tamarins and marmosets; both as a group and individually they showed no preference between the two stimuli (t(47)=.008, p=.994; individuals spent between 47.9% and 52.4% on the side of the square wave envelope). Attack velocity, at least in the way that we manipulated it, thus does not appear to be responsible for the lullaby preference observed in Experiment 1.

5. Experiment 4: Music vs. Silence

Given that both tamarins and marmosets seem to prefer slow over fast tempos, one question of interest is whether they “like” slow tempo stimuli. One way to address this question behaviorally is by presenting subjects with a choice between various pleasant musical stimuli and silence. For purposes of comparison we tested human adults on this contrast, in addition to both species of monkey.

5.1 Method

We tested 8 human adults (18-21 years old, 6 female, 2 male), 4 cotton-top tamarins, and 5 common marmosets.

As described previously (McDermott & Hauser, 2004), human subjects were placed in a room divided in half by a stripe taped to the floor. There was a speaker on each side of the room; each speaker played a particular sound when the subject was in the corresponding half of the room. The human subjects were told only that they had to stay within the confines of the room for the designated period of 5 minutes. After the experiment the human subjects were debriefed and were asked to rate how much they liked the stimulus on a scale of –3 (disliked) to 3 (liked).

For stimuli, we used the same lullaby as in Experiment 1, as well as a field recording of a lullaby sung in German (provided by Sandra Trehub), and a MIDI rendition of a Mozart string concerto (K458, in B flat major). After the first experiment with the flute lullaby (with 6 sessions per subject) yielded large effects, we ran the subsequent experiments with the sung lullaby and Mozart concerto with only 4 sessions per monkey subject, and with fewer subjects, as this provided adequate statistical power.

5.2 Results and Discussion

Human subjects consistently preferred the musical stimuli over silence as measured behaviorally (flute lullaby: $t(7)=2.67$, $p=.032$; sung lullaby: $t(7)=9.47$, $p<.0001$; Mozart concerto: $t(7)=3.36$, $p=.01$; Fig. 4). The ratings data were consistent with this behavioral effect, significantly greater than zero on average (flute lullaby: mean = 1.5, $t(7)=4.58$, $p=.0025$; sung lullaby: mean = 1.9, $t(7)=4.7$, $p=.002$; Mozart concerto: mean = 2.38, $t(7)=12.98$, $p<.00001$). In contrast, both tamarins and marmosets strongly and consistently preferred silence over the musical stimuli (flute lullaby - tamarins: $t(23)=-5.26$, $p<0.0001$; marmosets: $t(29)=-4.6$, $p<0.0001$; sung lullaby – tamarins: $t(11) = -3.93$, $p=0.003$, 3 animals; marmosets: $t(7) = -8.48$, $p= 0.0006$, 2 animals; Mozart concerto – tamarins: $t(15)=-5.33$; $p<.0001$, 4 animals; marmosets: $t(15)=-4.69$; $p=.0003$, 4 animals; Fig. 4). The monkey aversion to music was also evident if our analysis was restricted to data from the first session of each experiment, with animals spending more time with silence than with music in every case (n=22; the proportion of time spent on the music side ranged from 15.19% to 45.79%). Given this pattern of results, it seems unlikely that the monkeys find lullabies or other musical stimuli rewarding in the same way that people do.

6. Experiment 5: Energy Control

Given that the monkeys had relatively aversive responses to even the most pleasant of musical stimuli (at least as measured in our apparatus with a comparison to silence), several new issues emerged with respect to our tempo results in Experiments 1 and 2. For example, the fast click trains contained more acoustic energy per unit time than did the slow trains. The overall aversion to acoustic stimulation exhibited in Experiment 4 might, therefore, predict a preference for slow click trains over fast that could be unrelated to tempo. We therefore conducted an experiment with stimuli that varied in tempo but were equated in acoustic energy.

6.1 Method

We tested four cotton-top tamarins and one marmoset from Experiment 1a.

The stimuli were 800 Hz pure tones with a square wave envelope, as used in Experiment 3. The slow tempo stimulus had an envelope of 1Hz, producing 60 0.5 sec tone pulses per minute. The fast tempo stimulus had an envelope of 6.66 Hz, producing 400 0.075 sec tone pulses per minute. The stimuli thus had the same amount of acoustic energy per unit time, while differing in their tempos, which again were approximately matched to the tempos of the stimuli of Experiment 1.

Each subject was run for 4 sessions with one stimulus assignment followed by 4 sessions of the reversed assignment.

6.2 Results and Discussion

All animals spent more time on the slow side (Fig. 5: ranging from 54.78% to 64.02%); this effect was significant for the tamarins as a group (t(31)=3.54, p=.001) and for the marmoset individually (t(7) = 3.3; p=.01). The preference for slow tempos over fast persisted even when stimuli were equated for acoustic energy.

7. Experiment 6: Chirp Trains

One question of interest is whether the tempo effect evident in tamarins and marmosets bears any relation to the effect of tempo on humans. Fast tempos are more arousing than slow tempos to humans, and although humans do not always prefer slower tempo stimuli, a similar difference in arousal could underlie the effect we found in our nonhuman primate subjects.

Another less interesting explanation, however, is again suggested by the results of Experiment 4. Although Experiment 5 ruled out an explanation of our tempo effects in terms of differences in total acoustic energy, if the individual sounds composing our stimuli are themselves unpleasant to the animals, then slow tempo stimuli might be preferred simply because they have fewer aversive acoustic events. The preference for slow tempos over fast could then be merely an indication that the animals do not like the onsets of the sounds we chose for stimuli. To control for this possibility, we sought to generate stimuli composed of non-aversive sounds, using elements of tamarin and marmoset vocalizations.

7.1 Method

We tested 6 tamarins and 3 marmosets, all of whom had previously demonstrated a preference for soft over loud noise. We were able to generate vocalization-based stimuli that produced non-aversive responses in 5 of the 6 tamarins and 1 of the 3 marmosets, thereby yielding 6 animals suitable for further testing. All of these animals, apart from one of the tamarins, had been tested in the previous experiment with click trains, and had shown a preference for the slower tempo.

We started with vocalizations that each species often produces in non-threatening situations. Tamarins often emit "chirps" when presented with food (Cleveland & Snowdon, 1981); presentation of food elicits chirp-like calls in marmosets as well. Tamarins also produce a "combination long call" that functions for long distance communication, and this vocalization begins with a brief chirp. We took recordings of these two kinds of chirps from various members of the tamarin and marmoset colonies, and created trains of them at various repetition rates. We then ran a series of pilot experiments (with 6 sessions total per subject) in which we presented subjects with a choice between slow tempo chirp trains (50 or 60 chirps per minute) and silence, searching for stimuli that our subjects would at least prefer equally to silence. The monkeys for whom such stimuli were found were then run on an experiment with fast tempo chirp trains (250 per minute) vs. silence, and then again on slow chirp trains (50 per minute) vs. silence, counterbalanced for order, again with 6 sessions per subject. The higher chirp rate of 250 per minute was chosen to approximately match the fastest rates that we observed naturally in the animals. For both tests we recorded the number of times the animals switched from left to right in our apparatus, as it seemed that this measure of activity might be informative with respect to their arousal level.

The non-aversive slow chirp trains were composed of either the chirps from a cagemate's combination long calls (n=3 tamarins), the cagemate's food chirps (n = 1 marmoset), or the food chirps from an unrelated, long deceased tamarin (n = 2 tamarins). These six animals were then tested on fast and slow chirp trains composed of these chirps (the slow chirp trains were repeated to avoid selection bias and order effects). One of the six animals was not run on the full series of tests and her data are omitted from the group analysis. The data from the other 5 animals are combined.

7.2 Results and Discussion

None of the animals spent significantly different amounts of time with their customized slow chirp trains (50 or 60 chirps/min, depending on the animal) compared to silence, individually or as a group (t(29)=-0.18, p=.85). However, when the same chirps were presented at a faster rate of 250 per minute, the same animals consistently preferred silence over the chirps (t(29)=-5.99, p<0.00001; Fig. 6). When tested again on slow chirp trains composed of the same chirps, the same animals again spent equal amounts of time with the chirps as with silence (t(29)=-0.77, p=.45; Fig. 6). This effect is not well-accounted for via the cumulative effect of aversive sounds, as the chirps are non-aversive when played at slow rates. It also seems unlikely that the effect is due to the fast trains being less ecologically valid, as we have frequently observed the animals make rapid bursts of chirps at rates close to our fast presentation rate. Rather, it appears that both tamarins and marmosets find fast tempos aversive. We also observed a significant difference in the average number of times the animals moved from one side of the apparatus to the other – they averaged 21.86 side switches per session (SE=2.22) during the fast tempo sessions, and 27.56 (SE=2.17) for the slow tempo sessions (t(70)=1.84, p=.03). This is consistent with the hypothesis that the fast tempo stimuli are aversive, and that the animals are less likely to revisit the side with the aversive stimuli once they have been there.

8. Conclusions

In previous work we reported evidence that nonhuman primate spontaneous acoustic preferences differ dramatically from those of humans (McDermott & Hauser, 2004), and the experiments described in the present paper largely support this idea. Where humans choose to hear music over silence, monkeys choose the reverse, suggesting that they do not find such stimuli pleasurable or relaxing. However, despite the apparently aversive response monkeys have to many musical stimuli, we found for the first time evidence that they do have nontrivial preferences for some musical stimuli over others, and our results suggest that tempo is a critical variable.

Is the monkey response to tempo homologous to the human response to tempo? Further work is needed on this topic, but our results at least leave this possibility open. Humans obviously do not always prefer slow tempos to fast, but differences in temperament could cause tamarins and marmosets to find arousing stimuli aversive, whereas to humans they would merely be stimulating. Many stressful events in the natural environment, such as fights and storms, feature rapid sequences of acoustic events, and it is thus conceivable that animals have come to associate such stimuli with high levels of arousal. Future work using direct measures of arousal could provide further support for this hypothesis. It is also interesting to note that the alarm calls of tamarins and marmosets consist of short broadband bursts repeated at very high rates (Fig. 7 shows one such call from a tamarin). This acoustic structure is common to certain types of alarm calls in species ranging from monkeys and squirrels to birds (Marler, 1955), and could be related to the response nonhuman animals have to fast-tempo stimuli. Taken as a whole, however, the body of work on music perception in nonhuman primates suggests fundamental differences in the way they respond to musical stimuli compared to humans. Our preferences for certain kinds of music may reflect a unique evolutionary history of selection on perceptual mechanisms linked up with our emotional and motivational systems.

Acknowledgments

We thank Altay Guvench and Matt Kamen for help building the apparatus, Joe Presbrey for programming the sound presentation software, Anna Volkova for providing the flute lullaby stimulus, Sandra Trehub for providing the sung lullaby stimulus, and Tonja Machulla, Joy Liang, Cori McLean, Calvin Yeh, Meredith Brown, Nadja Oertelt, Sam Lasser, Adena Schachner, Chris Ludwick and Kara Furman for help running the experiments.

References

Balch, W. R., & Lewis, B. S. (1999). Music-dependent memory: The role of tempo changes and mood mediation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1354-1363.

Cleveland, J., & Snowdon, C. T. (1981). The complex vocal repertoire of the adult cotton-top tamarin, Saguinus oedipus oedipus. *Zeitschrift fur Tierpsychologie, 58*, 231-270.

D'Amato, M. R. (1988). A search for tonal pattern perception in cebus monkeys: Why monkeys can't hum a tune. *Music Perception, 5*(4), 453-480.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Hevner, K. (1937). The affective value of pitch and tempo in music. *American Journal of Psychology, 49*, 621-630.

Hulse, S. H., Cynx, J., & Humpal, J. (1984). Absolute and relative pitch discrimination in serial pitch perception by birds. *Journal of Experimental Psychology: General, 113*(1), 38-54.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities: Re-examination of the "Mozart effect". *Music Perception, 20*, 151-171.

Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America, 108*(6), 3073-3078.

Krumhansl, C. L. (1996). A perceptual analysis of Mozart's piano sonata K. 282: Segmentation, tension and musical ideas. *Music Perception, 13*(3), 401-432.

Marler, P. (1955). Characteristics of some animal calls. *Nature, 176*, 6-7.

McDermott, J., & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition, 94*(2), B11-B21.

McDermott, J., & Hauser, M. D. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23*(1), 29-59.

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21*(1), 77-88.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6*(7), 669-673.

Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music, 20*(1), 15-28.

Watanabe, S., & Nemoto, M. (1998). Reinforcing property of music in Java sparrows (Padda oryzivora). *Behavioural Processes, 43*(2), 211-218.

Watanabe, S., & Sato, K. (1999). Discriminative stimulus properties of music in Java sparrows. *Behavioural Processes, 47*(1), 53-57.

Wright, A. A., Rivera, J. J., Hulse, S. H., Shyan, M., & Neiworth, J. J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129*(3), 291-307.

Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature, 383*(6595), 29.

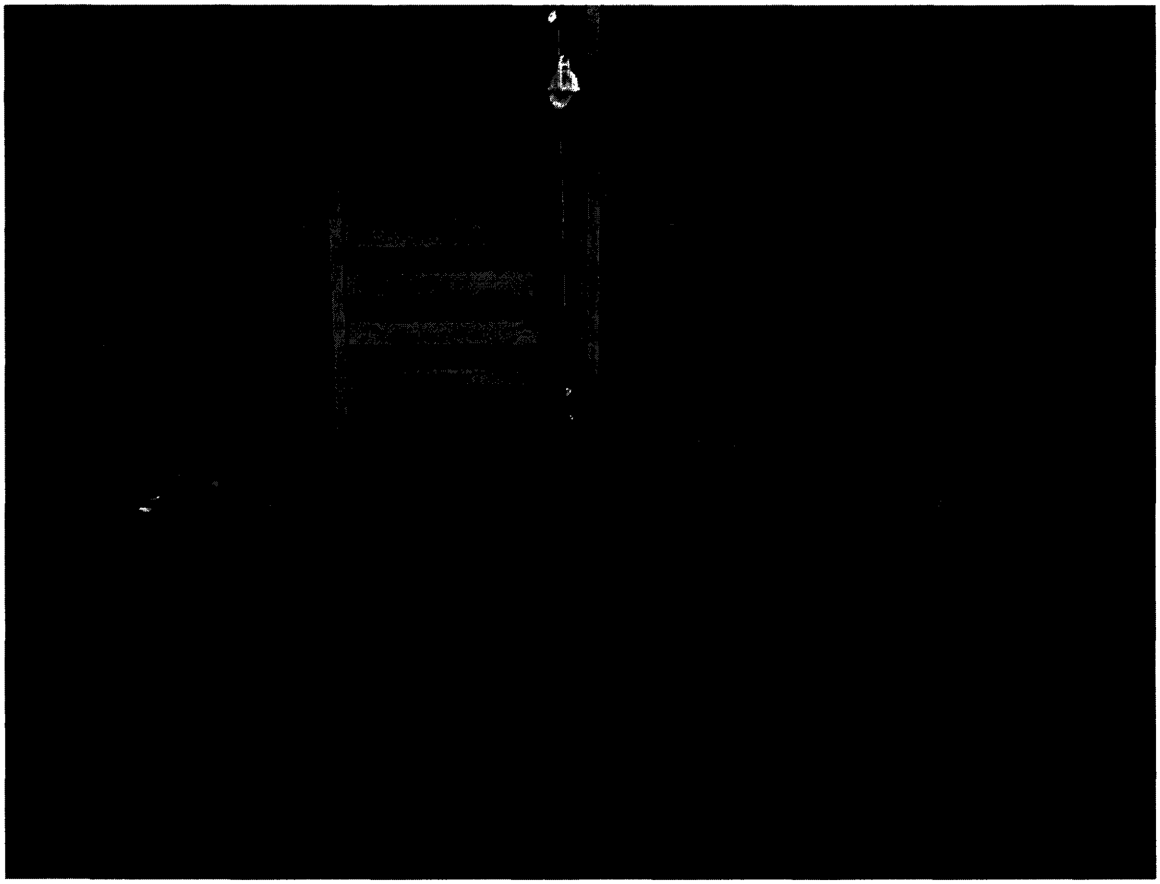

Fig. 1. Photo of the apparatus used in the tamarin and marmoset experiments.

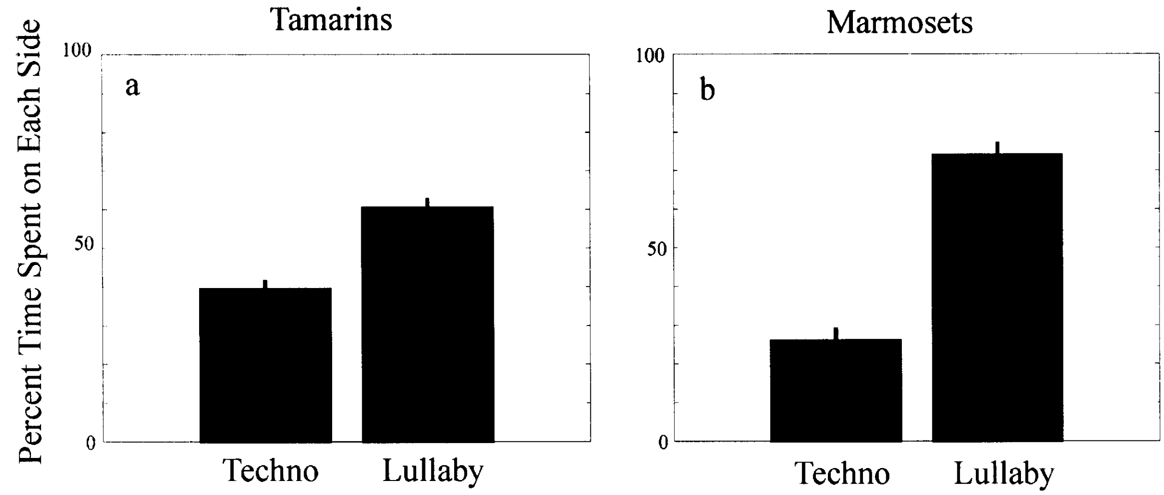

Fig. 2. Results of Experiment 1a, comparing a lullaby played on a flute with electronic "techno" filtered to have the same amplitude envelope. The bars plot the average data from the tamarins and marmosets run in the experiment. Error bars here and elsewhere denote standard errors. (a) Results for tamarin subjects. (b) Results for marmoset subjects.

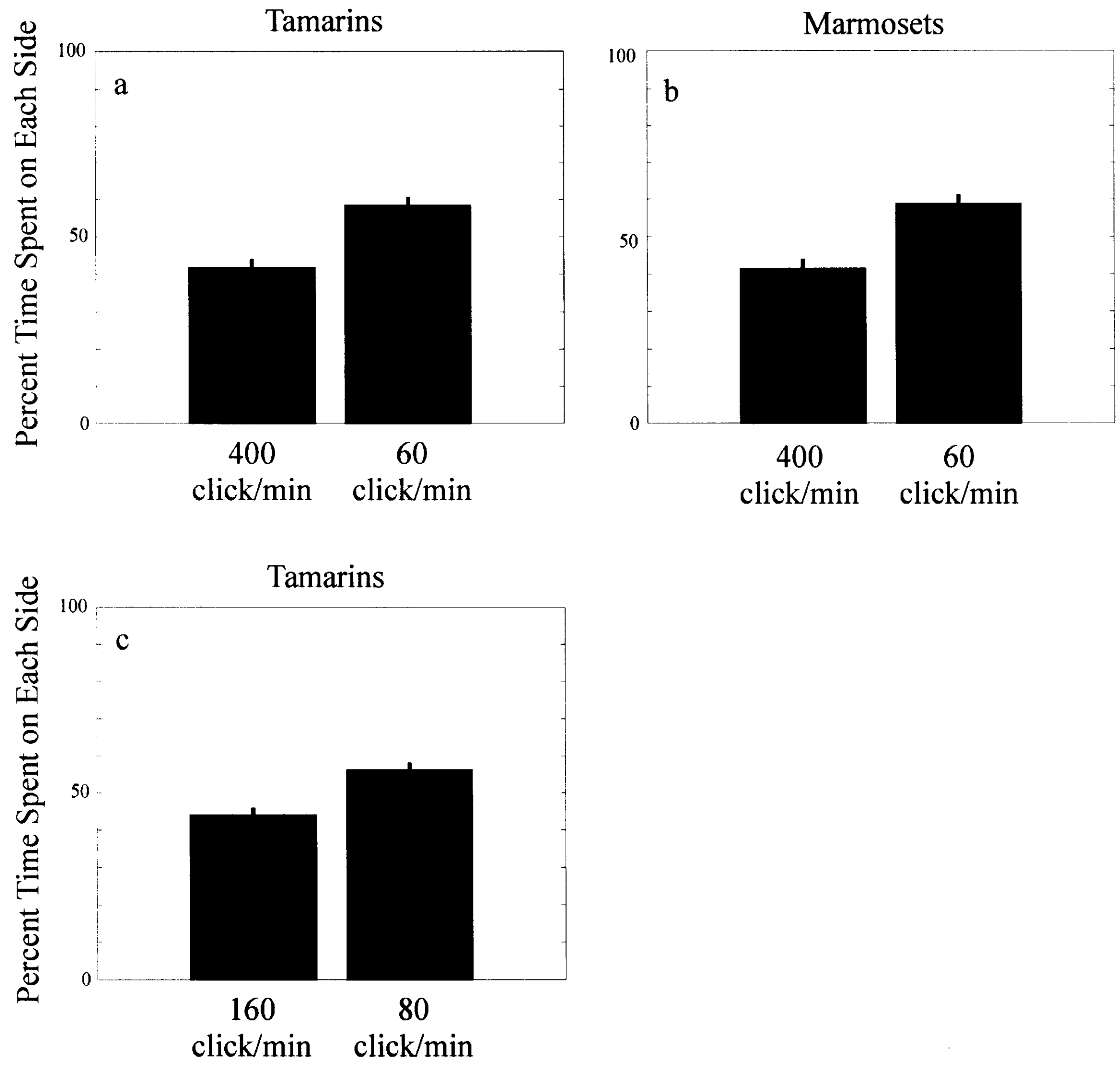

Fig. 3. Results of Experiment 2, comparing slow and fast click trains. (a) Results for tamarin subjects, 60 vs. 400 clicks per minute. (b) Results for marmoset subjects, 60 vs. 400 clicks per minute. (c) Results for tamarin subjects, 80 vs. 160 clicks per minute.

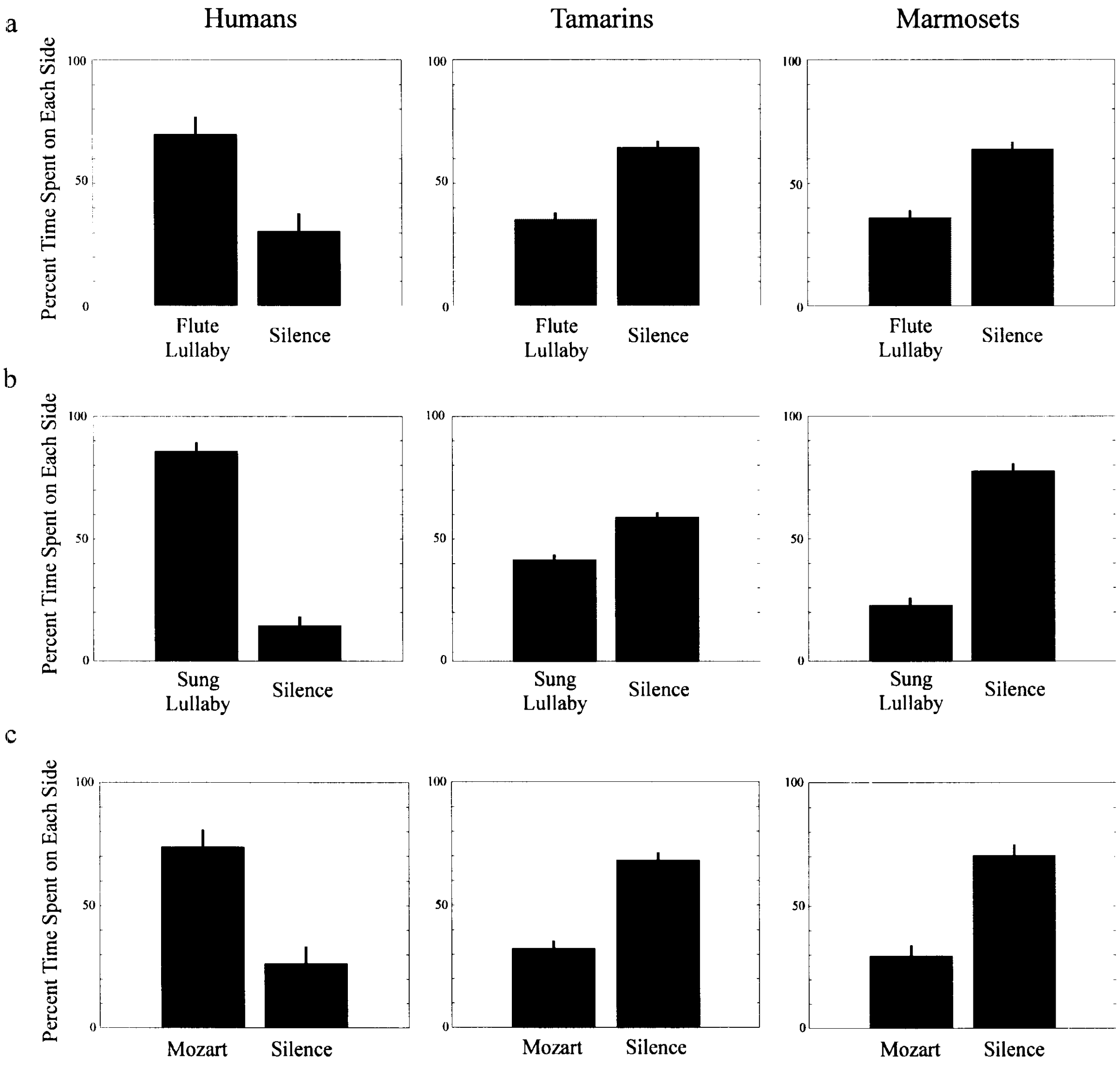

Fig. 4. Results of Experiment 4, comparing various musical stimuli to silence for human, tamarin and marmoset subjects. (a) Results for an instrumental lullaby played on a flute. (b) Results for a lullaby sung in German. (c) Results for a Mozart string concerto.

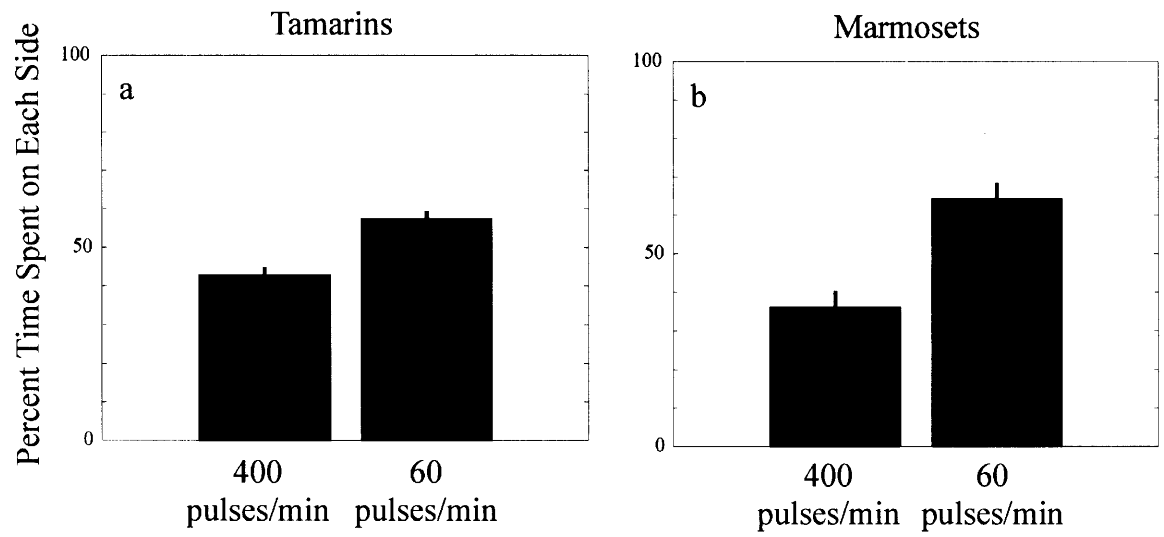

Fig. 5. Results of Experiment 5, comparing fast and slow tempo stimuli equated for acoustic energy. (a) Results for tamarins. (b) Results for a single marmoset subject.

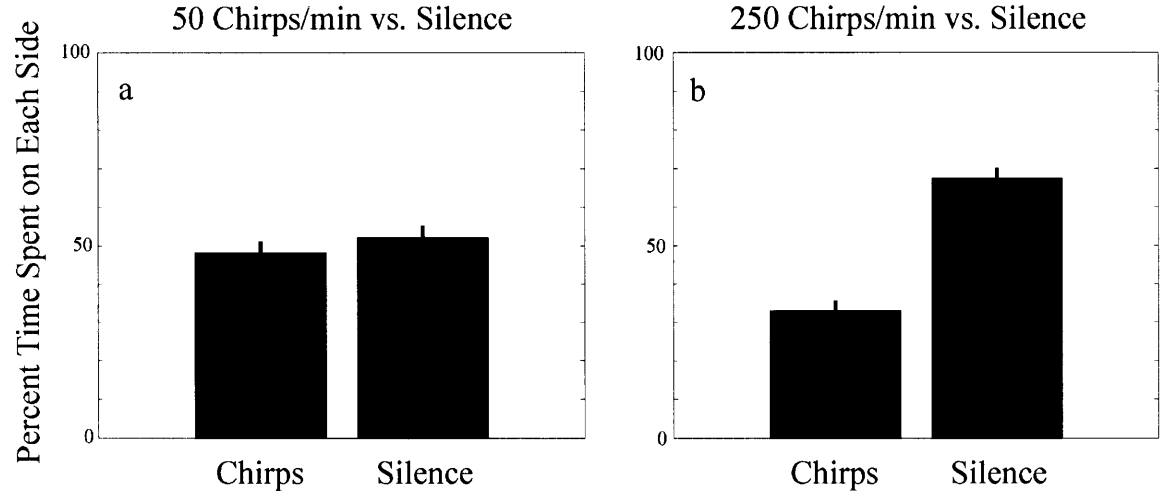

Fig. 6. Results of Experiment 6, comparing slow and fast chirp trains to silence, for 4 tamarins and 1 marmoset. (a) Results for slow chirp trains vs. silence. (b) Results for fast chirp trains vs. silence.

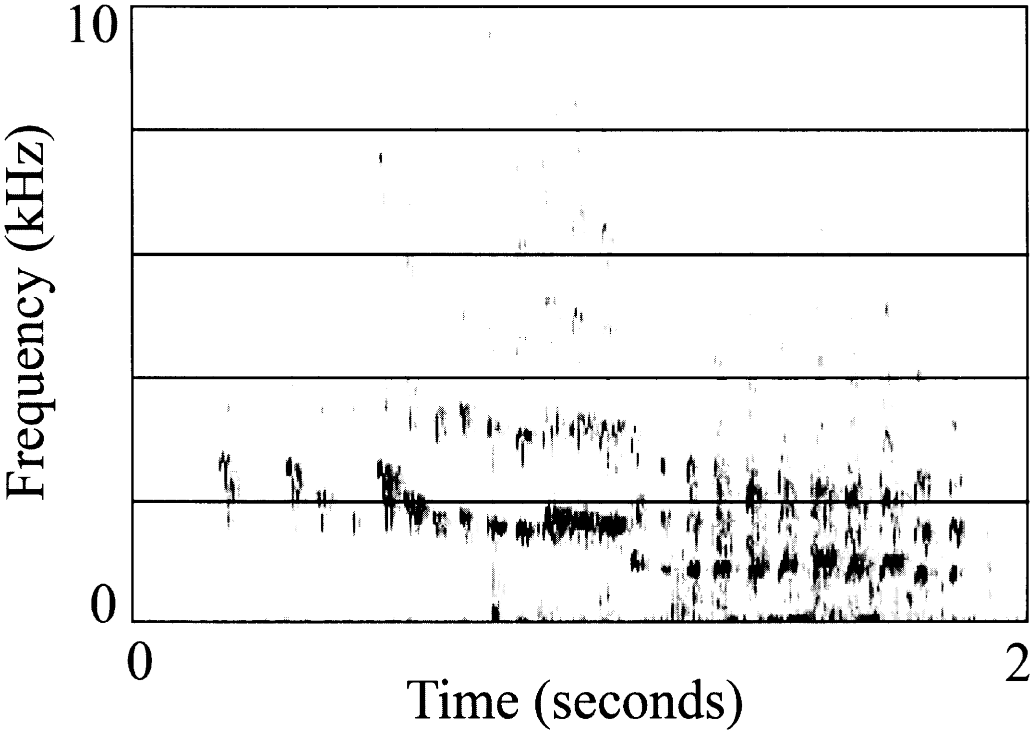

Fig. 7. Spectrogram of a tamarin alarm call.

Chapter 4:

Can music exposure induce preferences for consonance in animals? Evidence from pet dogs

Josh McDermott
Department of Brain and Cognitive Sciences, MIT

Abstract

The origins of the preference for consonant over dissonant musical intervals have been debated since the time of Pythagoras, with some scholars favoring an innate basis, and others contending it is learned from exposure to music. To gain insight into the evolutionary origins of the consonance preference, we explored whether music exposure would produce similar preferences in pet dogs, who receive exposure to music via their owners. We find that dogs fail to exhibit a consonance preference despite the benefit of extensive exposure to music, and despite exhibiting a preference for soft over loud intensities when tested with the same method. Our results suggest that if humans acquire their preference for consonance from incidental exposure to music, they do so with the aid of acquisition mechanisms that are not shared by dogs, and that perhaps are unique to humans.

Introduction

The origins of music, long regarded as one of evolution's great mysteries, have recently become a focus of research and debate (I. Cross, 2001; Fitch, 2006; Huron, 2001; McDermott & Hauser, 2005; Merker, 2000). One approach is to explore whether any characteristics of music have an innate basis, and the distinction between consonance and dissonance has been of considerable interest in this regard. As famously noted by Pythagoras, some combinations of musical notes generally sound better, or more consonant, than others (Butler & Daston, 1968). The contrast between consonant and dissonant pairs of notes, i.e. intervals, is central to Western music, influencing harmony (Dahlhaus & Gjerdingen, 1990), musical tension (Bigand, Parncutt, & Lerdahl, 1996), and scale structure (Huron, 1994). Given their importance in music, it is no surprise that consonance and dissonance have received much attention in auditory research (Helmholtz & Ellis, 1954; Plomp & Levelt, 1965; Terhardt, 1974). By now it is generally agreed that the ability to *discriminate* consonant from dissonant intervals derives in part from differences in how they excite the peripheral auditory system (Fishman et al., 2001; Plomp & Levelt, 1965; Tramo et al., 2001). This discrimination ability is thus presumed to be shared by nonhuman animals, as has been confirmed in birds and monkeys (Hulse et al., 1995; Izumi, 2000).

In contrast, the basis of the *preference* for consonance over dissonance remains unclear (McDermott & Hauser, 2005). Some have supposed that this is learned from music exposure, and that the preference for consonance is therefore the arbitrary result of enculturation (Boulez, 1971; W. J. Dowling & Harwood, 1986). However, recent evidence indicates that infants as young as 2 months of age prefer consonant to dissonant intervals (Trainor & Heinmiller, 1998; Trainor et al., 2002; Zentner & Kagan, 1996, 1998), suggesting that the preference may be present from birth. Conversely, some species of nonhuman primates appear to lack a preference for consonance (McDermott & Hauser, 2004). This apparent species difference could indicate that the consonance preference is innate and unique to humans, making it a candidate for an adaptation for music (McDermott & Hauser, 2005). Alternatively, the preference observed in human infants could be the product of their exposure to music. Infants have obviously heard less music than typical adults, but even brief amounts of exposure have been shown to influence sensitivity to musical structure (Hannon & Trehub, 2005b), and could plausibly affect preferences as well (Saffran, Loman, & Robertson, 2000). The laboratory animals previously tested (McDermott & Hauser, 2004) did not have the benefit of this musical exposure, and it is an open question as to whether their preferences for sounds would be influenced by it.

To explore the effect of music exposure on consonance preferences in animals, we turned to the domesticated dog. Pet dogs receive incidental exposure to music through many of the same channels that infants do, and in similar circumstances, raising the question of whether they would display any of the same musical sensitivities. No psychophysical tests pertaining to consonance have been conducted in dogs to our knowledge, but canine auditory sensitivity has been measured to be similar to that of humans albeit with an extended range of high frequency sensitivity (Heffner, 1983). Dogs have also been shown to discriminate vowels (Baru, 1975), indicating sensitivity to at least one perceptual dimension of importance to humans. To test whether pet dogs had a preference for consonance over dissonance, we adapted a method previously used to measure acoustic preferences in rats, birds and monkeys, in which animals control what they hear by virtue of their location in a confined space (H. A. Cross, Holcomb, & Matter, 1967; McDermott & Hauser, 2004; McDermott & Hauser, 2006; Watanabe & Nemoto, 1998). To ensure that our method was appropriate for measuring preferences in dogs, we ran a control condition in which animals were presented with a choice between loud and soft white noise, on the assumption that the loud noise would be aversive. We also ran analogous experiments in humans to verify that they would exhibit the expected preferences given a similar method.

Method

Procedure

Subjects were placed in a room, part of which was cordoned off with a partition and divided in half with a stripe of tape, as shown in Fig. 1. Speakers were located on either side of the experimental space, covered in black cloth. Each experiment involved a comparison between two stimuli; one stimulus was assigned to the left speaker, and the other to the right. The speakers were controlled with a computer outside the room, and subjects were observed via a video feed. At the start of a test, subjects were given a small piece of food at the center of the experimental space; the experimenter then left the room. Once the animal moved to the left or the right of the dividing line, the experimenter started the stimulus playback. The stimulus for a particular side played continuously as long as the animal was on that side, and switched as soon as it switched sides. The animal's position in the room thus determined which sound it heard. Individual tests lasted 5 minutes. A consistent difference in the amount of time spent with one stimulus compared to the other was taken to indicate a preference between the stimuli.

Each dog was brought in by its owner for an experimental session lasting approximately one hour. After they arrived, dogs were allowed to explore the experimental space until the owners judged them to be comfortable. They were then run in a single familiarization test to introduce them to the experiment format. The stimuli in the familiarization run were low amplitude nonmusical sounds, intended merely to get the animals used to sound presentation in the experimental space. Each animal was then run on two consecutive sessions each of two experiments: one contrasting consonant and dissonant musical intervals, and the other loud and soft white noise. To rule out order effects, half the subjects completed the consonance/dissonance experiments first, and half the loud/soft experiments first. For a given experiment, there were two possible assignments of stimuli to the left and right sides of the room. To rule out side biases as contributing factors to the results, each animal was run once on each side assignment, the order of which was counterbalanced across animals. Data from the two side assignments was combined in our analysis.

The success of our paradigm depended critically on a subject's movement from one side of the room to the other, because in order for a subject to make a putative choice between stimuli, they had to have experienced both stimuli at least once. By far the biggest obstacle to obtaining useful data from our canine subjects was their occasional tendency to stop moving and lie down in one place. We therefore adopted a criterion whereby data was included for a given subject only if they switched from one side of the room to the other at least once in each of the four test sessions. By this criterion we obtained useful data from exactly half of our subjects.

The human experiments were closely matched to the animal experiments. Subjects were told they would be left in the experimental space for periods of 5 minutes, during which they could do as they wished subject to the constraint of not sitting down. No other instructions were given. Each subject was run in one test with consonant and dissonant stimuli, and one test with loud and soft white noise stimuli. The order of the experiments and the side assignment was counterbalanced across subjects. Stimuli were identical to those used with the canine subjects. Subjects were run in the same room used for the dog experiments.

After the test sessions ended, human subjects were asked to rate their preference between the two pairs of stimuli on a scale of 3 to –3, with 3 meaning they preferred the stimulus on the right and –3 meaning they preferred the stimulus on the left. These were sign reversed where appropriate and averaged to yield a preference rating for consonance over dissonance and soft over loud white noise.

Subjects

Our canine subjects (Canis familiaris) were pets belonging to MIT students and staff. Ages ranged from 1-13 years, with a mean of 6.26 (SE = 1.18). We did not select according to breed. 32 animals participated in the experiments. Data from 16 of these were discarded for not switching sides in at least one of the test sessions, as described above. Of the 16 dogs whose data we analyzed, 6 were male and 10 were female. Some of the dogs were adopted as puppies (which we defined as less than four months of age, n=7), some were acquired as adults from shelters or breeders (n=4), and the remainder were adopted as adults prior to living with other human families (n=5). Owners were asked to estimate how much exposure to music their dogs received on a daily basis. For the 16 dogs included in our analysis, these estimates ranged from a low of 30 minutes to a high of 12 hours a day, with a mean of 3.83 hours (SE = 0.99).

Our 10 human subjects were MIT undergraduates; 4 were male and 6 were female.

Stimuli

The loud and soft white noise stimuli were 62 and 78 dB in level, as measured with a sound meter at the center of the room. The consonant and dissonant stimuli were random sequences of musical intervals played with complex synthetic tones. The consonant intervals were octaves, fifths and fourths, and the dissonant intervals were minor seconds, tritones, and minor ninths. The bass note of the intervals was always the A above middle C, such that all frequency components in the stimuli were well within the audible range for dogs (Heffner, 1983), and overlapped with the frequency range of typical dog vocalizations (Yin, 2002). Both sets of intervals were 68 dB in level.

Results

Results of the experiments with human subjects are shown in Figure 2. As expected, human subjects exhibited robust preferences for soft over loud noise ($t(9) = 6.75$, $p < .0001$), and for consonance over dissonance ($t(9) = 5.0$, $p<.001$) as measured by the time spent with one stimulus compared to the other. The preference ratings mirrored these behavioral results, averaging 2.2 for soft over loud noise ($t(9) = 6.74$, $p < .0001$), and 1.8 for consonance over dissonance ($t(9) = 5.01$, $p<.001$). The magnitude of the consonance preference did not significantly differ from the intensity preference via either measure (behavior: $t(9) = 1.06$, $p = .32$; ratings: $t(9) = 0.8$, $p=.44$), at least not with our moderate sample size.

Results of the experiments with canine subjects are shown in Figure 3. Dogs showed a highly significant preference for soft over loud noise as measured behaviorally ($t(31) = 7.07$, $p <$ 10^-7). In contrast, no such preference was observed for consonance over dissonance ($t(31) = 0.58$, $p = .56$). To ensure that order effects were not diluting the results, either due to fatigue or delay in learning that the acoustic environment was subject-controlled, we separately analyzed the consonant/dissonant results for the subjects who completed that experiment first, and for those who completed it last. Neither group showed a significant trend (consonant/dissonant first: $t(15) = 0.83$, $p = .42$; consonant/dissonant last: $t(15) = 0.01$, $p = .99$). In contrast there were significant effects in the loud/soft experiment irrespective of order (loud/soft first: $t(15) = 7.02$, $p<$10^-5; loud/soft last: $t(15) = 3.9$, $p=.001$). We also separately analyzed the dogs that were adopted when they were still puppies. We again found significant preferences for soft over loud ($t(13) = 3.89$, $p = .001$), but not for consonance over dissonance ($t(13) = -0.83$, $p = 0.42$). Many of the other dogs were probably also exposed to music as puppies, but this group clearly had exposure during this period in their lives, and still showed no preference between the two classes of musical intervals.

Data from individual animals are plotted in Figure 4. Animals nearly always avoided the loud noise; in comparison, they seem to have behaved randomly with respect to the consonant/dissonant distinction.

Conclusions

Although it remains unclear whether music exposure plays a role in the consonance preferences observed in human infants, such exposure is apparently insufficient to trigger similar preferences in domesticated dogs. Even dogs who were adopted as puppies, and who thus received musical exposure from a very young age, did not exhibit consistent preferences for consonant over dissonant intervals. Clearly, the musical experience of pet dogs, while extensive, is not identical to that of the average human infant. Infants are routinely sung to by their parents (Sandra E. Trehub et al., 1997), and although some dog owners report engaging in this behavior, it is unlikely to be as widespread as its infant-directed counterpart. Notably, however, infant-directed singing lacks harmony, as there is usually only a single parent doing the singing. The exposure of human infants to consonance is thus likely to occur predominantly through the passive playback of musical materials (television, radio etc.), which dogs also receive. If a consonance preference were the product of mere exposure (Zajonc, 1968), it would thus seem that dogs' exposure would be as likely to produce it as that of humans.

It might also be argued that the presence of other humans during incidental music exposure could critically influence infants' internalization of music, perhaps by conditioning positive associations. Out of all the various species of nonhuman animals, dogs, having been bred to value human company (Serpell, 1995), are perhaps most likely to be similarly sensitive to the social contexts in which music is frequently enjoyed (Hare, Brown, Williamson, & Tomasello, 2002). Moreover, the fact that dogs receive much of their music exposure in the company of humans suggests they ought to benefit from these social influences to much the same extent that human infants do. We did not test whether these experiences cause music to be associated with positive connotations for dogs, but if they do, such associations do not seem to induce preferences for consonance.

The consonance preferences observed in human infants thus appear to be the product of something not shared by dogs, and which might be uniquely human. It is possible that the effects observed in infants are dependent on musical exposure, but if so, we would suggest that they implicate acquisition processes that are attuned at least somewhat to musical structure. Exposing other species to the same stimuli apparently does not produce the same preferences, suggesting that these acquisition processes, if they exist, may be unique to humans. Such processes might be music-specific, or could instead be side effects of other mechanisms, such as those for language acquisition (Vouloumanos & Werker, 2004). Alternatively, consonance preferences in human infants may be independent of musical experience, and might represent part of an adaptation to enable humans to appreciate music, the putative function of which remains to be determined.

Acknowledgements

I thank Molly Jabas, Meredith Brown, and Carrie McDonough for help setting up and conducting the experiments, John Canfield for technical assistance, and all the dog owners who volunteered their animals for this study.

References

Baru, A. V. (1975). Discrimination of synthesized vowels [a] and [i] with varying parameters (fundamental frequency, intensity, duration, and number of formants) in dog. In G. Fant & M. A. A. Tatham (Eds.), *Auditory Analysis and Perception of Speech* (pp. 91-101). New York: Academic Press.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics, 58*(125-141).

Boulez, P. (1971). *Boulez on music today*. Cambridge, Mass.,: Harvard University Press.

Butler, J. W., & Daston, P. G. (1968). Musical consonance as musical preference: A cross-cultural study. *Journal of General Psychology, 79*, 129-142.

Cross, H. A., Holcomb, C. G., & Matter, W. W. (1967). Imprinting or exposure learning in rats given early auditory stimulation. *Psychonomic Science, 7*, 233-234.

Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Science, 930*, 28-42.

Dahlhaus, C., & Gjerdingen, R. O. (1990). *Studies in the Origin of Harmonic Tonality*. Princeton, NJ: Princeton University Press.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. Orlando: Academic Press.

Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., et al. (2001). Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology, 86*(6), 2761-2788.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, in press.*

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, 102*, 12639-12643.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The domestication of social cognition in dogs. *Science, 298*, 1636-1639.

Heffner, H. E. (1983). Hearing in large and small dogs: Absolute thresholds and size of the tympanic membrane. *Behavioral Neuroscience, 97*, 310-318.

Helmholtz, H. v., & Ellis, A. J. (1954). *On the sensations of tone as a physiological basis for the theory of music* (2d English , ed.). New York,: Dover Publications.

Hulse, S. H., Bernard, D. J., & Braaten, R. F. (1995). Auditory discrimination of chord-based spectral structures by European starlings (Sturnus vulgaris). *Journal of Experimental Psychology: General, 124*(4), 409-423.

Huron, D. (1994). Interval-class content in equally tempered pitch-class sets: Common scales exhibit optimum tonal consonance. *Music Perception, 11*(3), 289-305.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Science, 930*, 43-61.

Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America, 108*(6), 3073-3078.

McDermott, J., & Hauser, M. (2004). Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition, 94*(2), B11-B21.

McDermott, J., & Hauser, M. D. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception, 23*(1), 29-59.

McDermott, J., & Hauser, M. D. (2006). Nonhuman primates prefer slow tempos but dislike music overall. *Cognition, In press.*

Merker, B. (2000). Sychronous chorusing and human origins. In B. Merker & N. L. Wallin (Eds.), *The origins of music.* (pp. 315-327). Cambridge, MA, US: The MIT Press.

Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America, 38*, 548-560.

Saffran, J. R., Loman, M. M., & Robertson, R. R. (2000). Infant memory for musical experiences. *Cognition, 77*(1), B15-23.

Serpell, J. e. (1995). *The Domestic Dog, its evolution, behavior and interactions with people.* Cambridge, UK: Cambridge University Press.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55*, 1061-1069.

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21*(1), 77-88.

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception, 20*(2), 187-194.

Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. *Annals of the New York Academy of Science, 930*, 92-116.

Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology, 51*(4), 385-396.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the priveleged status of speech for young infants. *Developmental Science, 7*(3), 270-276.

Watanabe, S., & Nemoto, M. (1998). Reinforcing property of music in Java sparrows (Padda oryzivora). *Behavioural Processes, 43*(2), 211-218.

Yin, S. (2002). A new perspective on barking in dogs (Canis familiaris). *Journal of Comparative Psychology, 116*(2), 189-193.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*, 1-27.

Zentner, M. R., & Kagan, J. (1996). Perception of music by infants. *Nature, 383*(6595), 29.

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development, 21*(3), 483-492.

Figure 1. Photograph of experimental setup.

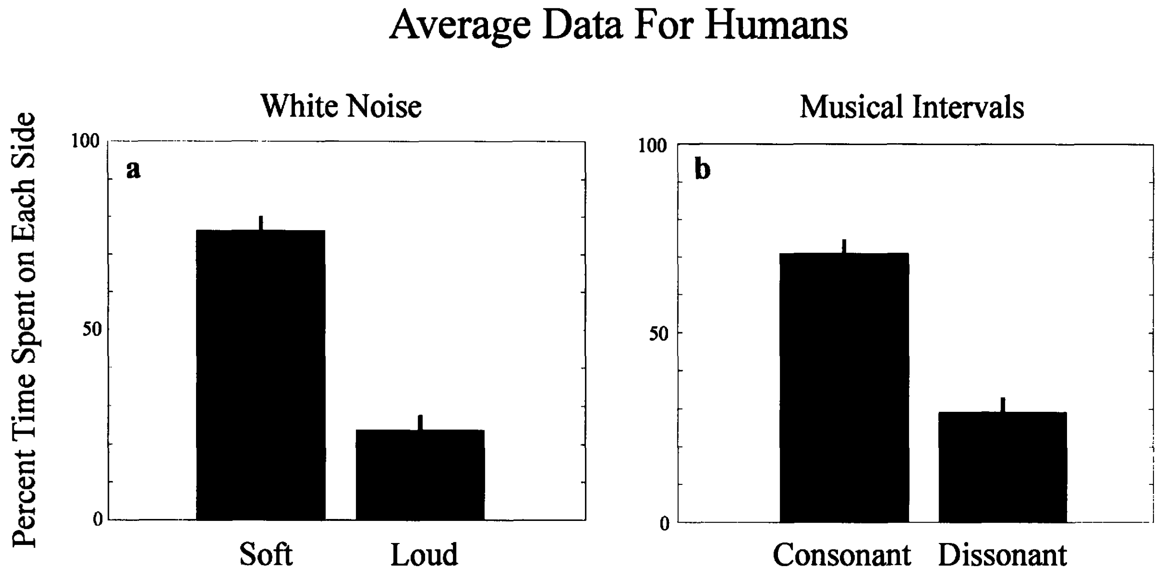

Figure 2. Average results from human subjects, for the experiment with (a) loud and soft white noise and (b) consonant and dissonant intervals. Error bars denote standard errors.

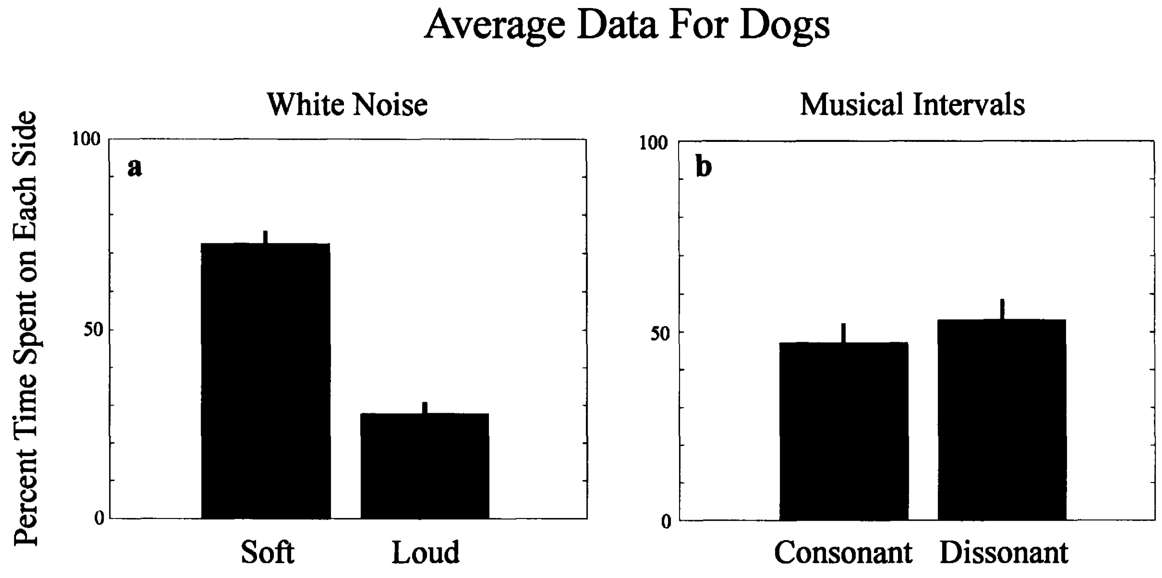

Figure 3. Average results from canine subjects, for the experiment with (a) loud and soft white noise and (b) consonant and dissonant intervals. Error bars denote standard errors.

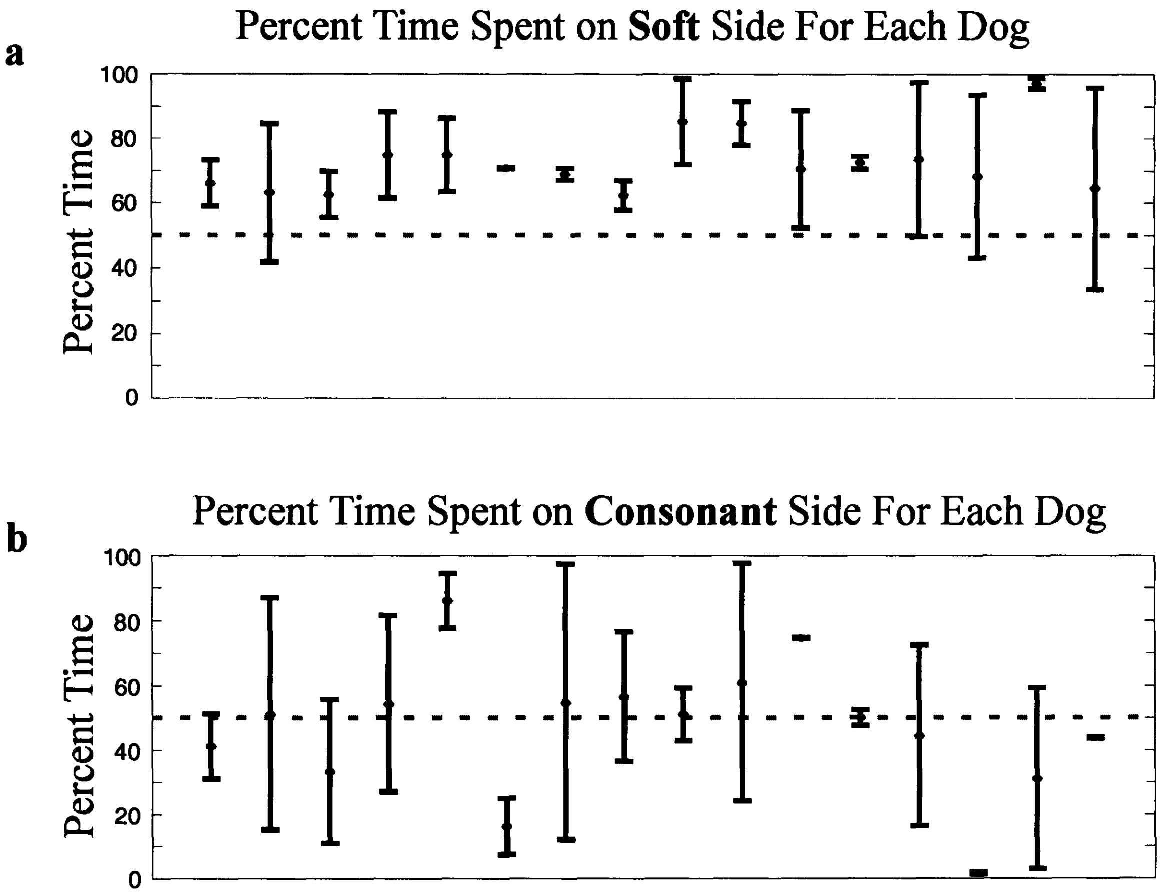

Figure 4. Results from individual dogs for the two experiments. Individual data points denote the average time spent by an individual animal on (a) the soft side in the white noise experiment or (b) the consonant side in the musical interval experiment. Data points are listed in the order in which the animal was run, such that the nth data point in each of the two plots represents data from the same dog. Error bars denote half the difference between the results from the two test sessions run for each animal in each experiment, such that the ends of the error bars represent the results from the individual sessions.

Chapter 5:

Mere Exposure Effects for Music in a Nonhuman Primate

Josh McDermott
Department of Brain and Cognitive Sciences, MIT

Marc Hauser
Department of Psychology, Harvard University

Abstract

The relative importance of innate constraints and cultural exposure in shaping music perception has been debated for centuries. Recent findings indicate that human infants possess an impressive array of musical abilities, consistent with the presence of innate predispositions. Alternatively, the abilities observed in infants might originate from the exposure they inevitably receive prior to and during the earliest post-natal stages. To test the extent to which exposure to music can influence music preferences through domain-general, mere exposure effects, we exposed a colony of common marmoset monkeys to musical stimuli over long periods of time, subsequently testing them for their music preferences. Exposure to random sequences of consonant intervals and to Mozart concertos both produced preferences for the familiar exposure stimulus over other, unfamiliar stimuli, but only in infant monkeys; adults were apparently unaffected. Even in infants, however, the preferences induced were specific to the exposure stimuli. When tested for preferences for consonance over dissonance using novel stimuli, both infants and adults failed to display a preference. The preference exhibited by human infants and adults therefore appears to minimally depend on acquisition mechanisms that are not shared by nonhuman primates, and that may selectively target musical structure. Alternatively, the consonance preference might be innate in humans, and potentially part of an adaptation for music.

keywords: consonance, dissonance, music, preference, monkey, mere exposure, critical period

Introduction

Much of the debate over the evolutionary origins of music centers on whether any aspects of music perception are innate, and therefore potential products of natural selection (McDermott & Hauser, 2005). One obvious source of evidence for innateness is cross-cultural studies (Kessler et al., 1984). Unfortunately, such studies of music perception have become practically difficult to conduct because of the ubiquity of Western popular music worldwide. A more promising and productive alternative has been to probe the perception of young infants (Sandra E. Trehub, 2003), whose exposure to music dwarfs that of adults. Recent findings indicate that human infants encode melodies (S. E. Trehub et al., 1984), perceive meter (Hannon & Trehub, 2005a), and prefer consonant to dissonant intervals (Trainor et al., 2002; Zentner & Kagan, 1996) much the way that adults do. These results are consistent with the possibility that humans are born with such abilities and that they manifest themselves largely independent of musical input. Alternatively, the musical abilities found in human infants might be learned from their few months of music exposure. Such a dependence on exposure need not rule out innate music-related machinery, as in exposure-dependent acquisition of language (Chomsky, 1986). Unlike language, however, the case for domain-specific acquisition mechanisms for music is at present less than clear.

Studies with nonhuman animals have the potential to help shed light on these issues, as animals do not make music, and therefore by definition must lack domain-specific mechanisms for the perception, acquisition, and production of music (McDermott & Hauser, 2005). In addition, the music exposure of laboratory animals, unlike human infants, can be eliminated or carefully titrated to reveal the extent to which music-related effects depend on exposure to music. In this paper we examine the effects of music exposure on music preferences in nonhuman primates. It is well known that "mere" exposure to a stimulus can produce preferences for that stimulus over unfamiliar alternatives (Zajonc, 1968), and that this effect can occur for music (Meyer, 1903). Such effects have been documented in human infants for specific musical pieces (Saffran et al., 2000), and there is one report of mere exposure effects for in rats (H. A. Cross et al., 1967). What remains unclear is whether mere exposure can account for preferences for general features of music, such as consonance.

Although the role of music exposure is potentially critical to most aspects of music perception, we will focus specifically on the preference for consonant over dissonant intervals. Central to the music of Western cultures (Dahlhaus & Gjerdingen, 1990; Huron, 1991), the consonant/dissonant distinction has been widely discussed since at least the time of the Greeks. The preference is widespread among Western adults (Butler & Daston, 1968), and thought to be present in infants as young as two months of age (Trainor et al., 2002), raising the possibility that it might have an innate basis. We have previously found that nonhuman primates fail to demonstrate spontaneous preferences for consonant over dissonant intervals, suggesting that the preference is perhaps unique to humans. Alternatively, the absence of the preference in laboratory animals could be due to their lack of exposure to music, which both adult and infant humans have in abundance.

One key motivation for testing the effects of exposure on animals' preferences for consonance, as opposed to some other feature of music, is that it seemed conceivable to replicate in most respects the exposure that infants receive to consonance. This is less true of other aspects of music; the ubiquity of infant-directed singing by parents (Sandra E. Trehub et al., 1993) is obviously not something we might hope to mimic with animal subjects, and such singing might plausibly play an important role in how infants internalize music. However, exposure to consonance rarely if ever occurs through such channels, because infant directed singing typically consists of solo performances, which by definition lack harmony. Consonance exposure thus seems likely to occur predominantly through passive exposure, which we can readily impose on nonhuman animals.

Prior to conducting this study one of us (McDermott, 2006) tested pet dogs for consonance preferences, motivated by the same set of issues. Rather than experimentally manipulate music exposure, McDermott took advantage of the fact that dogs receive incidental music exposure from their owners through many of the same channels that young humans do. Unlike human subjects tested with the same method, however, and despite receiving an average of nearly 4 hours of music exposure a day as reported by their owners, dogs displayed no evidence of a preference for consonance. Given that the incidental and varied music exposure of pet dogs apparently did not induce a preference for consonance, we adopted a slightly different approach in the present paper - one of brute force, maximizing exposure in the hopes that this might produce an effect. We manipulated the music exposure of a colony of common marmosets in two phases. In the first phase, the colony received random sequences of consonant intervals 12 hours a day, as a test of whether massive exposure to nonstop consonance might induce a consonance preference. In the second phase, to test the possible importance of musical structure and variability, the colony was exposed to a set of Mozart concertos. Because it seemed conceivable that there might be critical or sensitive periods for the acquisition of auditory material (Nakahara, Zhang, & Merzenich, 2004; Newport, 1990), we tested infants as well as adults.

Our subjects in the present study, common marmosets, are a species of new World monkey, native to the South American rain forest. The auditory system of the marmoset is similar to that of humans so far as we know. The audiogram is reported to cover roughly the same range (Seiden, 1958). Their vocalizations are typically high in pitch, but the recent discovery of low frequency pitch neurons in marmosets (Bendor & Wang, 2005) suggests that a wide spectrum of sounds is behaviorally relevant to this species.

To measure preferences for one acoustic stimulus over another we used a method that had been previously used to measure acoustic preferences in rats (H. A. Cross et al., 1967), birds (Watanabe & Nemoto, 1998), and dogs (McDermott, 2006), and that we had previously used to test for consonance and tempo preferences in tamarins and marmosets (McDermott & Hauser, 2004; McDermott & Hauser, 2006).

General Methods

Subjects

We tested 11 common marmosets - 5 adults (3 male, 2 female) and 6 infants (3 male, 3 female). 4 of the infants were born one month prior to the start of the first phase of the exposure, and 2 were born four months after it began. Infants were between 4-8 months at the time of testing. Due to the colony's gradual departure from the host lab towards the end of the study, not all animals were available for all experiments.

Procedure

The apparatus (a V-shaped maze, Fig. 1) and procedure were identical to that of McDermott and Hauser (2006). We placed speakers (PolkAudio Atrium 45p) at the end of each branch of the maze; each speaker and branch was paired with a different sound. Subjects were initially placed at the entrance to the maze. The experimenter then left the room and raised the door to the maze by means of a pulley, allowing the animal to enter. We placed small pieces of food in both branches of the maze, equidistant from the entrance, to entice them to enter; the animals always ate both pieces of food. When the subject moved into a branch for the first time, the experimenter started the playback. The stimulus for a particular side played continuously as long as the animal was on that side, and switched as soon as it switched sides. The animal's position in the maze thus determined which sound it heard. Testing sessions lasted 5 minutes. After a certain number of sessions with one sound on the left side and another on the right, generally separated by a full day, the sound-side pairing was reversed, and the animals were run for the same number of additional sessions. Data before and after the reversal were combined to distinguish stimulus preferences from side biases, and the initial side assignment was always evenly divided between animals. The number of sessions for a single animal in a single experiment was usually 6 or 8. It was always fixed in advance of the experiment.

The computer program that turned the sounds on and off generated a list of the times at which one sound was started and the other was stopped. These times were almost identical to those extracted by blind observers from video recordings of 10 sessions randomly selected from Experiments 3 and 6 ($r= .99$). The total time that a subject spent on each side was computed from these switch times, and never differed by more than 3 sec from the totals computed from the video-derived switch times. A statistically significant bias to spend more time in one branch than the other was taken as evidence for a preference for one sound over the other. Throughout, we test for statistical significance using one sample, two-tailed t-tests, testing if the average proportion of time spent with one of the two stimuli differed from 50%. Thresholds for significance were set at $p = 0.05$.

Infant marmosets generally will not come out of their cages until they are at least 3-4 months of age, and even then appear stressed when they are away from their parents for long periods of time. To obtain useful data from the infant monkeys, we began by familiarizing them with the experimental apparatus in the absence of acoustic stimulation. They were placed in the apparatus for periods of up to 5 minutes, and occasionally given food. Over time the infants became progressively calmer in the apparatus, and typically after a month or so we deemed them ready for their first experiment.

Stimuli

The exposure stimulus of Phase 1 was a random sequence of octaves, fifths and fourths, the pitch, timbre and duration of which varied. The bass note of each interval was drawn randomly from the set [A5, B5, C#5, D5, E5, F5], with fundamental frequency ranging from 440 to 740 Hz (notes were drawn from the equal-tempered scale). The notes in each interval were complex tones with 10 frequency components. All stimulus frequencies thus fell well within the range of audibility for the marmoset (Seiden, 1958). The timbre was varied by randomly setting the amplitude envelope of the complex tones composing each interval to one of three profiles, and by randomly setting the amplitude spectrum of the tones to one of three profiles. The duration of each interval was randomly set to 0.75, 1.5 or 2.25 sec. The stimulus played from 7am every morning until 7pm every night. The average intensity of the stimulus was fixed at approximately 62 dB as measured with a sound level meter held at the monkeys' cages.

The Mozart exposure stimulus of Phase 2 was a sequence of MIDI renditions of movements from Mozart piano concertos and string quartets, lasting two hours in total (K.281, second movement; K.311, second movement; K.330, second movement; K.122, first movement; K.475, second movement; K533, second movement; K.570, second movement; K.175, second movement; K.453, second movement; K.458, third movement; K.80, first movement; K.156, second movement; K.172, second movement; K.593, second movement). We intentionally chose movements played at slow tempos, as our previous research had suggested that, when given a choice, marmosets prefer slow tempos to fast (McDermott & Hauser, 2006). This was repeated six times a day for 12 hours of stimulation, again between 7am and 7pm, mirroring the marmosets' wake-sleep cycles. The average intensity of the stimulus was again fixed at 62 dB.

Note that throughout this study the animals were being run in other experiments, and so were taken out of the homeroom several times a day, for periods of time ranging from 5 to 30 minutes, for between 1-5 times a day. They thus heard the exposure intermittently, but for the majority of the 12-hour exposure period.

Experiment Order

To aid clarity, the order in which we discuss experiments is not always the order in which they were conducted. Moreover, due to developmental constraints, the different groups of subjects did not always complete each experiment at the same time. Fig. 2 depicts the timeline of experiments and exposure phases for the adults and the two groups of infants.

Phase 1: Consonance Exposure

The first phase of the study involved exposing the marmosets to random sequences of consonant intervals, with the expectation that this might induce preferences for consonance via mere exposure. This stimulus lacked much of the temporal structure found in typical music, but contained non-stop consonance. Despite the lack of musical structure, the stimulus was judged to be pleasant by human observers; 8 Harvard undergraduates who were given a choice between this exposure stimulus and silence in a human analogue of the animal experiments (see our previous publications for detailed methods (McDermott & Hauser, 2004; McDermott & Hauser, 2006)) spent significantly more time with the exposure stimulus than with silence (74.89% on average; $t(7) = 10.22$, $p < 10^{-4}$). Afterwards, when asked to rate it on a scale of 3 (pleasant) to –3 (unpleasant), they rated it significantly higher than 0 (0.88 on average; $t(7) = 2.97$, $p = .02$), though not as high as actual pieces of music (see data in McDermott & Hauser (2006)).

Experiment 1: Consonance vs. Dissonance in Adult Monkeys

Before beginning the homeroom exposure, we tested the 5 adult marmoset subjects to assess whether they, like our previously studied tamarin subjects, would fail to exhibit a preference for consonance over dissonance. Results confirmed this prediction (Fig. 3a; $t(39) = -0.30$, $p = 0.77$; 8 sessions per animal). We had previously confirmed that the 5 adult marmoset subjects demonstrated a preference for soft over loud white noise (McDermott & Hauser, 2006), suggesting that the method had some validity as a measure of preference. Before beginning tests with our infant subjects, we demonstrated this intensity effect in them as well (see Experiment 2). Previously (McDermott & Hauser, 2004), we showed that humans display the expected preference for consonance when tested with a comparable method, suggesting that if a similar preference was present in our nonhuman primate subjects, our method ought to demonstrate it.

Once the pretests with consonance and dissonance were complete, we began the first exposure phase, exposing our subjects to consonant intervals 12 hours a day, and testing them after one month, two months, and six months of exposure.

Methods

The test stimuli were random sequences of equal-tempered consonant and dissonant intervals in A major, with a sinusoidal amplitude envelope and a fixed duration of 1.5 sec. The intervals were composed of complex tones with 10 frequency components. Humans exhibit a pronounced preference for similarly constructed consonant stimuli (McDermott & Hauser, 2004).

At the time of the second test, one of the five adult marmosets was sick, so only four were run. In the other tests all five animals were tested. In the consonance tests, we ran 8 sessions per animal; in the loud/soft noise test, 3 subjects were run for 6 sessions, and 2 subjects for 4 sessions, because these effects turned out to be quite large.

Results

In no case was there a significant difference between the time spent with consonance and the time spent with dissonance (Fig. 3b; one month: $t(39) = -1.32$, $p= 0.19$; two months: $t(31) = -0.26$, $p= 0.79$; six months: $t(39) = 0.20$, $p = 0.84$). The proportion of time spent by individuals on the consonant side in the three tests ranged from 43.0% to 55.09%, and never significantly differed from 50%. We also analyzed the data combined across the three tests both for the entire group of subjects and for each subject individually. In neither case was there ever a significant preference for one stimulus over the other (group: 48.41% spent with consonance, $t(111) = -0.89$, $p=0.38$; individuals ranged from 46.5% to 52.57%, with t values never in excess of 1.22, and the corresponding p values never less than 0.23). Following the last of the consonance tests, we conducted a control experiment with white noise stimuli that differed in level by 8dB, to confirm that our subjects had not habituated to our method. All subjects consistently spent more time on the side of the softer white noise, ranging from 60.08% to 88.19% (Fig. 3c; t(25) = 10.25, p<10^-9). The robustness of this effect with a relatively small difference in amplitude suggests that our measure was reasonably sensitive, and that the absence of an effect with consonant and dissonant intervals reflects the absence of a preference between these stimuli in our marmoset subjects.

Experiment 2: Consonance vs. Dissonance in Infant Monkeys

Given prior research on auditory plasticity (Nakahara et al., 2004) and various other lines of evidence for critical periods for learning in both humans and nonhuman animals (Hannon & Trehub, 2005b; Hess, 1973; Hubel & Weisel, 1970; Knudsen, 2004; Newport, 1990), it seemed plausible that infant subjects might be more affected by music exposure than adults. We therefore tested the four infant marmosets that were born one month before the start of the exposure, to check whether the exposure during their infancy might have had an effect.

Methods

We ran each animal for 8 sessions. The experiment was conducted six months after the start of Phase 1. The stimuli were identical to those used in Experiment 1.

Results

When tested with white noise stimuli differing in level by 8 dB, all four infants showed pronounced tendencies to spend more time on the side of the soft noise, with proportions ranging from 67.01% to 77.58% (Fig 4a; t(31) = -9.45, p<10^-9; 8 sessions per subject). Like the adults, however, the infants showed no inclination to spend more time with consonance compared to dissonance (Fig. 4b; t(31) = -0.34, p = 0.73). Apparently the exposure stimulus was insufficient to induce a general preference for consonance, even in infant marmosets.

Experiment 3: Consonance vs. Dissonance with Exposure Stimulus

Given the prior report of mere exposure effects for musical stimuli in rats (H. A. Cross et al., 1967), we wondered if the marmosets would exhibit preferences if tested with stimuli that were near replicas of the exposure stimulus. The intervals in the test stimuli of Experiments 1 and 2 were fixed in A major, and were composed of tones that had a fixed envelope, amplitude spectrum, and duration. In contrast, the intervals in the exposure stimulus varied in pitch, envelope and duration. The exposure stimulus thus contained all the intervals of the test stimuli in exact form, but contained other intervals as well. As a result it was much more variable in pitch, timbre, and temporal structure, and was clearly distinct from the test stimuli that we used. We therefore tested whether our marmoset subjects would demonstrate a preference for consonance when tested with stimuli drawn from the same distribution as the exposure stimulus.

Methods

The consonant test stimuli were drawn from the same distribution as the consonant exposure stimulus. The dissonant test stimulus was similarly generated, but with the intervals changed to minor seconds, tritones, and minor ninths (as were used in the dissonant test stimulus of Experiments 1 and 2).

We tested our 5 adult subjects and 4 infants – 3 of the group who were born just prior to the start of the exposure, and 1 of the pair who were born four months after the beginning. We ran each subject for 8 sessions.

Before testing the two youngest infants, we first ran them in a control experiment with white noise stimuli differing in amplitude by 8 dB, as had been done earlier with the older group of infants. Both of the youngest infants spent more time on the side of the soft noise, 62.54% and 75.87%, respectively ($t(11) = -3.8$, $p= 0.003$; 6 sessions per subject).

Results

The adults again showed no evidence for a preference for the consonant stimulus over the dissonant one (Fig. 5a; $t(39) = -0.06$, $p = 0.96$; proportion of time spent with consonance by individual subjects ranged from 45.46% to 52.92%). The infants, in contrast, showed a significant tendency to spend more time on the consonant side (Fig. 5b; $t(31) = 2.19$, $p < 0.05$). This effect was larger in some subjects than in others, and in one subject was not apparent, with the four animals individually spending 50.27%, 53.22%, 55.32%, and 61.91% respectively. The data suggest that at least some of the infants internalized the exposure stimulus, and preferred it to a comparable dissonant version.

Experiment 4: Consonant Exposure Stimulus vs. Noise Control

To replicate the preference for the exposure stimulus, we sought to compare it to a second unfamiliar stimulus. We chose a random sequence of white noise bursts that were faded in and out to match the temporal parameters of the consonant exposure stimulus, and low-pass filtered to mimic some of the spectral variation.

Methods

The cutoff frequency for each noise burst was randomly set to one of 9 values ranging between 550 and 8820 Hz. We tested 4 of our adult subjects and 5 of the infants, 3 of the older group and the two in the young group, with 8 sessions per subject.

Results

We had not tested marmoset subjects on this stimulus contrast prior to beginning the exposure. To ensure that any preference was not due to a preexisting bias, we tested 3 adult tamarin subjects. In previous work we found that tamarins and marmosets exhibited similar preferences in every case tested with this method (McDermott & Hauser, 2006), suggesting that tamarins constitute a reasonable control group for our purposes. As expected, the tamarins failed to show a difference in time spent with the exposure stimulus compared to that spent with the noise control, both individually and as a group ($t(23) = -0.22$, $p = 0.83$; the proportion of time spent with consonance ranged from 46.07% to 52.03% for individual subjects.

We then tested adult and infant marmosets with the same stimuli. As shown in Fig. 6, the adult marmosets again failed to exhibit a preference for the exposure stimulus (t(31) = 0.04, p = 0.97; proportion of time spent with exposure stimulus by individuals ranged from 46.19% to 51.62%). In contrast, however, the infants consistently spent more time on the side of the exposure stimulus than on the side of the noise control (t(39) = 5.05, p < 10^-4), with the proportion of time spent with the exposure stimulus ranging from 59.32% to 63.84% in individual subjects. The rank ordering of the animals according to the proportion of time spent with the exposure stimulus revealed that the two animals who spent the most time with the exposure stimulus in Experiment 3 again did so in the present experiment. No difference was apparent in the strength of the effect in the three infants who were born a month before the exposure's start compared to the two infants who were born four months in and who therefore likely received in utero exposure (63.31% and 62.41% of time spent with exposure stimulus, respectively; t(38) = -0.17, p=.86).

It thus appears that infants, but not adults, preferred the musical stimulus to which they were exposed over unfamiliar comparison stimuli. This preference seems to be fairly specific to the exposure stimulus, as it did not generalize to the consonant test stimuli of Experiment 2 that were distinct from the exposure stimulus, even though the sounds composing the test stimuli were embedded in the exposure stimulus. The results of Experiment 3 suggest that the infants internalized the particular intervals used in the exposure stimulus, as they tended to prefer this, albeit weakly, to a dissonant stimulus that was otherwise similar. However, it would appear that the infant monkeys also learned other features of the exposure stimulus that were not present in the test stimuli of Experiment 2 (e.g. the pitch range, or the rhythmic structure), and without these additional features present in the test stimuli, consonance was not preferable to dissonance. The hypothesis that the infants internalized other features of the exposure stimulus in addition to the interval structure might also explain why the preference for the exposure stimulus over the noise control was larger and more robust than the preference over the dissonant stimulus, as the noise control differs in more respects from the exposure stimulus than did the dissonant test stimuli of Experiment 3.

Experiment 5: Consonant Exposure Stimulus vs. Silence

Given that the infant monkeys seemed to prefer the consonant exposure stimulus to unfamiliar alternatives, the question was raised as to whether they would in any sense "like" the exposure stimulus. In a previous study we reported that marmosets and tamarins both avoid musical stimuli when given a choice between such stimuli and silence, suggesting they find such stimuli relatively aversive. To test whether the exposure stimulus would produce different results, we presented our adult and infant subjects with a choice between the exposure stimulus and silence.

Methods

We ran 3 of our adult subjects and 3 of the infant subjects in 8 sessions each.

Results

Both the adults and infants spent more time with silence than with the exposure stimulus (adults: t(23) = -2.86, p = .009; infants: t(23) = -2.26, p = .03). There was an insignificant trend for the infant monkeys to spend more time with the exposure stimulus than the adults (43.38%, SE = 2.93, vs. 38.75%, SE = 3.93, respectively), consistent with the idea that the exposure had rendered the stimulus less aversive than it was to the adults. The infant subjects individually spent 39.47%, 43.02%, and 50.06% of the time on the side of the exposure stimulus. The subject who spent the most time with the exposure stimulus also showed the largest preferences in Experiments 3 and 4. Overall, however, the results suggest that even extensive exposure was insufficient to completely reverse the aversive response to musical stimuli.

Phase 2: Mozart Exposure

Exposure to nonstop consonance induced preferences for the particular stimulus infant marmosets were exposed to, but did not seem to induce general preferences for consonance over dissonance. In a second attempt to induce consonance preferences, we exposed the same colony of marmosets to a set of Mozart concertos, reasoning that the variability and structure of paradigmatic Western music might help to produce a more general preference. Exposure to the Mozart concertos began immediately following the conclusion of the consonance exposure. The colony was exposed for 6 weeks prior to testing, which occurred over the following two weeks, during which the exposure continued in the homeroom.

We conducted two experiments, one to test whether the infant monkeys had again internalized the exposure stimulus, and a second to test whether they could generalize to other consonant musical stimuli.

Experiment 6: Mozart vs. Atonal Control

To test whether the Mozart exposure stimulus had induced a mere exposure effect, we presented the monkeys with a choice between a Mozart concerto and a matched atonal piece of music, which lacked the consonance of the Mozart piece and most other aspects of its musical structure.

Methods

The Mozart concerto was the second movement of K453, one of those used in the exposure stimulus. The atonal stimulus was generated by perturbing the notes of the Mozart concerto such that they were drawn with equal probability from a whole note scale. This manipulation destroyed the sense of tonality, and replaced most of the consonant intervals with dissonant ones, but left the overall pitch range intact. The note onset times were also randomly perturbed by an amount drawn uniformly between 0 and 1 beats, to remove the metrical structure. The atonal stimulus was intended to differ from the Mozart stimulus in many ways, while maintaining the pitch range and overall sound level.

Because some of the animals were no longer available at the time of testing, the post-exposure tests were conducted on 2 adults and 3 infants (the 2 of the older group of infants that spent the largest amounts of time with the exposure stimulus in Experiments 3 and 4, and 1 of the younger group) out of the original subject pool. To maintain power with these smaller numbers, each animal was run in 10 sessions. The order of Experiments 6 and 7 was counterbalanced across the animals who participated, such that 2 of the infants and 1 of the adults participated in Experiment 6 first.

Results

To ensure that any preference for Mozart depended on the exposure, we tested 4 of the 5 adult marmoset subjects with the same stimuli prior to beginning the Mozart exposure phase. This test occurred 14 weeks after the start of the consonance exposure phase. Despite the stimulus differences in consonance, scale structure and meter, the marmosets showed no preference for the Mozart concerto over the atonal control stimulus ($t(32) = 0.003$, $p = 0.998$; proportion time spent with Mozart by individuals ranged from 46.83% to 53.28%; 8 sessions each).

As shown in Fig. 7a, this preference was unchanged in the adult marmosets following 6 weeks of exposure to Mozart ($t(19) = 0.39$, $p = 0.70$). The two individuals spent 48.33% and 55.83% of the time with the Mozart; neither was close to significance ($t(9) = -0.31$, $p = .77$; $t(9) = 0.63$, $p = .55$). The infants, in contrast, all spent more time with the familiar Mozart concerto compared to the atonal control stimulus (Fig. 7b; $t(29) = 5.0$, p <10^-4). Each individual showed a significant effect, spending 55.71%, 58.67%, and 67.63% of the time with the Mozart, respectively ($t(9) = 2.34$, $p = .04$; $t(9) = 2.57$, $p = .03$; $t(9) = 4.14$, $p = .0025$).

Our results thus suggest that the Mozart exposure, like the consonance exposure before it, induced a preference for the exposure stimulus in infant monkeys, but not adults. The pattern of results here and in Experiments 3 and 4 is suggestive of a critical or sensitive period for the acquisition of auditory material. This experiment was not intended to reveal what subjects learned about the exposure stimulus with any precision, and the preference could be due to any of the many differences between the stimuli, ranging from general properties such as consonance, tonality, and temporal regularity to the particular features of the Mozart concerto we tested.

Experiment 7: Consonant vs. Dissonant Melodies

Given that the Mozart exposure apparently induced a preference for Mozart in the infants, we wondered whether it might have induced a more general preference for consonant musical stimuli. In the interest of maximizing the chances of getting an effect, we used stimuli that had previously been used to demonstrate consonance preferences in human infants (Zentner & Kagan, 1996, 1998). These stimuli were folk melodies composed in counterpoint, and thus contained some of the musical structure found in classical music, while consisting exclusively of consonant or dissonant intervals.

The stimuli were sequences of the two melodies used by Zentner and Kagan to test for consonance preferences in young infants – European folk songs played in either major and minor thirds (the consonant stimulus) or minor seconds (the dissonant stimulus) on a synthesizer. The consonant and dissonant versions had the same upper voice. Each melody was 35 sec long. The melodies were played repeatedly in random order with the constraint that a given melody did not repeat more than once in a row. We ran 4 adults and 4 infants in 8 sessions each.

Neither the adults nor the infants showed any tendency to spend more time with the consonant melodies compared to the dissonant ones (Fig. 8; adult: $t(31) = 0.43$, $p = 0.67$, proportion of time spent by individuals with consonant melodies ranged from 46.28% to 54.12%; infant: $t(31) = -0.49$, $p = 0.63$, proportion of time spent by individuals with consonant melodies ranged from 43.76% to 54.25%). To help ensure that the animals had not somehow habituated to the stimuli through repeated exposure, we separately analyzed the first two and first four sessions per animal. In neither case was there a trend to spend more time with the consonant versions of the melodies (first two sessions: 47.31%, SE = 5.1%; first four sessions: 46.35%, SE = 4.47%). Despite the various musical similarities between the test stimuli and the Mozart concertos, the animals did not exhibit any preference for the consonant over the dissonant melodies. This suggests that even massive exposure to paradigmatic Western music is insufficient to induce a generalized consonance preference in our marmoset population.

Discussion

The goal of this paper was to examine the effects of music exposure on a colony of nonhuman primates, with the specific purpose of testing whether music exposure might induce preferences for consonance over dissonance. We conducted two periods of exposure followed by various preference tests. In the first period the monkeys were exposed to 12 hours a day of consonant intervals; in the second period they heard Mozart concertos. Our findings suggest two main conclusions. First, infant marmosets are far more likely to exhibit mere exposure effects for auditory stimuli than are adult marmosets. In three experiments we found that infant monkeys consistently chose to hear familiar stimuli over unfamiliar, whereas the adults were just as consistently indifferent. Second, the preferences induced are apparently fairly specific to the exposure stimuli. We found that neither the consonance nor the Mozart concertos produced a general preference for consonance.

The presence of the effects in infants but not adults suggests a critical or sensitive period for the acquisition of auditory material. Such critical periods for learning have been previously documented behaviorally in other species, notably songbirds (Marler, 1970), and the various species that exhibit imprinting (Hess, 1973). There is also considerable neurophysiological evidence for increased auditory plasticity in cats and rats during infancy (Nakahara et al., 2004; L. I. Zhang, Bao, & Merzenich, 2001). Critical periods are also well known in language acquisition (Bickerton, 1975; Curtiss, 1977; Johnson & Newport, 1989; Newport, 1990), and there is some evidence that they exist in the domain of music as well (Hannon & Trehub, 2005b). To our knowledge the present results are the first evidence that critical periods exist in nonhuman primates for auditory learning. It remains to be seen whether the effects observed in infant marmosets are homologous to the critical periods observed in human language acquisition, and whether they are specific to auditory stimuli. The effects of music exposure that have been observed in human infants occur with relatively small amounts of exposure (e.g. 20 minutes a day for 2 weeks) (Hannon & Trehub, 2005b; Saffran et al., 2000), and it would be interesting for future studies to test whether nonhuman primates will exhibit exposure effects with similarly minimal stimulation.

The one prior study on mere exposure effects for music in animals that we are aware of was conducted with newborn rats (H. A. Cross et al., 1967), and although the authors did not also test adult rats for comparison, their results are thus consistent with ours. The rats in that study were exposed to Mozart pieces for an extended period of time, after which they were found to prefer unfamiliar Mozart pieces to unfamiliar Schoenberg pieces. We were unable to test whether our exposure induced preferences of any generality in our monkey subjects, as their imminent departure from the host lab limited the number of tests we could conduct. However, our results suggest a method to probe representations of auditory stimuli in animals, by exposing them to a stimulus and then probing the stimulus representation with preference measurements. It would be of great interest to test the extent to which the preference for the familiar stimulus might generalize to other examples from the same genre, transpositions to different pitch ranges, timbre changes, and so forth.

Although the apparent presence of a sensitive period in marmosets would suggest at least the possibility of rudimentary similarities between the acquisition of auditory knowledge in nonhuman and human primates, the failure of even massive exposure to induce a consonance preference suggests a fundamental difference. We have every reason to believe that the consonant and dissonant test stimuli were readily discriminable to the animals, most obviously because they preferred a consonant stimulus that shared the statistical properties of the exposure stimulus over a comparable dissonant stimulus, suggesting that they could perceive the difference between the consonant and dissonant intervals. That animals can discriminate consonance from dissonance is further supported by neurophysiological studies in other nonhuman primates indicating that consonant and dissonant intervals produce distinct patterns of activity in the auditory system (Fishman et al., 2001; Tramo et al., 2001), and by behavioral studies in Japanese macaques (Izumi, 2000) and songbirds (Hulse et al., 1995) suggesting discrimination of consonance from dissonance. It would thus appear that the failure to induce consonance preferences does not reflect an inability to discriminate the test stimuli. Another possibility is that the massive amounts of exposure experienced by our animals caused them to overlearn the specific stimuli they were exposed to, impairing generalization. Future studies with briefer and/or more varied exposure stimulation could help to clarify this issue. However, the previously reported finding that pet dogs apparently lack a preference for consonance (McDermott, 2006) suggests that even highly varied exposure to music (as varied as the tastes of dog owners) in positive social settings is not enough. Rather, we would suggest that there is a species difference that accounts for the failure of marmosets to acquire a preference for consonance.

Given that the origins of the consonance preference in humans remain unclear, at present there are two obvious possibilities. The first is that some aspect of the human preference for consonance over dissonance is innate, and that animals fail to exhibit a similar preference because they lack this innate component. The second is that infants acquire their consonance preference from exposure to music, but do so with the aid of acquisition mechanisms that cause them to encode musical structure in ways that animals do not. Such acquisition mechanisms might merely involve a predisposition to attend to music. It is well known that infants are positively affected by music (Sandra E. Trehub, 2003), but the same seems not to be true of nonhuman primates, at least those we have tested thus far (McDermott & Hauser, 2006). An innate interest in music, coupled with the sensitive period mere exposure effects documented here, might produce the preferences observed experimentally in human infants. Either way, our results are consistent with the possibility of uniquely human music-related mechanisms that could conceivably have evolved to enable the appreciation, acquisition, and practice of music.

Acknowledgments

The authors thank Kara Furman and Chris Ludwick for help running the experiments.

References

Bendor, D. and X. Wang (2005). "The neuronal representation of pitch in primate auditory cortex." Nature **426**: 1161-1165.

Bickerton, D. (1975). Dynamics of a creole system. Cambridge, Cambridge University Press.

Butler, J. W. and P. G. Daston (1968). "Musical consonance as musical preference: A cross-cultural study." Journal of General Psychology **79**: 129-142.

Chomsky, N. (1986). Knowledge of language : its nature, origins, and use. New York, Praeger.

Cross, H. A., C. G. Holcomb, et al. (1967). "Imprinting or exposure learning in rats given early auditory stimulation." Psychonomic Science **7**: 233-234.

Curtiss, S. (1977). Genie: A psycholinguistic study of a modern day "wild child". New York, Academic Press.

Dahlhaus, C. and R. O. Gjerdingen (1990). Studies in the Origin of Harmonic Tonality. Princeton, NJ, Princeton University Press.

Fishman, Y. I., I. O. Volkov, et al. (2001). "Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans." Journal of Neurophysiology **86**(6): 2761-88.

Hannon, E. E. and S. E. Trehub (2005). "Metrical categories in infancy and adulthood." Psychological Science **16**(1): 48-55.

Hannon, E. E. and S. E. Trehub (2005). "Tuning in to musical rhythms: Infants learn more readily than adults." Proceedings of the National Academy of Sciences **102**: 12639-12643.

Hess, E. H. (1973). Imprinting: Early experience and the developmental psychobiology of attachment. New York, Van Nostrand Rheinhold.

Hubel, D. H. and T. N. Weisel (1970). "The period of susceptibility to the physiological effects of unilateral eye closure in kittens." Journal of Physiology **206**: 419-436.

Hulse, S. H., D. J. Bernard, et al. (1995). "Auditory discrimination of chord-based spectral structures by European starlings (Sturnus vulgaris)." Journal of Experimental Psychology: General **124**(4): 409-423.

Huron, D. (1991). "Tonal consonance versus tonal fusion in polyphonic sonorities." Music Perception **9**(2): 135-154.

Izumi, A. (2000). "Japanese monkeys perceive sensory consonance of chords." Journal of the Acoustical Society of America **108**(6): 3073-3078.

Johnson, J. S. and E. L. Newport (1989). "Critical period effects in second-language learning: The influence of maturational state on the acquisition of English as a second language." Cognitive Psychology **21**: 60-90.

Kessler, E. J., C. Hansen, et al. (1984). "Tonal schemata in the perception of music in Bali and in the West." Music Perception **2**(2): 131-165.

Knudsen, E. I. (2004). "Sensitive periods in the development of the brain and behavior." Journal of Cognitive Neuroscience **16**: 1412-1425.

Marler, P. (1970). "A comparative approach to vocal learning: Song learning in white-crowned sparrows." Journal of Comparative Physiological Psychology **71**: 1-25.

McDermott, J. (2006). "Can music exposure induce preferences for consonance? Evidence from pet dogs." Submitted.

McDermott, J. and M. Hauser (2004). "Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate." Cognition **94**(2): B11-B21.

McDermott, J. and M. D. Hauser (2005). "The origins of music: Innateness, uniqueness, and evolution." Music Perception **23**(1): 29-59.

McDermott, J. and M. D. Hauser (2006). "Nonhuman primates prefer slow tempos but dislike music overall." Cognition **In press**.

Meyer, M. (1903). "Experimental studies in the psychology of music." American Journal of Psychology **14**: 456-478.

Nakahara, H., L. I. Zhang, et al. (2004). "Specialization of primary auditory cortex processing by sound exposure in the "critical period"." Proceedings of the National Academy of Sciences **101**(18): 7170-7174.

Newport, E. L. (1990). "Maturational constraints on language learning." Cognitive Science **14**: 11-28.

Saffran, J. R., M. M. Loman, et al. (2000). "Infant memory for musical experiences." Cognition **77**(1): B15-23.

Seiden, H. R. (1958). Auditory acuity of the marmoset monkey (Hapale jacchus), Princeton University.

Trainor, L. J., C. D. Tsang, et al. (2002). "Preference for sensory consonance in 2- and 4-month-old infants." Music Perception **20**(2): 187-194.

Tramo, M. J., P. A. Cariani, et al. (2001). "Neurobiological foundations for the theory of harmony in Western tonal music." Annals of the New York Academy of Science **930**: 92-116.

Trehub, S. E. (2003). "The developmental origins of musicality." Nature Neuroscience **6**(7): 669-673.

Trehub, S. E., D. Bull, et al. (1984). "Infants' perception of melodies: The role of melodic contour." Child Development **55**(3): 821-830.

Trehub, S. E., A. M. Unyk, et al. (1993). "Adults identify infant-directed music across cultures." Infant Behavior and Development **16**(2): 193-211.

Watanabe, S. and M. Nemoto (1998). "Reinforcing property of music in Java sparrows (Padda oryzivora)." Behavioural Processes **43**(2): 211-218.

Zajonc, R. B. (1968). "Attitudinal effects of mere exposure." Journal of Personality and Social Psychology **9**: 1-27.

Zentner, M. R. and J. Kagan (1996). "Perception of music by infants." Nature **383**(6595): 29.

Zentner, M. R. and J. Kagan (1998). "Infants' perception of consonance and dissonance in music." Infant Behavior and Development **21**(3): 483-492.

Zhang, L. I., S. Bao, et al. (2001). "Persistent and specific influences of early acoustic environments on primary auditory cortex." Nature Neuroscience **4**: 1123-1130.

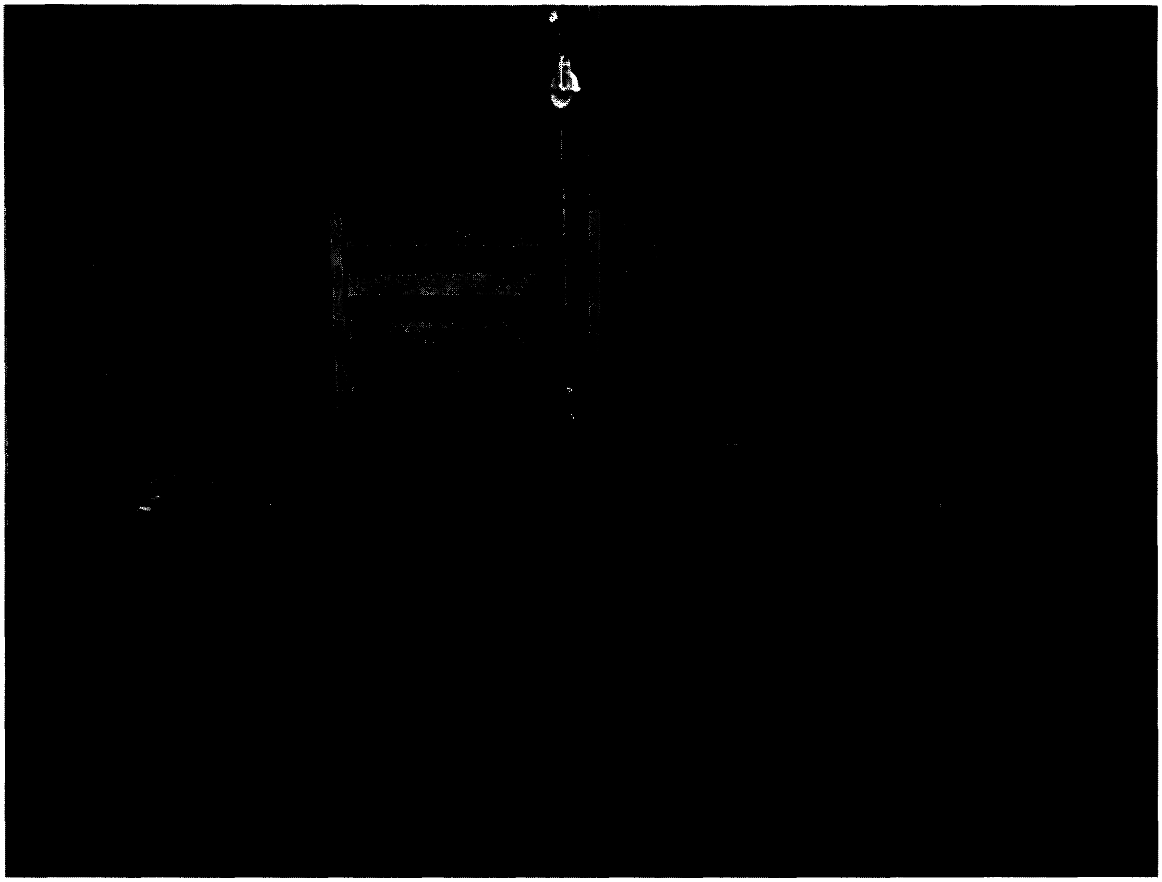

Figure 1. Photo of the apparatus used in the experiments.

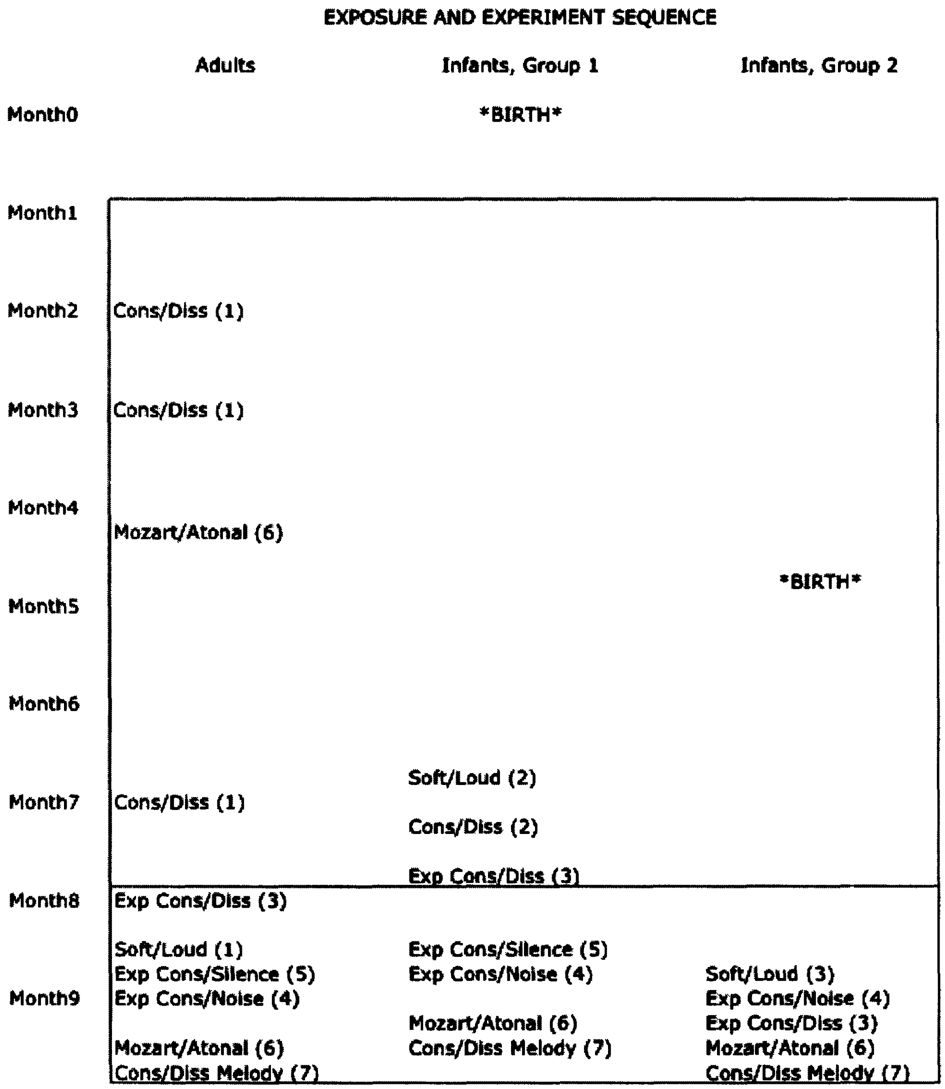

Figure 2. Timeline of exposure phases and experiments for the adult and infant marmosets. The vertical dimension represents time. The top and bottom outlined regions denote the first (consonance) and second (Mozart) exposure phases. The stimuli descriptions denote the experiments conducted on each group of subjects at each point in time. The number in parentheses after the stimuli descriptions is the number the corresponding experiment was assigned in the text.

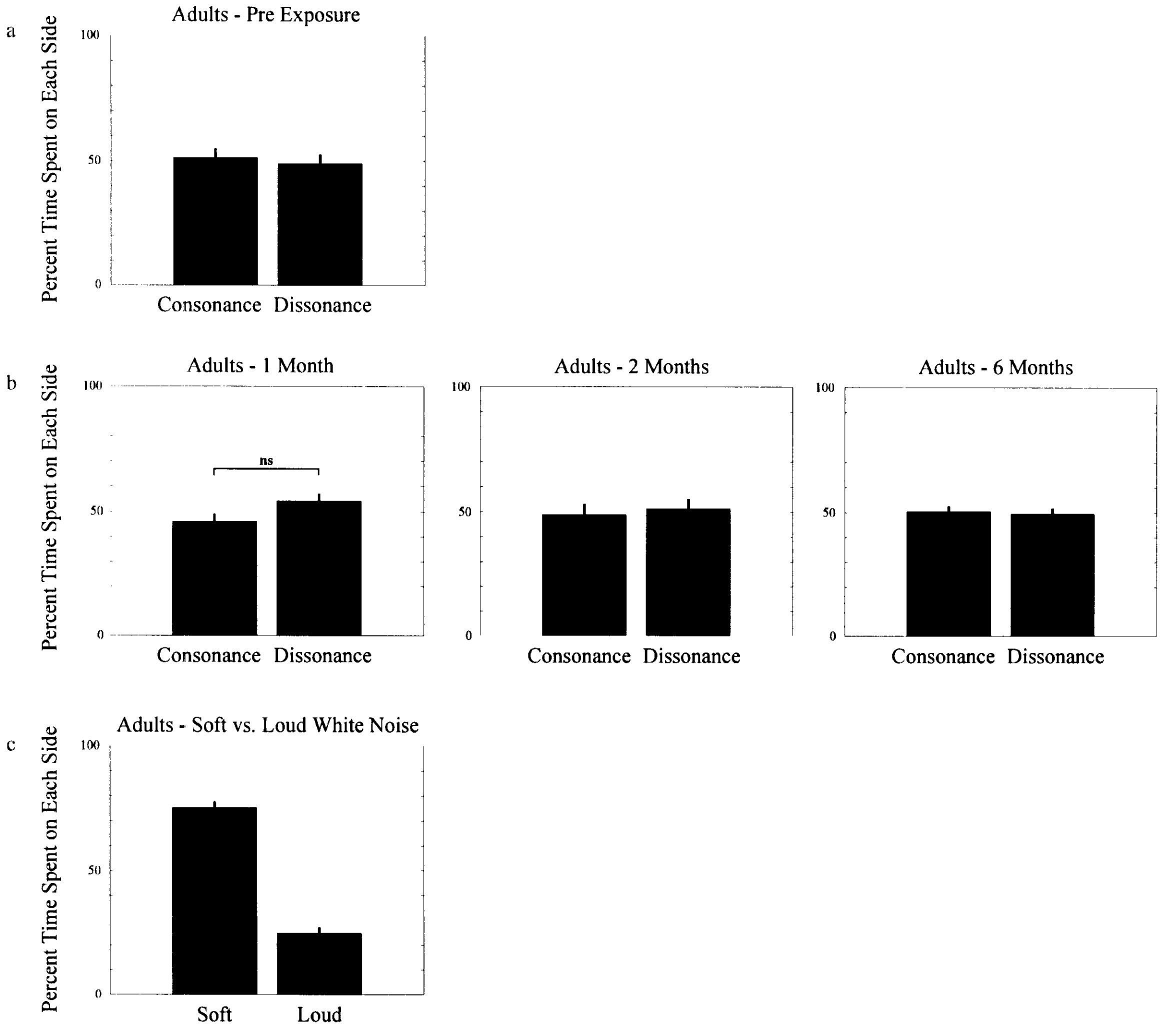

Figure 3. Results of Experiment 1, testing adult marmosets for consonance preferences before and after consonance exposure. The bars plot the average data from the five marmosets run in the experiment. Error bars here and elsewhere denote standard errors. (a) Results from before the exposure. (b) Results from 3 tests at 1 month, 2 months, and 6 months into the exposure period. (c) Results from a test with loud and soft white noise conducted after the last consonance test, to verify that the animals had not habituated to the method.

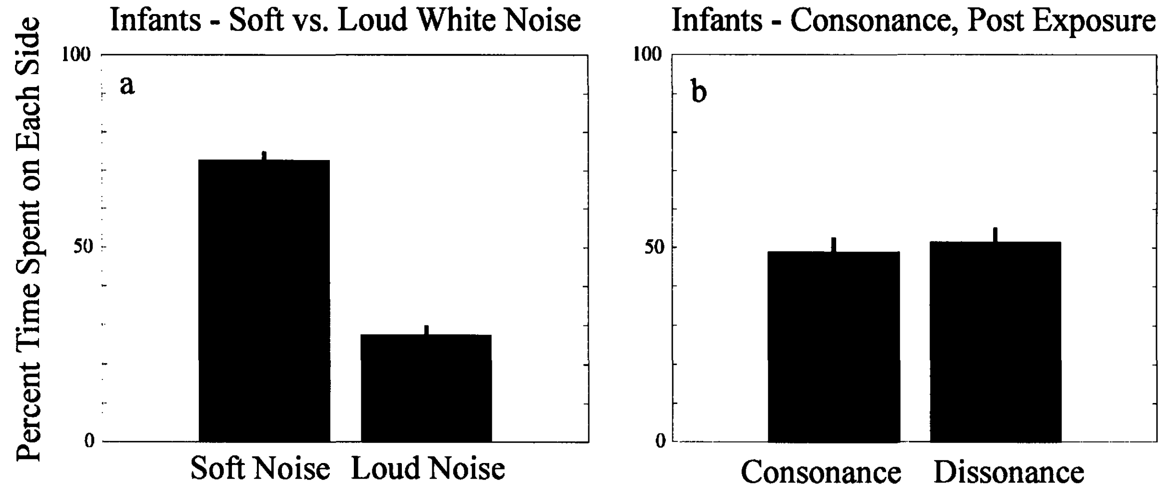

Figure 4. Results of Experiment 2, testing infant marmosets for consonance preferences after consonance exposure. (a) Results for an initial test with loud and soft white noise. (b) Results for tests with consonance and dissonance.

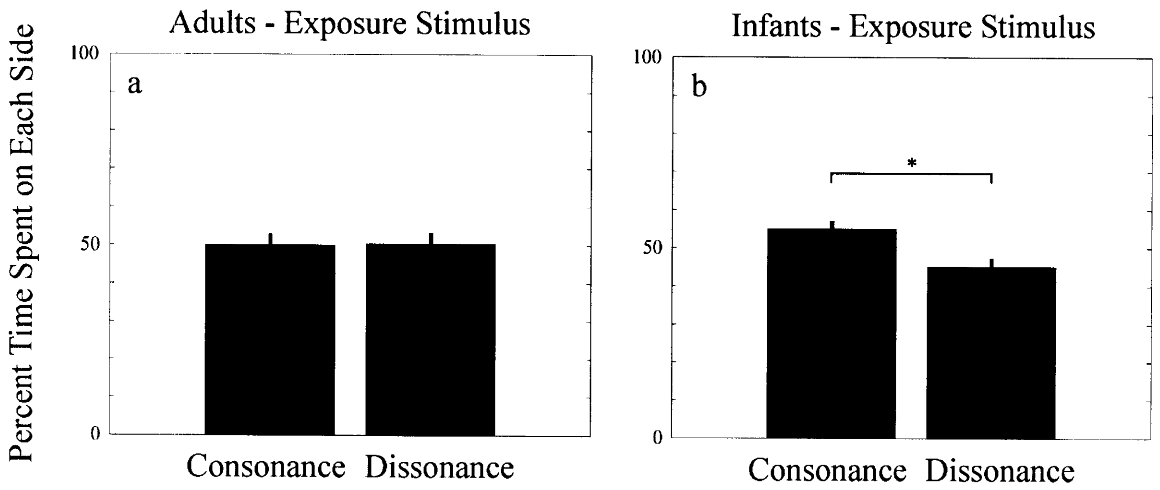

Figure 5. Results of Experiment 3, comparing consonance and dissonance in stimuli with the statistics of the exposure stimulus. (a) Results for adult marmosets. (b) Results for infant marmosets.

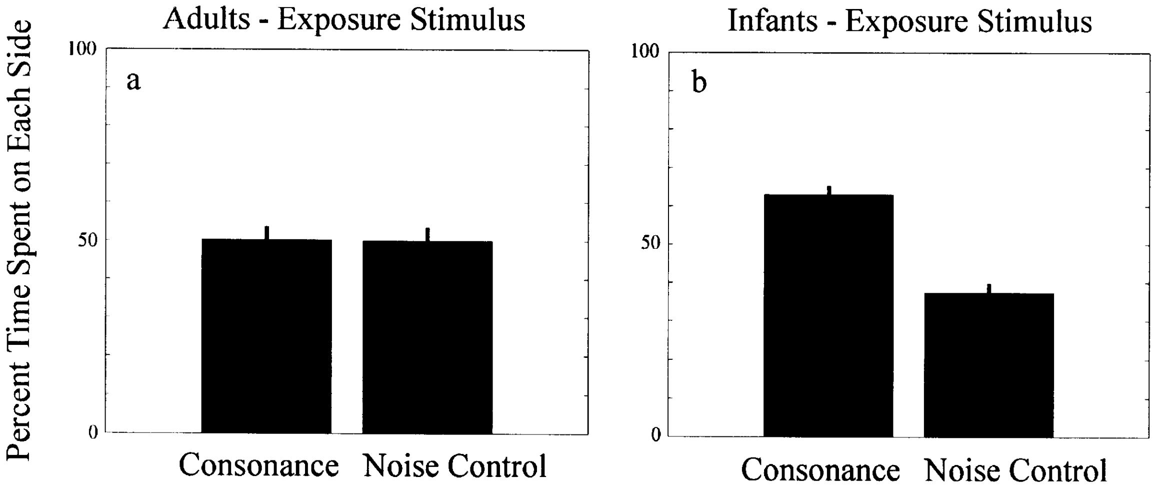

Figure 6. Results of Experiment 4,comparing the consonant exposure stimulus to a noise control with the same temporal structure. (a) Results for adult marmosets. (b) Results for infant marmosets.

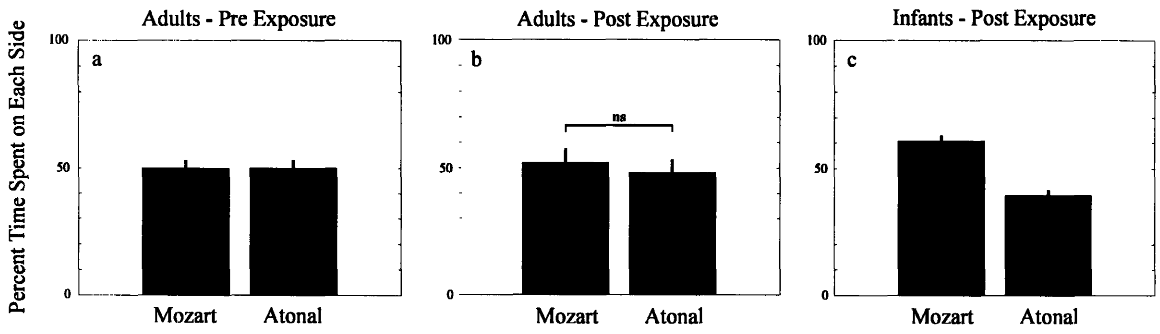

Figure 7. Results of Experiment 6, comparing one of the exposure Mozart concertos with an atonal control stimulus. (a) Results for adult marmosets, before exposure. (b) Results for adult marmosets, after exposure. (c) Results for infant marmosets, after exposure.

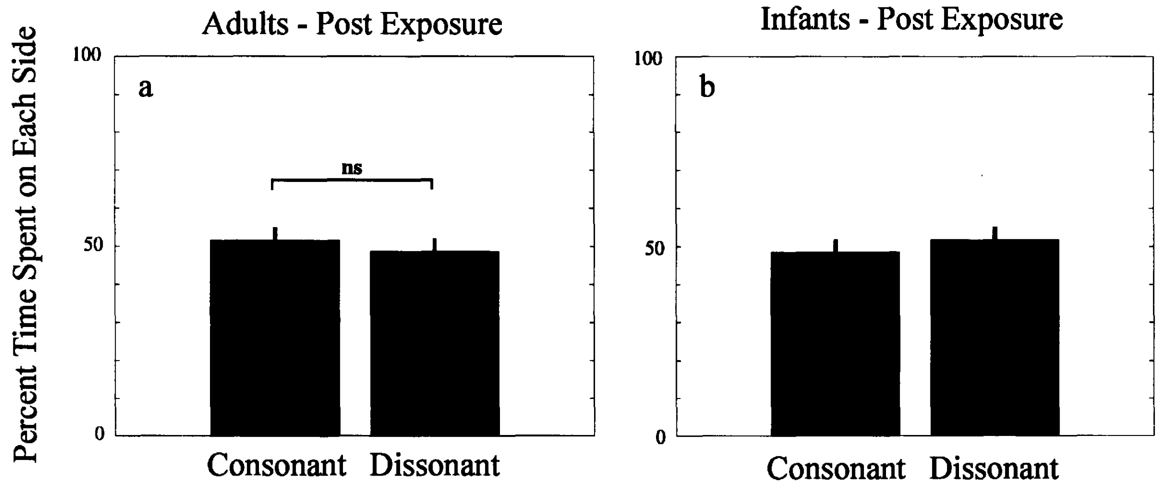

Figure 8. Results of Experiment 7, comparing melodies in consonant and dissonant counterpoint. (a) Results for adult marmosets. (b) Results for infant marmosets.

Chapter 6:

Conclusions

We embarked on the work described in this thesis with the hope of shedding light on music's origins by conducting experiments. We argued that candidates for adaptations selected to enable musical behavior would be found in aspects of music that are innate, uniquely human, and specific to music – innate because adaptations must involve changes to the genetic code, uniquely human because nonhuman animals do not make music, and specific to music because otherwise some other function could have fueled the evolutionary process. Empirical results can push the debate forward, we argued, by providing evidence for or against these three properties of a music-related trait. We have focused on answering the uniqueness question for music-related preferences for sounds, by testing for their presence in nonhuman animals. There were several reasons for this focus. Foremost, the aesthetic response to music is arguably its most important psychological effect, and is therefore of some intrinsic interest. Certain elementary preferences (for consonance over dissonance, for instance) may also underlie more complex phenomena in music cognition, such as the induction and resolution of musical tension. The origins of such preferences thus have broader importance for understanding music perception. An additional advantage to studying preferences lay in the robust methods that existed to test for sound preferences in nonhuman animals (H. A. Cross et al., 1967; Watanabe & Nemoto, 1998); comparable methods for measuring discrimination, for instance, are much more difficult to implement in practice.

Much of the work dealt with one specific preference, that for consonant over dissonant musical intervals. The preference for certain combinations of notes over others has been a subject of fascination since at least the time of the Greeks, and is central to Western music. Several recent studies indicate that the consonance preference is present in very young infants, suggesting it might have an innate basis (Trainor & Heinmiller, 1998; Trainor et al., 2002; Zentner & Kagan, 1996). To clarify its evolutionary origins, we tested for its presence in nonhuman animals. As best we can determine, nonhuman animals lack the consonance preference found in human adults and infants. We tested two species of new World monkeys, cotton-top tamarins and common marmosets, as well as domesticated dogs. Although all three species exhibited robust preferences for other stimulus contrasts using the same method, when given a choice between consonant and dissonant intervals, they were at chance. We replicated this finding several times, with both higher pitched stimuli to match the frequency range of the monkey vocalizations, and tonal and atonal music. In no case do the animals demonstrate a preference for consonance over dissonance. It appears from these studies that nonhuman animals do not have an innate preference for consonance.

Of course, laboratory animals differ from humans not only genetically but also in the exposure they receive to music. All humans are exposed to music to varying degrees from birth, and thus even infant subjects have the benefit of many hours of music exposure. To test whether comparable exposure might induce preferences in nonhuman animals, we first tested pet dogs, who receive exposure to music through many of the same channels that children do, courtesy of their owners. Despite receiving an average of nearly 4 hours of music exposure per day as estimated by their owners, pet dogs failed to prefer consonant over dissonant intervals. We then tested a colony of marmosets monkeys, exposing them to consonant intervals in one phase, and Mozart pieces in another. Although the exposure stimuli were internalized by the infant monkeys and produced a mere exposure effect, these preferences did not generalize to other consonant stimuli.

Given these results, a species difference seems likely to be responsible for the absence of a consonance preference in animals. What might this difference be; why does music exposure fail to induce preferences for consonance in animals? One possibility is that exposure is not at issue - music exposure might not be responsible for the consonance preference found in humans, which could be innate. Alternatively, humans might develop their preference for consonance from exposure to music, but with the aid of acquisition mechanisms that nonhuman animals lack. At this point we have no direct evidence from humans to distinguish between these hypotheses, but one additional feature of the animal response to music makes the latter possibility at least plausible. We repeatedly found that both tamarins and marmosets chose to listen to silence rather than various types of music, including lullabies. These results suggest that nonhuman primates are not particularly attracted to even the most pleasant musical stimuli. In contrast, it is well known that infants have an affinity to music. The positive infant response to music presumably accounts for the universality of infant directed singing (S. Trehub, 2000), and may also underlie infant preferences for singing compared to speech (Nakata & Trehub, 2004). A liking for music may be independent of auditory experience, as congenitally deaf children who receive cochlear implants seem to prefer musical accompaniment to visual stimulation over its absence (Trehub, personal communication). An attraction to music could perhaps account for the hypothetical development of consonance preferences in human infants, as this might speed or otherwise facilitate the extraction of statistical regularities like consonance. Lacking an inclination to attend to music, nonhuman primates might encode different aspects of the same stimuli (for instance, crude pitch statistics). It is also possible that infants possess specific acquisition mechanisms for music above and beyond their interest in it, perhaps akin to those involved in language acquisition. In any case, our results suggest a species difference of some sort.

Although the results of our experiments generally indicate that nonhuman animals respond to music quite differently than do human infants and adults, there are a few potential points of similarity and possible homology. One of these lies in the striking differences we found between the effect of musical stimulation on infant and adult marmosets. Infants consistently chose to listen to familiar over unfamiliar stimuli, whereas adults were indifferent. The exposure occurred in the colony homeroom, which presumably was a positive environment for both age groups, but had very different effects depending on the maturity of the subjects we tested, suggesting the involvement of some sort of critical or sensitive period. The effect is at least superficially similar to sensitive period effects found in humans for music (Hannon & Trehub, 2005b) and language acquisition (Curtiss, 1977; Newport, 1990), raising the possibility that such effects derive in part from domain-general critical or sensitive periods inherited from our nonhuman ancestors.

Also potentially homologous is the effect of tempo. We found that tamarins and marmosets consistently chose to listen to slow over fast tempo stimuli, even when the stimuli did not differ in the average amount of acoustic energy. It is possible that nonhuman primates find fast tempos to be arousing, and that they prefer less arousing stimuli when given the choice. Humans obviously often enjoy arousing music, but the species of monkeys we tested have markedly different temperaments, being much less calm overall, at least as judged informally by us. Future experiments with more direct measures of arousal might help to bolster this interpretation, but at this point the possibility of homologous responses to tempo remains open.

Overall, however, our results indicate considerable divergence between humans and nonhuman animals where music is concerned. The timbral preferences found in humans seem not to be present spontaneously in nonhuman animals or learned from exposure, and the animals we tested apparently find music relatively aversive. Coupled with evidence that animals do not readily hear relative pitch (McDermott & Hauser, 2005), the present results indicate that key aspects of music perception are not shared by nonhuman animals. It would obviously be informative to obtain data from additional species, in particular apes, who are more closely related to humans. It should also be noted that there are numerous musical phenomena that have yet to be studied in animals, notably many interesting components of rhythm, both perceptual and productive. But at present there is a strong possibility of uniquely human perceptual and/or acquisition abilities for music. There is also some evidence suggesting some of these abilities are innate in humans. Our results thus leave open the possibility that some of these abilities evolved to enable the appreciation and production of music.

References

Cross, H. A., C. G. Holcomb, et al. (1967). "Imprinting or exposure learning in rats given early auditory stimulation." Psychonomic Science **7**: 233-234.

Curtiss, S. (1977). Genie: A psycholinguistic study of a modern day "wild child". New York, Academic Press.

Hannon, E. E. and S. E. Trehub (2005). "Tuning in to musical rhythms: Infants learn more readily than adults." Proceedings of the National Academy of Sciences **102**: 12639-12643.

McDermott, J. and M. D. Hauser (2005). "The origins of music: Innateness, uniqueness, and evolution." Music Perception **23**(1): 29-59.

Nakata, T. and S. E. Trehub (2004). "Infants' responsiveness to maternal speech and singing." Infant Behavior and Development **27**(4): 455-464.

Newport, E. L. (1990). "Maturational constraints on language learning." Cognitive Science **14**: 11-28.

Trainor, L. J. and B. M. Heinmiller (1998). "The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance." Infant Behavior and Development **21**(1): 77-88.

Trainor, L. J., C. D. Tsang, et al. (2002). "Preference for sensory consonance in 2- and 4-month-old infants." Music Perception **20**(2): 187-194.

Trehub, S. (2000). Human processing predispositions and musical universals. The origins of music. B. Merker and N. L. Wallin. Cambridge, MA, US, The MIT Press: 427-448.

Watanabe, S. and M. Nemoto (1998). "Reinforcing property of music in Java sparrows (Padda oryzivora)." Behavioural Processes **43**(2): 211-218.

Zentner, M. R. and J. Kagan (1996). "Perception of music by infants." Nature **383**(6595): 29.

Appendix:

Consonance vs. Dissonance With High Pitched Stimuli

Introduction

Although the auditory sensitivity of callitrichids appears to cover roughly the same frequency range as that of humans (Seiden, 1958), their vocalizations are much higher in pitch, generally with fundamental frequencies between 1 and 2 kHz (Cleveland & Snowdon, 1982). It thus seemed conceivable that the animals might be less interested or responsive to lower frequency sounds. Most of our tests for consonance preferences used stimuli in the typical range for human music, with fundamental frequencies ranging from middle C (262 Hz) up to just under 1 kHz. The test stimuli were composed of complex tones with 10 harmonics, and thus contained frequencies well above this range, but the fundamentals were certainly well below the communicative range for these animals. The presence of "pitch" neurons tuned to low frequency fundamentals (Bendor & Wang, 2005) suggests that our stimuli were processed by the auditory system, but it nonetheless seemed prudent to test for consonance preferences with stimuli closer to the vocal range of tamarins and marmosets. We therefore ran three adult tamarins (one male) in an experiment with high pitched consonant and dissonant stimuli.

Methods

The stimuli were similar to the consonant exposure stimuli, which the marmosets exhibited a preference for in Experiments 3 and 4 of Chapter 5. The consonant intervals were again octaves, fifths or fourths, and the dissonant intervals were minor seconds, tritones and minor ninths. The fundamental of the bass note of the intervals ranged between 800 Hz and 1345.4 Hz in one or two semitone steps, and the amplitude spectrum and envelope of the tones was also varied (see Chapter 5 for details). Given that the other note of the interval was as much as 13 semitones higher, these stimuli approximately covered the vocal range of the tamarin. Each animal was run for 8 sessions using the method described in Chapters 2, 3, and 5.

Results

As shown in Figure 1, we again see no sign of a preference for consonance ($t(23) = -0.83$, $p=0.42$). The three individuals each spent 48.53%, 48.68%, and 47.82% of the time with the consonant stimulus. This suggests that the pitch range was not the determining factor in our inability to demonstrate consonance preferences in tamarins and marmosets.

References

Bendor, D. and X. Wang (2005). "The neuronal representation of pitch in primate auditory cortex." Nature **426**: 1161-1165.

Cleveland, J. and C. T. Snowdon (1982). "The complex vocal repertoire of the adult cotton-top tamarin, Saguinas oedipus oedipus." Zeitschrift fur Tierpsychologie **58**: 231-270.

Seiden, H. R. (1958). Auditory acuity of the marmoset monkey (Hapale jacchus), Princeton University.

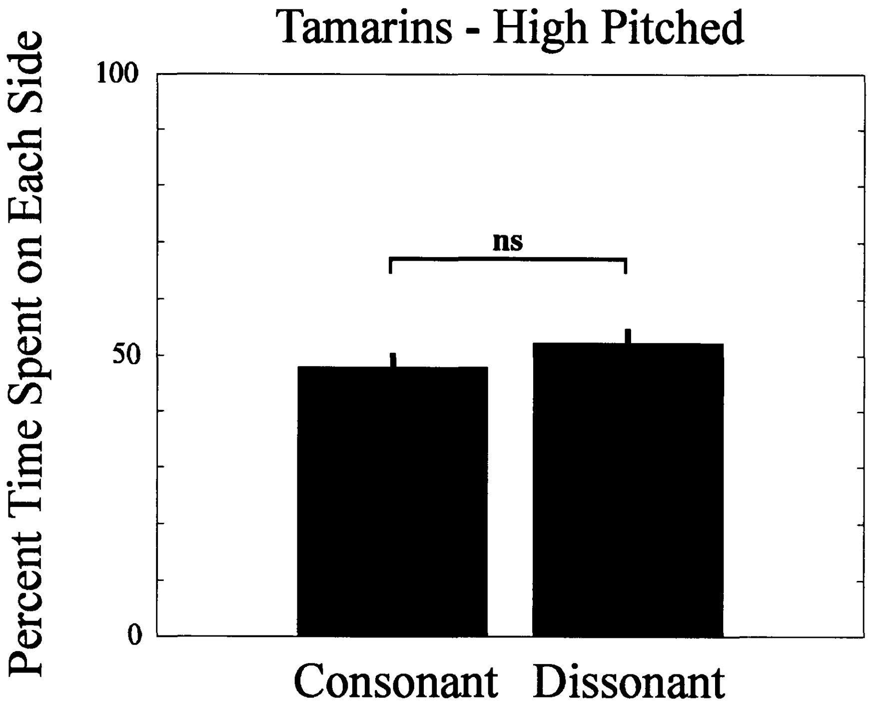

Figure 1. Results of experiment with high-pitched consonant and dissonant stimuli. Error bars represent standard errors.