学位申請論文

泉 明宏

主論文

Thesis

Auditory Perceptual Organization in Japanese Monkeys (Macaca fuscata): Some Properties Concerning Sequential and Harmonic Organization

Akihiro Izumi

Contents

Chapter 1: General Introduction

Sounds are mixed in the air, and perceivers have to resolve a sequence of sounds into elements and reconstruct these elements into streams: sequences of auditory events. Bregman (1990) argued the importance of such a process as auditory scene analysis. This analysis is a process of recovering separate description of each separate thing in the environment. In other words, it is the analysis of auditory world to discrete streams. In this process, it is necessary to select which elements are grouped into the same stream or not. Humans have been revealed to employ many contextual cues for this analysis; e.g., frequency separation, event density, location, and level (Rogers & Bregman, 1993, Rogers & Bregman, 1998, Bregman, 1990). The properties of such processes in humans may be adaptive to their auditory world characterized with spoken language and music. Scene analysis may also be important for non-human animals to perceive their species-specific auditory world. To compare the properties of auditory perceptual organization between humans and evolutional proximate, especially non-human primates, may help understanding the emergence of human language.

Both spoken language and music, which are prominent features of humans, depends on fine process of auditory perception. For example, children diagnosed with language impairment often exhibit a severe deficit in the temporal discrimination of brief time intervals of sounds (e.g., Tallal & Piercy, 1973; Wright et al., 1997). It looks natural to think that these features of humans have been evolved with the evolution of auditory system. Despite the importance of auditory perception in hominization, extensive studies have not been conducted with non-human primates. This may partly because of the difficulty in monkeys to acquire auditory tasks. D 'Amato and Salmon (1982), for example, reported that cebus monkeys required far more trainings than rats to acquire the auditory discrimination.

In the present study, some features of auditory perceptual organization in monkeys were investigated. Psychophysical studies on the analytic aspects of audition have generally shown relatively lower performance in non-humans. For example, frequency discrimination thresholds in non-human primates are typically higher than those in humans (e.g., Stebbins, 1973; Sinnott, Petersen, & Hopp, 1985; Sinnott, Owren, & Petersen, 1987b; Kojima, 1990). Intensity discrimination thresholds in non-human primates are compatible to or somewhat higher than those of humans (e.g., Stebbins, 1973; Clopton, 1972; Sinnott et al., 1985; Kojima, 1990). Monkeys show higher thresholds than humans in tone duration discrimination (Sinnott, Owen, & Petersen, 1987a). In terms of the process of perceptual organization, there have been relatively few investigations.

Japanese monkeys, a species of the subjects in the present study, possess rich variety of vocalizations related to their social behavior (Itani, 1963; Green, 1975). Their vocalizations usually contain rich harmonics and temporal modulation, and it implies monkeys have fine process of perceptual organization adaptive to their complex auditory world. Monkeys were well trained to auditory go/no -go serial discrimination task, and the same monkeys participated through the study.

Sequential organization in the auditory perception of monkeys is investigated in Chapter 2 to 4. Table 1 represents studies concerning sequential organization in human infants and other vertebrates. In this respect, monkeys were shown to encounter difficulty in perceiving frequency contour (Moody, May, Cole, & Stebbins, 1986;

Moody & Stebbins, 1989; D'Amato, 1988). This implies the character of sequential organization in monkeys is somewhat different from that in humans. In Chapter 2, gap-duration discrimination threshold was measured with varying pre-gap and post-gap tone frequency. In this task monkeys were required to integrate information from discrete frequency channels. In Chapter 3, tone-sequence discrimination task similar to the task in Bregman and Rudnicky (1975) with human subjects was applied to monkeys to examine stream segregation. In Chapter 4, melody perception was investigated. Previous studies of melody perception in cebus monkeys concluded monkeys discriminated melodies primarily by absolute pitch rather than by relative pitch. (see D'Amato, 1988). In the present study, stimuli were presented with transposition to make monkeys attend to cues other than absolute pitch.

In chapter 5, perception of harmonic relation was investigated. Table 2 shows some perceptual characteristics observed in human infants and other vertebrates. Although recent studies have revealed that human infants show many similarities to adults in perceiving harmonic relations, there have been few studies with non-humans except studies with perception of missing fundamental. In the present study monkeys were trained to discriminate frequency intervals of chords. The task was anticipated to reveal the effect of sensory consonance in the chord discrimination of monkeys.

Through these studies the author expected to figure out some convergences and divergences between the characteristics of auditory perceptual organization in humans and monkeys. The results were discussed in terms of the phylogeny of human auditory perception.

Chapter 2: The Effect of Marker Frequency Disparity on the Discrimination of Gap Duration in Monkeys

Abstract

Duration discrimination thresholds of the silent interval (gap) between two successive tones (markers) were measured in four Japanese monkeys. The task was serial discrimination, and monkeys were required to release the lever when the gap duration decreased from 200 ms. Monkeys successfully acquired the task, and gap thresholds of monkeys were revealed to be larger than previous data with human subjects. Gap thresholds were not affected by marker frequency when the two markers were identical in frequency, though the thresholds increased when large frequency differences existed between markers. The effect of marker frequency disparity on gap thresholds in monkeys was discussed in terms of the difficulty in integrating information from discrete frequency channels.

2.1 Introduction

There are psychophysical studies that have compared auditory perception in humans and other primates, and these have generally shown relatively lower performance in non-humans. Frequency discrimination thresholds in non-human primates are typically higher than those in humans (e.g., Stebbins, 1973; Sinnott, Petersen, & Hopp, 1985; Sinnott, Owren, & Petersen, 1987b; Kojima, 1990). Sinnott and Brown (1993a, 1993b) provide evidence that there are qualitatively different frequency-analyzing mechanisms in monkeys and humans. Intensity discrimination thresholds in non-human primates are compatible to or somewhat higher than those of humans (e.g., Stebbins, 1973; Clopton, 1972; Sinnott et al., 1985; Kojima, 1990). Monkeys show higher thresholds than humans in tone duration discrimination, and are more sensitive to duration increments than duration decrements (Sinnott, Owen, & Petersen, 1987a). Fragility of auditory short -term memory in monkeys has also been revealed (Kojima, 1985; Colombo & D'Amato, 1986).

In terms of comparative studies of auditory perception, there have been relatively few that have focused on perceptual organization, especially organization of discrete frequency components. Sounds are analyzed into frequency channels in the cochlea, so the process of integrating information from discrete frequency channels is necessary to perceive global patterns, e.g., pattern of frequency modulation or melodic contour. One such study investigated the sensitivity of macaque monkeys to frequency-modulated tones and concluded that monkeys responded to discrete frequency cues rather than to the presence of modulation (Moody, May, Cole, & Stebbins, 1986; Moody & Stebbins, 1989). In another study, cebus monkeys were trained to discriminate tunes (structured sequences of tones) and concluded that performance was controlled primarily by local cues rather than by frequency contours (D'Amato & Salmon, 1982, 1984; D'Amato, 1988). The consensus of these studies on the perception of frequency-modulated tones and tunes is that monkeys seem not to perceive frequency contours, that is, global patterns. This inability of monkeys can be hypothesized that monkeys fail to treat the tone sequence as an integrated pattern . In other words, the whole pattern is split into discrete frequency channels and the global frequency contour was lost.

Bregman (1990) argued the importance of the process of analyzing auditory world to discrete streams, sequences of auditory events, as "auditory scene analysis". In this process, it is necessary to select which elements are grouped into the same stream and which are not. This process in human audition is revealed to be affected by the similarity of component sounds and other contextual cu es (Bregman, 1978, 1990; Rogers & Bregman, 1993, 1998), and the ability of integrating information from frequency channels may underlie this process.

The present study measured the duration discrimination threshold of the silent interval (gap) between two tones (markers). Many studies reported that gap detection or duration discrimination thresholds of humans increase with increments of the frequency difference between the tones (Perrott & Williams, 1971; Williams & Perrott, 1972; Divenyi & Danner, 1977; Divenyi & Sachs, 1978; Grose & Hall, 1996; Phillips, Taylor, Hall, Carr, & Mossop, 1997; Fitzgibbons, Pollatsek, & Thomas, 1974). These increments of thresholds have been discussed as a result of the difficulty in integration. The present study varied the frequency difference between the two tones and investigated effect on gap thresholds. If the increment of frequency difference between the two markers induces an increase in gap threshold, it may be hypothesized as being the result of a difficulty in integrating information from discrete frequency channels.

2.2 Method

2.2.1 Subjects

Subjects were four Japanese monkeys (Macaca fuscata) who were experimentally naive. Two of them were males (I, J; both 4 years old) and the other two were females (H, 5 years old; K, 4 years old). They were housed in individual cages with water freely available. Depending on their performance, they were deprived of food by having their body weights reduced to 90 % of their free-feeding weights. The use of these subjects adhered to the Guide for the care and use of laboratory primates (1986) of the Primate Research Institute, Kyoto University.

2.2.2 Apparatus

Tests took place in a double-walled sound-attenuating chamber (Tracoustics RE-246A). As shown in figure 1, a speaker, a lever, and a feeding box, connected to a universal feeder (Davis Scientific Instruments UF-100), were on one side of the experimental box (50 cm wide x 60 cm deep x 70cm high) within the chamber. A cue lamp (miniature bulb) was mounted above the lever. The behavioral procedure and data collection were controlled by a personal computer (Epson PC-486HX). Two function generators (Leader LFG-1300; Iwatsu SG-4211) generated two pure tones. Each tone passed through an electric switch (Rion SB-04A), an attenuator (Leader LAT-45), and a band-pass filter (NF Electronics E-3201B) set to 1/3 octave above and below the stimulus frequency. Tones were then fed to a mixer (Yamaha KM802), amplified, and presented via the speaker. Auditory stimulus level was calibrated with a sound level meter (Rion NA-80) and a microphone (Rion UC-33P) placed at the position of the subject's head.

2.2.3 Procedure

Behavioral procedure. Discrimination thresholds of gap duration decrements were measured using a serial discrimination task (Fig. 2). When the cue lamp above the lever was turned on, a subject could initiate a trial by holding a lever down. After 0.5 s, the background stimulus was presented variable times (one to four times) at a rate of once per 1.5 s. The background stimulus consisted of two successive tones, called pre- and post-gap markers, separated by a 200-ms gap. Each marker intensity was 60 dB SPL. Marker duration was 200 ms with 5 ms rise-fall time. Following the repetitive presentation of the background stimulus, a target stimulus, which contained a shorter gap (200 ΔT ms) than the background stimulus, was presented once. In each session, consecutive three values of ΔT were chosen from 40, 60, 80, 120, 160, and 200 ms, and they were presented according to the method of constant stimuli. These values were adjusted in order that monkeys may respond correctly in about half of the trials.

Releasing the lever within 1.5 s of target onset was defined as a “hit” and was rewarded immediately with a small piece of sweet potato. A reward was followed by a 5-s inter-trial interval (ITI), during which the cue lamp above the lever was turned off. Continuing to press the lever for more than 1.5 s after the onset of the target stimulus was defined as a “miss”, and a 3-s ITI was delivered after the subject released the lever. Releasing the lever before the presentation of the target stimulus was defined as a “false alarm”, and a 10-s ITI was delivered. These ITIs were chosen to reduce false alarms. Pressing the lever during an ITI was punished with resetting the ITI. A session finished when the subject made 100 correct responses. Each subject had one session in a day.

Stimulus conditions. Thresholds were determined under frequency-identical (Fi) and frequency-different (Fd) conditions shown in Figure 3. In Fi conditions, the frequencies of the pre- and post-gap markers were identical and were set at 500, 1000, or 2000 Hz (stimulus conditions of 500-500, 1000-1000, and 2000-2000, respectively). In Fd conditions, the two markers were different in frequency. The frequency of the pre-gap marker was fixed at 500 Hz, and the frequency of the post-gap marker was 600, 700, 1000, or 2000 Hz (stimulus conditions of 500-600, 500-700, 500-1000, and 500-2000, respectively). Fd conditions were done after Fi conditions were completed. Stimulus conditions were fixed in a session and were varied from session to session. Data were collected for each Fi and Fd condition until the average threshold of the consecutive two sessions was higher than that of the previous two sessions, and thresholds of the final four sessions were used to calculate the final threshold.

Catch trials. To monitor false alarm (FA) rate, 20 % of the trials in a session were catch trials. On these trials, the target stimulus with the largest ΔT in the session was presented after the background stimuli were repeated five times. Releasing the lever during the period of the second to fifth background stimulus in catch trials was recorded as a "false alarm" Session FA rate was calculated in each session as $F / (B + F \quad T)$, where F is the cumulative number of false alarms in catch trials, B is the cumulative number of background stimuli before releasing the lever in catch trials, and T is the cumulative number of catch trials. When FA rate was .15 or higher, responses were regarded as indicating a lack of stimulus control and the data from the session was discarded.

Threshold estimation. To monitor the gross change of threshold, a threshold of a session was considered to be the ΔT corresponding to hit rate of .5 on the psychometric function by linear interpolation. These thresholds were used only for the judgment of the termination of the experiment with a stimulus condition.

For further analysis, repetition number of background stimuli was treated as a separate factor with four conditions. The hit and FA rate for each stimulus condition and background repetition were calculated and transformed to d’ values according to signal detection theory. To avoid infinite d’ (i.e., either 0 or 1 for hit or FA rate), each hit and FA rate was transformed as [(original hit or FA rate) .5] x .99 + .5. A psychometric function was constructed, and a threshold was considered to be the ΔT corresponding to d’ of 1.5 on the function by using least square method.

2.3 Results

Figure 4 represents examples of individual psychometric functions. Figure 5 shows gap duration thresholds in various stimulus conditions and repetition numbers of background stimuli. Gap thresholds of Fi conditions in the present study were comparable to the auditory duration discrimination thresholds of monkeys with standard duration of 200 ms (Sinnott et al. 1987a), and they were above that of humans (e.g., Abel, 1971; Divenyi & Danner, 1977; Penner, 1976).

A two-way analysis of variance with stimulus condition and number of background repetition revealed that there were significant main effects of stimulus condition [$F(6, 18) = 9.88$, $p < .01$] and background repetition [$F(3, 9) = 16.36$, $p < .01$], and there was not significant but weak interaction [$F(18, 54) = 1.78$, $p < .1$]. Tukey’s HSD test revealed that the threshold at the stimulus condition of 500 -2000 was significantly higher than the threshold at the condition of 500-1000 ($p < .05$) and all other conditions ($p < .01$). This means that the thresholds of monkeys were not affected by marker frequency when the two markers were identical in frequency, though the thresholds increased when large frequency differences existed between markers. Threshold in the case of one background stimulus was significantly higher than that of the other background repetitions ($p < .01$). Threshold of four background conditions were also significantly higher than that of two background ($p < .01$) and three background conditions ($p < .05$).

2.4 Discussion

High temporal resolution in humans is thought to be important in processing spoken language. People diagnosed with language impairment often exhibit a severe deficit in the temporal discrimination of brief time intervals (e.g., Tallal & Piercy, 1973; Wright et al., 1997). To identify and discriminate a two-tone sequence correctly, for example, children with language impairment required longer intervals between the two tones than did normal children (Tallal & Piercy, 1973). Compared to human speech, vocalization in monkeys may not require high temporal resolution to perceive. Japanese monkeys are known to possess two types of contact calls, coos, which are distinguished by the temporal position of the frequency inflection from rising to falling (Green, 1975; Zoloth et al., 1979; Beecher, Petersen, Zoloth, Moody, & Stebbins, 1979; May, Moody, & Stebbins, 1988, 1989). Hopp, Sinnott, Owren and Petersen (1992) tested the ability of monkeys to discriminate the temporal position of the frequency peak in coo calls, and the discrimination thresholds were comparable to the gap duration thresholds in the present study. In a study of vocal exchanges between Japanese monkeys, two consecutive coo calls of different callers within short time interval were shown to be similar in their band of frequency modulation, though whole duration or the contours of frequency modulation had no correlation (Sugiura, 1993, 1998; Sugiura & Masataka, 1995).

Threshold in the case of one or four background repetition was higher than that of two or three repetitions. Sensitivity to time interval is known to increases with repetition of the background interval in human subjects (Drake & Botte, 1993), and higher threshold in the condition of one background stimulus in the present study can be explained as a similar phenomenon. Increment of threshold with four-background repetition may due to guess responses, because maximum repetition number of background stimulus was five in catch trials. Monkeys can get a reward with releasing lever after four background stimuli at chance level of .5.

Increments of gap thresholds with the increments of the frequency separation may be interpreted as a result of the difficulty of monkeys in integrating information from two markers when large frequency differences are imposed between them. Monkeys may be able to integrate information from two frequency components when their frequency difference is not so large. Previous studies indicate a larger difference in marker frequency or a shorter standard gap duration than the parameters used in the present experiment are required to cause increments in human gap thresholds (Divenyi & Danner, 1977; Divenyi & Sachs, 1978). In other words, humans may be better able to integrate information across frequency channels. As mentioned earlier, monkeys are shown to possess lower frequency resolution than humans (e.g., Sinnott et al., 1985), so it is interesting that the results of the present study indicate that gap thresholds of monkeys are more affected by frequency differences compared to humans.

Human infants are known to group perceptually sounds according to similarity in pitch as well as adults (Demany, 1982; Thorpe & Trehub, 1989). Such tendency was observed also in European starlings (Braaten & Hulse, 1993; Hulse, MacDougall-Shackleton, & Wisniewski, 1997). It is speculated that many species of vertebrates inherently possess the ability of perceptual integration across frequency channels, and what may be species-specific, however, is the characteristics of this integration process. From the results of the present study, perceptual integration in monkeys seems to be more susceptible to frequency separation than in humans. Compared to monkeys, humans may possess a more flexible ability to both integrate and split frequency components. For humans, a sound containing a rapid change in frequency is usually best to be integrated in perceiving spoken language or music. For monkeys there may be no need to perceive like humans because sounds with large frequency transitions may by usually emanating from discrete sources in their auditory world.

Chapter 3: Auditory Stream Segregation in Monkeys: Capturing Distracter Tones

Abstract

Bregman and Rudnicky (1975) conducted an experiment of auditory sequence discrimination with humans, and showed the effect of frequency proximity on the auditory stream segregation. In the present study, similar task was applied to Japanese monkeys to test whether stream segregation is observed in their audition. Two monkeys acquired the task, and one of them showed qualitatively similar performance as humans.

3.1 Introduction

In auditory perception, sounds received by the ear are analyzed to discrete streams: sequences of auditory events. Bregman (1990) argued the importance of such process as auditory scene analysis. In this process, it is necessary to select which elements are grouped into the same stream and which are not. Humans have been revealed to employ many contextual cues for this analysis, e.g., frequency separation, event density, location and level (Rogers & Bregman, 1993, Rogers & Bregman, 1998, Bregman, 1990).

There have been few studies concerning such perceptual process in non-human animals. For example, Hulse, MacDougall-Shackleton, and Wisniewski (1997) showed that European starlings can detect signals (target songs) embedded in noises (other songs). This implies starlings can perceptually reconstruct target songs from noises as well as humans, but what kind of cues starlings use for the reconstruction is not revealed.

In the previous study examined monkeys' ability to discriminate gap duration between two tones, the performance of monkeys became worse when large frequency difference were imposed between two tones (Izumi, 1999). Many studies have reported similar tendency with humans (Perrott & Williams 1971; Williams & Perrott 1972; Divenyi & Danner 1977; Divenyi & Sachs 1978; Grose & Hall 1996; Phillips, Taylor, Hall, Carr, & Mossop 1997; Fitzgibbons, Pollatsek, & Thomas 1974), and these tendency have chiefly been discussed as a result of the difficulty in integration between frequency channels. It seems true that the ability of integrating between frequency channels has some role in the process of stream segregation, but there are discussions that stream segregation is independent of the effect shown in gap discrimination tasks (Neff, Jesteadt, & Brown, 1982).

Bregman and Rudnicky (1975) tested the streaming effect with tone-sequence discrimination task: discrimination from AB (ascending) to BA (descending). In the discrimination, humans feel difficulty when distracter tones were added before and after the sequence (XABX). Adding extra tones (captors) that are same frequency to the distracter tones (CCCXABXCC) make the discrimination easy. This effect of frequency proximity is discussed as a result of streaming; captor tones captured distracters in the same stream.

In the present study, similar discrimination task was applied to Japanese monkeys. Monkeys were initially trained to discriminate two two-tone sequences (ascending and descending). Transfer tests were then conducted with stimulus sequences containing captor and distracter tones. Transfer tests required monkeys to discriminate the two-tone sequences embedded in captor and distracter tones. If the process of perceptual organization in monkeys employs frequency proximity as in humans, similar effects of distracter and captor tones are expected.

3.2 Method

3.2.1 Subjects

Subjects were three Japanese macaques who were previously used in the auditory gap duration discrimination task (Izumi, 1999). Two of them were females (H, 6 years old; K, 5 years old) and the other was a male (I; 5 years old). They were housed in individual cages with water freely available. Depending on their performance, they were deprived of food by having their body weights reduced to 90 % of their free-feeding weights. The use of these subjects adhered to the Guide for the care and use of laboratory primates (1986) of the Primate Research Institute, Kyoto University.

3.2.2 Apparatus

Tests took place in a double-walled sound-attenuating chamber (Tracoustics, RE-246A). A speaker, a lever, and a feeding box, connected to a universal feeder (Davis Scientific Instruments, UF-100), were on one side of the experimental box (50 cm wide x 60 cm deep x 70cm high) within the chamber. A cue lamp (miniature bulb) was mounted above the lever. Auditory stimulus was generated by a PC/AT compatible personal computer with a sound card (Digidesign, AudiomediaIII). To reduce noise, tones were filtered with low-pass of 6000 Hz and high-pass of 100 Hz (NF Electronics, 3624). They were then amplified and presented via the speaker. Auditory stimulus level was calibrated with a sound level meter (Rion, NA-80) and a microphone (Rion, UC-33P) placed at the position of the subject's head. The computer controlled behavioral procedure and data collection.

3.2.3 Procedure

Behavioral procedure. A go/no-go procedure was used with positive reinforcement operant conditioning. The task was basically tone-sequence discrimination. When the cue lamp above the lever was turned on, a subject could initiate a trial by starting to press the lever. After 0.5 s, the background stimulus was presented three times. Stimulus presentation rate was once per 2.5 s. Background stimuli consisted of two pure tones. The frequency of each tone was 2000 Hz and 2200Hz respectively, and the sequence was ascending; the tone of 2200 Hz followed to the tone of 2000 Hz with no silent interval. The duration of these tones were 75 ms with 10 ms rise-fall time.

Half of the trials were "change" trials, and a target stimulus followed to the background stimuli. The target stimulus was different from background stimulus in the order of tones; sequence of tones was descending. In transfer tests, target stimulus contained distracter and captor tones as shown in figure 6. To monitor false alarm rate, the other half of the trials were "catch" trials. In these trials, catch stimulus was presented after the presentation of backgrounds. The sequence of target tones in this stimulus was ascending as in the background stimulus. As well as target stimulus, catch stimulus contained distracter and captor tones in transfer tests. Following to a catch stimulus, a target stimulus with descending target tones was presented.

In both kinds of trials, subjects were rewarded immediately with a small piece of sweet potato when they release the lever during 2.5 s from the onset of the target stimulus. A reward was followed by an inter-trial interval (ITI) of 3 s. Releasing the lever outside this reward period was mildly punished with prolonged ITI (10 s). During an ITI, the cue lamp above the lever was turned off, and pressing the lever was punished with resetting the ITI. A trial was repeated after the ITI if the subjects released the lever before the presentation of target or catch stimulus.

In "change" trials, releasing the lever during 2.5 s from the onset of target stimulus was defined as a "hit". Continuing to press the lever for more than 2.5 s after the onset of the target stimulus was defined as a "miss". In "catch" trials, releasing the lever during 2.5 s from the onset of the catch stimulus was defined as a "false alarm". Continuing to press the lever for more than 2.5 s after the onset of the catch stimulus was defined as a "correct rejection"

A session finished when the subject completed 100 trials. Each subject had two sessions in a day.

Stimulus conditions. Monkeys were initially trained to discriminate two-tone sequences (AB to BA) for 25 sessions. After the training, three transfer tests were conducted. Each test contained three stimulus conditions shown in table 3. In the transfer test 1, these stimulus conditions were randomly presented in a session. Each subject completed 10 sessions. In the transfer test 2 and 3, only one stimulus condition was presented in a session and conditions were varied in rotation. Each subject completed 10 sessions for each condition.

The duration of each tone was 75 ms in training, transfer test 1 and 2. In transfer test 3, the duration of each target tone was 100 ms, and each distracter and captor tone was 50 ms. Elongated target tones were used to make the discrimination easier, and distracter and captor tones were shortened to avoid the extension of the whole length of stimulus sequences. The distracter tones were 1700 Hz in frequency, and they were presented immediately before and after the target tones. The frequency of a captor tone in condition 2 and 3 was 1700 Hz and 700 Hz, respectively. These distracter and captor tones were presented at temporally regular intervals; an interval between adjacent tones was 150 ms and 100 ms in transfer test 2 and 3, respectively. In stimulus condition 2, the frequency of captors was the same as that of distracters, and captor sequence was expected to absorb distracters.

Calculation of discrimination index. The hit and FA rate for each stimulus condition was individually averaged and transformed to d' values according to signal-detection theory.

3.3 Results

The performance of a monkey (K) did not become above chance level, so she did not proceed to the transfer test. The performances of two monkeys (H, I) are in Figure 7. The trends of the data from these monkeys seemed different, so data was individually analyzed with two-way analysis of variance (stimulus condition x transfer type). Monkey H showed significant main effect of transfer test type [$F(2, 4) = 19.98$, $p < .01$], but the effect of stimulus conditions was not significant. Tukey's HSD test revealed that the performance in transfer test 3 is significantly better than in the other two tests ($p < .01$). The performance of monkey I was significantly affected both by transfer type [$F(2, 4) = 10.15$, $p < .05$] and stimulus conditions [$F(2, 4) = 23.28$, $p < .01$]. The performance of transfer test 1 was significantly worse than that of transfer test 2 ($p < .05$) and 3 ($p < .01$). The performance with condition B was better than condition A ($p < .01$) and condition C ($p < .05$).

3.4 Discussion

The failure of monkey K to acquire the task may because of the difficulty of the task and high chance level. Monkeys could get a reward by releasing the lever at the third or fourth stimuli at chance level of 0.5. Practically, subject K showed a tendency to wait for the fourth stimuli and release the lever.

The effects of the type of transfer tests were somewhat different for two monkeys. It is reasonable to think that the performance became better with longer tone duration for monkey H and fewer stimulus conditions in a session for monkey I. The effect of introducing distracter and captor tones on their performance were also different. In monkey H, these stimulus conditions had no effect. In monkey I, adding captors that are same frequency to the distracter tones ease the discrimination in any types of transfer tests. This tendency in monkey I is in accord with that in humans, and can be explained as a result of stream segregation.

The reason why the performance of monkey H was not affected by distracter and captor tones may be that the monkey used somewhat different strategy for the task. In the study of Bregman and Rudnicky (1975), subjects were instructed to discriminate two-tone sequences regardless of distracter and captor tones. In the present study monkeys were initially trained to discriminate two-tone sequences to make them attend to the sequences. In the following transfer tests, monkeys might feel difficulty in discriminating two-tone sequences embedded in the sequences with distracter and captor tones. Possibly they abandoned to attend to the two-tone sequence, and became to use more global cues. For example, it looks consistent that monkey H attended to four-tone sequence: two target tones and two distracters.

Further research is necessary to confirm the effect of streaming observed in monkeys I. To show clear effects of streaming and to induce monkeys attend to target tones, prolongation of captor sequence may effective. With similar stimulus as in the present study, the magnitude of captor effect in humans is shown to vary with temporal regularity of captor tones (Jones, Kidd, & Wetzel, 1981). Aside with the effect of frequency proximity, it is expected to assess the effect of other cues in the perceptual organization in monkeys.

Chapter 4: Relative Pitch Perception in Monkeys

Abstract

Three Japanese monkeys were trained to detect changes from rising to falling contours of three-tone sequences. Tone sequences were presented serially with transposition, so monkeys were urged to attend cues other than the absolute frequency of a component tone. Probe tests with novel sequences showed that monkeys discriminated by the relative pitch when the frequency ranges of sequences were within the training range; showing similar tendency as in previous studies with birds (e.g. Hulse, Cynx, & Humpal, 1984).

4.1 Introduction

The relations between frequencies of tones (rise or fall) define frequency contours of melodies. Perception of such relations, relative pitch perception, is prominent for humans from early stages of development (Demany & Armand, 1984; Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1987; Chang & Trehub, 1977). Studies have revealed that several species of non-human animals, birds and a dolphin, also perceive relative pitch (avian: Hulse, Cynx, & Humpal, 1984; Hulse & Cynx, 1985; MacDougall-Shackleton & Hulse, 1996; Cynx, 1995; Page, Hulse, & Cynx, 1989; Braaten, Hulse, & Cynx, 1990; Cynx, 1995; Cynx, Hulse, & Polyzois, 1986; dolphin: Ralston & Herman, 1995).

There have been no study demonstrated monkeys, evolutional proximate to humans, perceive relative pitch. D'Amato and Salmon (D'Amato & Salmon, 1982, 1984; D'Amato, 1988) tested perception of melody contour with cebus monkeys, and concluded monkeys discriminated melodies by local cues (absolute pitch of component tones) rather than by frequency contours. Similar in melody perception, macaque monkeys were shown to detect frequency-modulated tones by absolute-pitch cues rather than by the presence of modulation (Moody, May, Cole, & Stebbins, 1986; Moody & Stebbins, 1989). These studies suggest monkeys primarily use absolute cues if possible, but it is still unknown whether monkeys use relative-pitch cues if the uses of absolute cues are restricted.

In the present study, perception of relative pitch in monkeys was investigated with some methodological modification from previous studies. The task was serial discrimination, and monkeys were required to detect changes from rising to falling contours of three-tone sequences. Stimuli were transposed for each presentation to make monkeys attend cues other than the absolute frequency of a component tone. After the initial training, some types of probe tests were conducted to test whether monkeys practically used relative-pitch cues for their discrimination. Many studies with birds reported "frequency range constraint" of relative pitch perception; subjects were able to transfer relative pitch perception to novel sequences only within the absolute frequency range of the training phase (e.g. Cynx, Hulse, & Polyzois, 1986). To test whether such a phenomenon can be seen in monkeys, there were both within-range and outside-range stimuli in the probe tests.

4.2 Method

4.2.1 Subjects

Subjects were three Japanese macaques who were previously used in the auditory gap duration discrimination task (Izumi, 1999) and in some other auditory experiments. Two of them were females (H, 6 years old; K, 5 years old) and the other was a male (I; 5 years old). They were housed in individual cages with water freely available. Depending on their performance, they were deprived of food by having their body weights reduced to 90 % of their free-feeding weights. The use of these subjects adhered to the Guide for the care and use of laboratory primates (1986) of the Primate Research Institute, Kyoto University.

4.2.2 Apparatus

Tests took place in a double-walled sound-attenuating chamber (Tracoustics, RE-246A). A speaker, a lever, and a feeding box, connected to a universal feeder (Davis Scientific Instruments, UF-100), were on one side of the experimental box (50 cm wide x 60 cm deep x 70cm high) within the chamber. Two cue lamps (miniature bulb) were mounted above the lever and the feeding box. The lamp above the feeding box was illuminated throughout a session, and the other lamp above the lever was turned on when a trial can be initiated. Auditory stimulus was generated by a PC/AT compatible personal computer with a sound card (Digidesign, Audiomedia III). Generated stimuli were 16 bit and sampling rate of 22.05 KHz. The stimuli passed through a band-pass filter (NF Electronics, 3624) with low-pass of 6000Hz and high-pass of 300 Hz. They were then amplified and presented via the speaker. Auditory stimulus level was calibrated with a sound level meter (Rion, NA-80) and a microphone (Rion, UC-33P) placed at the position of the subject's head. The computer also controlled the behavioral procedure and data collection with a customized program.

4.2.3 Procedure

Stimulus conditions. Stimuli were sequences of three pure tones. Figure 8 shows schematic representation of stimuli in training and probe tests. The duration of each tone was 200 ms with 10 ms rise-fall time and there was no temporal gap between tones. Tone intensity was approximately 70 dB SPL. Frequencies of tones corresponded to that in equal-tempered scale.

There were two types of sequences: rising and falling. In training phase, both rising and falling sequences consisted of four sequences. The frequency of the middle tone for each rising and falling sequences was 494, 622, 784 and 988 Hz, and these sequences were referred as rising/falling -A, -B, -C, -D, respectively. Interval sizes of first-second and second-third tones were two semitones in each sequence.

In each probe test, a pair of rising and falling sequences (rising- and falling-probe) was introduced along with baseline stimuli that were identical to stimuli in training sessions (rising- and falling-baseline). Middle-tone frequency of a probe stimulus was 1244 Hz, 698 Hz, and 698 Hz in probe test 1, 2, and 3, respectively. Interval sizes of first-second and second-third tones of a sequence in probe test 1 and 2 were two semitones, and that in probe test 3 was four semitones.

Behavioral procedure. A go/no-go procedure was used with positive reinforcement operant conditioning. When the cue lamp above the lever was turned on, a subject could initiate a trial by starting to press the lever. After 0.5 s, the background stimulus was presented two to five times at a rate of once per 1.8 s. The sequences were randomly selected from four sequences (rising-A to -D) for each presentation to make monkeys attend cues other than the absolute frequency of a component tone.

Half of the trials were "change" trials, and a target stimulus followed to the background stimuli. The target stimulus was a falling sequence and was selected from four sequences (falling-A to -D). The other half of the trials were "catch" trials. In these trials, extra one rising sequence (catch stimulus) was presented before the presentation of a falling (target) sequence. A response rate for each falling sequence was calculated from the data in " change" trials, and a rate for each rising sequence was the response rate for each type of catch stimulus. In both types of trials, subjects were rewarded immediately with a small piece of sweet potato when they release the lever during 1.8 s from the onset of the target (falling) sequences. A reward was followed by an inter-trial interval (ITI) of 3 s. Releasing the lever outside the reward period was followed by an ITI of 5 s. During an ITI, the cue lamp above the lever was turned off, and pressing the lever during the interval was punished with resetting the ITI. A trial was repeated after the ITI if the subjects released the lever before the presentation of target or catch stimulus. In training phase, a session finished when the subject completed 96 trials.

Each subject had one or two sessions in a day. Training phase consisted of 15 sessions; former 12 sessions includes amplitude reduction of background stimuli. At the first session of training, amplitude of background stimuli was halved, and the amplitude was gradually increased from session to session. This manipulation was introduced to enable monkeys to use intensity cues for initial discrimination, and monkeys were anticipated to attend to relative pitch with the fading of intensity cues.

Following to 15 training sessions, three types of probe tests were conducted successively. In each probe tests, two probe stimuli were added to baseline stimuli that were identical to stimuli in training sessions. In a probe trial, a probe stimulus followed to the background stimuli. Any types of responses were not rewarded in these trials, and an ITI of 5 s followed to releasing the lever. A probe session was 120 trials (96 baselines and 24 probes), and each probe test consisted of three sessions. Other configurations in probe tests were the same as in training.

4.3 Results

Training. Monkeys successfully acquired tasks in the training phase. Figure 9 shows response rates for each stimulus and the number of background repetition in the last three sessions of training. Monkeys responded to falling sequences more frequently than to rising sequences [$t(2) = 13.39$, $p < .005$], and the general disparity of response rates between rising and falling sequences suggests their discrimination cannot be explained only by the perception of absolute pitch. A two-way analysis of variance (ANOVA; stimulus frequency x number of background repetition) for response rates with falling sequences revealed that there was a significant main effect of background repetitions [$F(3, 6) = 17.74$, $p < .005$]. The effect of stimulus frequency was not significant. The other two-way ANOVA (stimulus frequency x number of background repetition) for rising sequences revealed that there was a significant main effect of background repetitions [$F(3, 6) = 53.14$, $p < .001$]. Interaction between background repetition and stimulus frequency was also significant [$F(9, 18) = 2.62$, $p < .05$]. There was a simple main effect of the number of background repetition with stimulus conditions of rising-C [$F(3, 24) = 15.64$, $p < 0.001$] and D [$F(3, 24) = 9.19$, $p < 0.001$]. It implies discriminations of monkeys were partly affected by absolute pitches of stimuli. The first tone of the falling-D is higher than all first tones of rising sequences, and this can be a cue for discrimination. It looks reasonable to think monkeys partly used this cue for their discrimination, and they tended to respond rising sequences with higher frequencies.

Probe test 1. Figure 10a shows individual response rates for each baseline and probe stimuli in probe test 1. As in training phase, monkeys responded to falling-baseline sequences more frequently than to rising-baseline sequences [$t(2) = 12.62$, $p < .005$]. Response rates with both rising- and falling-probe stimulus were higher than that with rising-baseline [rising: $t(2) = 6.78$, $p < .05$; falling: $t(2) = 7.84$, $p < .01$]. The results means, at least, monkeys did not transfer their relative-pitch perception to the stimulus in new frequency ranges. As in training phase, their tendency to respond stimuli with new frequencies implies monkeys discriminated stimuli partly by the absolute pitch of the stimuli.

Probe test 2. Figure 10b shows individual response rates for each baseline and probe stimuli in probe test 2. Monkeys responded to falling-baseline sequences more frequently than to rising-baseline sequences [$t(2) = 16.78$, $p < .005$]. Response rate with falling-probe stimulus was higher than that with rising-baseline [$t(2) = 11.78$, $p < .005$]. Response rate with rising-probe was not different from that with rising-baseline. These response tendencies to probe stimuli imply monkeys discriminated sequences by the relative pitch.

Probe test 3. Figure 10c shows individual response rates for each baseline and probe stimuli in probe test 3. Monkeys responded to falling-baseline sequences more frequently than to rising-baseline sequences [$t(2) = 19.50$, $p < .005$]. Response rate with falling-probe stimulus was higher than that with rising-baseline [$t(2) = 207.63$, $p < .001$]. Response rate with rising-probe was not different from that with rising-baseline. The interval sizes between tones were enlarged in probe stimuli; nevertheless monkeys were shown to maintain their relative pitch perception.

4.4 Discussion

Together with above-chance performances for baseline stimuli through training and probe tests, performances for novel stimuli in probe test 2 and 3 suggest monkeys discriminated sequences by relative pitch. The results from probe test 1 did not show evidences for relative pitch perception, and it can be discussed as a result of frequency range constraint; monkeys were able to transfer relative pitch perception to novel sequences within the absolute frequency range of the training phase, but such perception did not transfer to sequences outside the trained range. However, frequency range constraint is not observed in studies with starlings by MacDougall-Shackleton and Hulse (1996) and with a dolphin (Ralston & Herman, 1995), and such phenomena can be task-dependent.

The results in the present study were not in accord with previous studies reported cebus monkeys discriminate melodies primarily by local (absolute) cues rather than by relative pitch (D'Amato & Salmon, 1982, 1984; D'Amato, 1988). Most possible reason why monkeys show relative pitch perception in the present study is that monkeys were restricted to use absolute cues by frequency transpositions of background stimuli. In tone-sequence discrimination tests in MacDougall-Shackleton and Hulse (1996), starlings responded to novel probe sequences on the basis of absolute pitch, however, they quickly learned to respond on the basis of relative pitch in transfer test s. Relative pitch perception in starlings and monkeys may appear only when absolute pitch cues were restricted to be used.

The results of the present study extended the universality of relative pitch perception. Then, how such percept works in monkeys' practical communication. Japanese monkeys have been reported to possess two types of contact calls, coos, which are distinguished by the temporal position of the frequency inflection from rising to falling (Green, 1975; Zoloth et al., 1979; Beecher, Petersen , Zoloth, Moody, & Stebbins, 1979; May, Moody, & Stebbins, 1988, 1989). Processing these vocalizations may involve relative pitch perception, but such types of vocalizations seem rare in their communication. For monkeys, most types of vocalizations may be distinguished by absolute cues, and this may relate to robust effects of absolute pitches in laboratory studies.

Chapter 5: Japanese Monkeys Perceive Sensory Consonance of Chords

Abstract

Consonance/dissonance affects human perception of chords from early stages of development (e.g., Schellenberg & Trainor, 1996). To examine whether consonance have some role in audition of non-humans, three Japanese monkeys (Macaca fuscata) were trained to discriminate simultaneous two-tone complexes (chords). The task was serial discrimination (AX procedure) with repetitive presentation of background stimuli. Each tone in a chord comprised of six harmonics, and intervals with complex frequency ratios of fundamental frequency (e.g. frequency ratio of 8:15 in major sevent h) resulted in sensory dissonant. The chords were transposed for each presentation to make monkeys attend cues other than the absolute frequency of a component tone. Monkeys were initially trained to detect changes from consonant (octave) to dissonant (maj or seventh). Following the successful acquisition of the task, transfer tests with novel intervals were conducted. In these transfer tests, the performances with detecting changes from consonant to dissonant (perfect fifth to major seventh; perfect fourth to major seventh) were higher than that with detecting reverse changes. These results suggested that the discriminations of monkeys depended chiefly on the consonance/dissonance of chords.

5.1 Introduction

Simple frequency ratios of tones (e.g., octave of 1: 2) have been show to be perceptually special for humans. Complex tones with simple frequency ratios are perceived more consonant compared to tones with complex ratios. This sense of consonance is referred as sensory consonance, and it has been discussed to derive from amplitude fluctuation (Plomp & Levelt, 1965; Kameoka & Kuriyagawa, 1969a, 1969b). Harmonic components of complex tones with complex frequency ratio tend to fall within an identical critical band, and the fluctuation resulted from the interaction between harmonics is perceived as a sense of dissonance. As well as adults, human infants also perceive sensory consonance (Schellenberg & Trainor, 1996). Further, infants prefer consonant more than dissonant (Zentner & Kagan, 1996, 1998; Trainor & Heinmiller, 1998).

In the absence of sensory consonance, simplicity of frequency ratios still affects human perception. Changes from successive pure tones with simple frequency ratios to those with relatively complex ratios are detected easier than changes from complex to simple ratios (Schellenberg & Trehub, 1994a). Deutsch (1973a, b) found interference of memory for tonal pitch, and the effects generalize across octave. Human children and infants show similar tendency of adults both with simultaneous and sequential intervals with pure tones (Schellenberg & Trehub, 1996;Trainor, 1997; Cohen, Thorpe, & Trehub, 1987; Demany & Armand, 1984). For example, three-month-old infants perceive a melody and its octave-transposition as similar (Demany & Armand, 1984).

There have not been many studies concerning perception of consonance and frequency ratios in non-human animals. Wide varieties of vertebrate species are known to perceive missing fundamental as well as humans (goldfish: Fay, 1970; cat: Heffner & Whitfield, 1976; European starling: Cynx & Shapiro, 1986; rhesus monkey: Tomlinson & Schwarz, 1988). This process of pitch extraction may be relevant to the perception of consonance because both processes include integration across frequency components. Blackwell & Schlosberg (1943) reported octave generalization in rats. This result is not generally accepted mainly because of the high frequency used in the experiment: such perceptual effects in humans lose below these frequency ranges (see Demany & Armand, 1984). Octave generalization was not observed in starlings (Cynx, 1993). Richards, Wolz, and Herman (1984) showed some example of octave transposition in vocal mimicry by a dolphin. In studies investigated perceptual invariance of chord structure, Hulse, Bernard, and Braaten (1995) tentatively suggested that starlings discriminated chords by the sensory consonance of stimuli.

In short, simple frequency ratios are perceptually special for humans include infants whether tones contain rich harmonics or not, whether tones are presented sequentially or simultaneously. For non-human animals, the effect of simple frequency ratio is not clear in any cases.

To systematically examine whether consonance have some role in the perception of non-human animals, the perception of two-tone interval was investigated with Japanese monkeys. Each tone in a stimulus comprised of six harmonics, so intervals with complex frequency ratios of fundamental frequency resulted in sensory dissonant. In Experiment 1, monkeys were initially trained to discriminate two intervals. They were then tested whether they perceive transposed chords (chords with same interval but different in absolute frequency) as similar. Experiment 2 was designed to examine whether the discrimination of chords depended on sensory consonance.

5.2 Experiment 1

Monkeys were initially trained to discriminate octave and major seventh (one-semitone difference). To test whether monkeys learned the chord structure, the discrimination was then transferred to other chords with new frequencies.

5.2.1 Method

5.2.1.1 Subjects

Subjects were three Japanese macaques (Macaca fuscata) who previously participated in the experiment of auditory gap duration discrimination (Izumi, 1999). Two of them were females (H, 7 years old; K, 6 years old) and the other was a male (I; 6 years old). They were housed in individual cages with water freely available. Depending on their performance, they were deprived of food by having their body weights reduced to 90 % of their free-feeding weights. The use of these subjects adhered to the Guide for the care and use of laboratory primates (1986) of the Primate Research Institute, Kyoto University.

5.2.1.2 Apparatus

Tests took place in a double-walled sound-attenuating chamber (Tracoustics, RE-246A). A speaker, a lever, and a feeding box, connected to a universal feeder (Davis Scientific Instruments, UF-100), were on one side of the experimental box (50 cm wide x 60 cm deep x 70cm high) within the chamber. Two cue lamps (miniature bulb) were mounted above the lever and the feeding box. The lamp above the feeding box was illuminated throughout a session, and the other lamp above the lever was turned on when a trial can be initiated. Auditory stimulus was generated by a PC/AT compatible personal computer with a sound card (Digidesign, Audiomedia III). Generated stimuli were 16 bit and sampling rate of 22.05 KHz. The stimuli passed through a band-pass filter (NF Electronics, 3624) with low-pass of 6000Hz and high-pass of 300 Hz. They were then amplified and presented via the speaker. Auditory stimulus level was calibrated with a sound level meter (Rion, NA-80) and a microphone (Rion, UC-33P) placed at the position of the subject's head. The computer also controlled the behavioral procedure and data collection with a customized program.

5.2.1.3 Procedure

Behavioral procedure. A go/no-go procedure was used with positive reinforcement operant conditioning. Figure 11 shows a schematic representation of a trial. When the cue lamp above the lever was turned on, a subject could initiate a trial by starting to press the lever. After 0.5 s, the background stimulus was presented two to five times at a rate of once per 2 s. These stimuli consisted of two simultaneous tones (lower tone and higher tone) comprised of six harmonics of equal intensity. The frequency ratio of the two tones was fixed, though the absolute pitch changed randomly for each presentation. The frequency of the lower tone in each chords were selected from 12 frequencies in the equal-tempered scale (A3-G4#). This transposition reduces the possibility that monkeys discriminate stimuli on the basis of absolute pitch. The duration of the stimulus was 500 ms with 10 ms rise-fall time. Stimulus intensity was approximately 60 dB SPL.

Half of the trials were "change" trials, and a target stimulus followed to the background stimuli. The target stimulus was different from background stimulus both in frequency and frequency ratio of tones. The frequency of the lower tone of a target stimulus was selected from the same frequency set as in the case of backgrounds. To monitor false alarm rate, the other half of the trials were "catch" trials. In these trials, extra one background stimulus (catch stimulus) was presented before the presentation of a target. In both kinds of trials, subjects were rewarded immediately with a small piece of sweet potato when they release the lever during 2 s from the onset of the target stimulus. A reward was followed by an inter-trial interval (ITI) of 3 s. Releasing the lever outside the reward period was mildly punished with prolonged ITI (10 s). During an ITI, the cue lamp above the lever was turned off, and pressing the lever during the interval was punished with resetting the ITI. A trial was repeated after the ITI if the subjects released the lever before the presentation of target or catch stimulus.

In "change" trials, releasing the lever during 2 s from the onset of target stimulus was defined as a "hit" Continuing to press the lever for more than 2 s after the onset of the target stimulus was defined as a "miss". In "catch" trials, releasing the lever during 2 s from the onset of the catch stimulus was defined as a "false alarm". Continuing to press the lever for more than 2 s after the onset of the catch stimulus was defined as a "correct rejection"

In training phase, a session finished when the subject completed 96 trials. In transfer tests, a session was 112 trials. Each subject had two sessions in a day.

Stimulus conditions. In the training phase, subjects were required to detect the change of intervals from octave to minor seventh. This was a change of interval size from 13 to 12 semitones, and consonance decreased. To progress transfer tests Subjects were required to achieve 80 % correct responses in successive two sessions.

Following to the training, a transfer test was done. In this test, chords of background stimuli were identical to that in training phase; the interval was an octave and the frequency of the lower tone in each chords were selected from 12 frequencies

(A3-G4#; 220-415 Hz). As target stimuli, 14 new chords with various intervals, including octave, were introduced (see Table 4). These target stimuli consisted of a lower tone with fixed fundamental frequency (A4; 440 Hz) and a higher tone with various fundamental frequencies (A4 to A5#; 440-932 Hz). If the discrimination depended on absolute frequency of chords, the performances of monkeys were not affected by frequency interval of new targets. If monkeys attended to interval size or consonance of chords, discriminating new chord with octave interval from backgrounds was predicted to be difficult.

Calculation of discrimination index. The hit and false alarm (FA) rate was averaged for each stimulus condition and subject. These values were then transformed to d' values according to signal-detection theory. To avoid infinite d' (i.e. either 0 or 1 for hit or FA rate), each hit and FA rate was transformed as (original value – 0.5) x 0.99 + 0.5.

5.2.2 Results and Discussion

Monkeys successfully acquired the task of interval discrimination. Though monkey H and I required only 4 sessions to criterion, monkey K required 16 sessions. Figure 12 shows the performances of monkeys in the transfer test. A two-way analysis of variance with stimulus condition and number of background repetition revealed that there were significant main effects of stimulus condition [$F(13, 26) = 8.87$, $p < .001$]. The effect of background repetition and interaction was not significant. Tukey's HSD test revealed that the performance with a target of an octave interval (12-semitone interval) was significantly lower than other intervals except unison (7- and 11-semitone intervals: $p < .05$; other intervals: $p < .01$). The performance with a target of unison interval was also lower than target intervals of 1 to 5, 9 and 13 semitones ($p < .01$).

The indices of frequency-ratio simplicity prepared by Schellenberg and Trehub (1994b) are also presented in Figure 12. The index was calculated as $[\log_e (X + Y)]^{-1}$, where X and Y is an integer in a frequency ratio in its simplest form. Higher values of the index correspond to greater simplicity of the frequency ratio and more consonant. While stimulus frequency used in the present study was in equal-tempered scale, a frequency ratio for the calculation of this index was in justly-tuned scale.

The result showed successful transfer of the discrimination to new target intervals except unison and octave intervals. The decrement of performance with octave target means the cue of monkeys to discriminate chords was not absolute frequency of a component or combination of two frequencies in a chord. The reason why unison was difficult to discriminate from octave may be that these intervals are almost in the same spectral pattern.

It was not significant but monkeys showed lower performance around perfect fifth interval (7 semitones). The frequency ratio of perfect fifth interval is 2:3; the chord is a consonant. The reason for this decrement of performance may be that the discrimination of monkeys, to some extent, depended on the consonance of chords. The performance also showed non-significant decrement with 11- and 12-semitone intervals. Because the sizes of these intervals are near to the octave intervals (13 semitones), this tendency can be explained as monkeys attended to the size of intervals. Although the results cannot be concluded, both consonance and interval size may affect the chord discrimination of monkeys.

5.3 Experiment 2

Following to Experiment 1, the same monkeys were examined whether their discriminations of tone intervals were affected by the consonance of chords. In this experiment, two pairs of consonant and dissonant intervals were used as stimuli (perfect fifth and major seventh; perfect fourth and minor second). The consonant and dissonant intervals reversed its role as background and target stimuli in a session; monkeys were required to discriminate change from consonant to dissonant and vice versa. This configuration may enable to see the effect of consonance in the discrimination independent of absolute interval size.

5.3.1 Method

The method used in this experiment was identical to that of experiment 1 except stimulus conditions. There were two pairs of consonant and dissonant chords shown in Table 5. The chords in each pair reversed their role as background and target stimuli in a session; monkeys were required to detect changes from perfect fifth to major seventh, from major seventh to perfect fifth, from perfect fourth to minor second, and from minor second to perfect fourth. All of these four conditions were presented in a session. The lower tone in each background chord was selected from 12 frequencies as in Experiment 1 (A3-G4#; 220-415 Hz), and the lower tone in each target chord was selected from three frequencies (C4#-D4#; 277-311 Hz). These frequencies were selected as all tones in target stimuli were presented in background stimuli. Though the size of interval difference between background and target stimuli was variable in Experiment1, the interval size always changed by three semitones in this experiment. A session finished when the subject completed 96 trials.

5.3.2 Results and Discussion

Figure 13 shows the performances of monkeys in the transfer test. A two-way analysis of variance with stimulus condition and number of background repetition revealed that there were significant main effects of stimulus condition [$F(3, 6) = 8.81$, $p < .05$]. The effect of background repetition and the interaction between stimulus condition and background repetition was not significant. Tukey's WSD test revealed that the performances with changes from perfect fifth to major seventh and from major seventh to perfect fifth were significantly higher than performances in the other two conditions ($p < .05$). In other words, the discrimination of chord changes from consonant to dissonant was easier than the change with reverse order.

The discrimination indices with changes from dissonant to consonant were below 0. This means monkeys tend to respond to background stimuli (dissonant) more frequently than target stimuli (consonant). The strategies of monkeys might be to respond when dissonant chords were presented regardless of the stimulus conditions.

5.4 General Discussion

The results can be concluded that monkeys discriminated chords on the sensory consonance after training. In human adults and infants, changes from successive or simultaneous intervals with simple frequency ratios to those with relatively complex ratios are detected easier than changes from complex to simple ratios (Schellenberg & Trehub, 1994a; Trainor, 1997). In the present study monkeys were initially trained to discriminate intervals of 13 semitones (consonant) to 12 semitones (dissonant), so it is not clear whether monkeys originally discriminate dissonant intervals from consonants easier than the reverse order.

One remaining problems is whether simplicity of frequency ratios without sensory consonance affects auditory perception in monkeys as well as in human adults and infants. This can be answered by studying perception of harmonic intervals composed of pure tones, or perception of melodic intervals with similar experiments as in the present study. Demany and Semal (1990) showed that human adults perceive melodic octave intervals more accurately than harmonic intervals at high frequency ranges of tones. It is interesting to investigate both harmonic and melodic intervals in monkeys. Sinnott and Brown (1993a, b) suggested that monkeys rely on temporal-coding mechanism less than humans do in perceiving lower frequencies. If the perception of frequency ratio depends on such coding system (see Ohgushi, 1978, 1983), frequency ratio simplicity may be difficult to perceive in monkeys.

It may be reasonable to think that sensory consonance has some significance in monkeys' auditory world. Of course monkeys' audition are independent of the knowledge of music. Perceptual organization is a candidate that perception of consonance participates in monkeys. Bregman and Doehring (1984) revealed that for humans simple harmonic relations induce perceptual fusion of two simultaneous pure tones.

Sounds are mixed in the air, and animals may analyze such sounds into suitable groups. For example, European starlings can detect target songs from song complexes (Hulse, MacDougall-Shackleton, & Wisniewski, 1997; Wisniewski & Hulse, 1997). Monkeys may have similar ability to perceive their auditory world. Vervet monkeys, for example, give acoustically different alarm calls to some distinct predators, and each type of calls evokes different and apparently adaptive responses from surrounding monkeys (Seyfarth, Cheney, & Marler, 1980a, 1980b; Struhsaker, 1967). For these monkeys, spectral patterns were found to be most salient to discriminate these calls (Owren, 1990a, b). To detect harmonic components embedded in ambient noises and to perceive these components as a discrete calls, consonance may have some role. The presence of dissonance, or possibly frequency ratio complexity, means the sound emerged from more than two sources. Monkeys may have knowledge that frequency of a harmonic component is a multiple of the fundamental frequency.

Chapter 6: General Discussion

Through the present study, some convergences between auditory perception in humans and monkeys were figured out. In this chapter, I point out some remaining problems to be solved and discuss in terms of the phylogeny of human auditory perception.

In Chapter 2, integration across pitch channels was investigated with gap-duration discrimination task. The gap thresholds of monkeys were affected by the frequency disparity between marker tones, and the results were discussed that monkeys, as well as humans, integrate information across frequency channels if the disparity of frequencies is not large. I tentatively suggested that humans are better able to integrate information across frequency channels than monkeys. Although direct comparison between species is necessary, it is probable that humans have more flexible ability to split or fuse sounds. For humans a sound containing a rapid change in frequency is usually necessary to be treated as a perceptual group in perceiving spoken language or music.

The present study failed to show clear evidences of stream segregation in monkeys (Chapter 3). Recent studies with starlings have demonstrated stream segregation (Braaten & Hulse, 1993; MacDougall-Shackleton, Hulse, Gentner, & White, 1998), and a replication of these studies with monkeys may be possible.

Results in Chapter 4 suggested relative pitch perception in monkeys. The results also showed robust effects of absolute frequency (frequency range constraint) as in studies with birds (e.g. Cynx, Hulse, & Polyzois, 1986). Relative importance of absolute pitch perception may be greater for humans than for birds and monkeys.

In Chapter 5, monkeys were revealed to be able to discriminate chords by the sensory consonance. One of the remaining problems is whether monkeys originally discriminate dissonant intervals from consonants easier than the reverse order. In the present study one direction of discrimination (consonant to dissonant) was initially trained. Such perceptual asymmetry can be studied with training both direction of discrimination equally. For further research, the effect of frequency ratio simplicity without sensory consonance must be investigated in non-humans. Preferences of non-humans to consonant and dissonant are also interesting to compare the results with humans (e.g., Zentner & Kagan, 1996, 1998; Trainor & Heinmiller, 1998).

In the present study, there were data implying differences between species. In Chapter 2, gap thresholds of monkeys seemed more affected by frequency differences compared to humans. As in the studies of starlings (e.g. Cynx, Hulse, & Polyzois, 1986), robust effects of absolute pitch were observed in tone-sequence discrimination; such effect have not been prominent in humans. Aside from these differences, the results in the present study generally prepared evidences that characteristics of audition in monkeys are similar to that in humans and in birds. These commonalities in humans and non-humans are, at least, independent of human language. Then, for what animals process such complex auditory patterns?

Species-specific vocal communication is a possible objective for such powerful processes. Vocalizations of monkeys practically contain rich harmonic components and frequency modulation. Not only for vocal comprehension itself, the objective of non-human audition may be extended processing sounds in their auditory world as Bregman (1990) argued. In this view, the possible reason of using such complex auditory patterns is that such calls are easily detected from ambient noises; complex patterns contain rich cues for the perceptual organization. Sound localization of a monkey becomes accurate with the increment of effective bandwidth: e.g., the magnitude of frequency modulation (Brown, Beecher, Moody, & Stebbins, 1978, 1979, 1980). Vocalizations of animals may effectively prepare cues for the perceptual organization in the species-specific auditory world.

Though it was not investigated in the present study, top-down process [i.e. schema-based segregation and integration referred by Bregman (1990)] may have some role in the perceptual organization of non-human audition. European starlings, for example, can detect target songs from song complexes as humans do (Hulse, MacDougall-Shackleton, & Wisniewski, 1997). It may be true that bottom-up processes, like the processes investigated in the present study, are involved in such percept, but a kind of top-down process may have some role. Studies on the perception of species-specific vocalizations may reveal such processes and enable to discuss the significance of perceptual properties in practical communication.

Although further accumulation of comparable data with humans and other animals is necessary, a speculation on the evolution of human audition may be possible. In the process of hominization, ancient humans might have been necessary to transmit much information via the vocal-auditory pathway. To meet this necessity, vocal system had improved and vocal repertories had increased. For example, Lieberman and his colleagues (e.g., Lieberman, 1968, 1985; Lieberman, Klatt, & Wilson, 1969; Lieberman, Crelin, & Klatt, 1972) suggested the importance of anatomical structure of vocal tract and its motor control in humans to evolve their speech. With the increment of vocal repertories, perceivers had had to detect these sounds embedded in the environment sounds. Ancestor humans are thought to have improved the resolution in their auditory system. As well as the improvement of resolution, ancestor humans might have improved an ability to perceptually organize incoming sound elements. Perceiving complex auditory patterns enabled them to enrich their vocal communication. The increment of vocal repertories, supported by the improvement of vocal-auditory system, might have been a basis for the emergence of language.

Acknowledgments

The author is grateful to Prof. Shozo Kojima for his guidance and encouragement throughout this study. Thanks are also to Sumiharu Nagumo for his technical support, and to other members of the Department of Behavioral and Brain Sciences, Primate Research Institute. This study was supported by Grants-in-Aid for Scientific Research by Ministry of Education, Science, and Culture, Japan.

References

Beecher, M. D., Petersen, M. R., Zoloth, S. R., Moody, D. B., & Stebbins, W. C. (1979). Perception of conspecific vocalizations by Japanese macaques. Brain, Behavior & Evolution, 16, 443-460.

Blackwell, H. R. & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. Journal of Experimental Psychology, 33, 407-419.

Braaten, R. F., & Hulse, S. H. (1993). Perceptual organization of auditory temporal patterns in European starlings (Sturnus vulgaris). Perception & Psychophysics, 54, 567-578.

Braaten, R. F., & Hulse, S. H., & Page, S. C. (1990). Discrimination and classification of rising and nonrising pitch patterns by the European starling (Sturnus vulgaris). Animal Learning & Behavior, 18, 352-364.

Bregman, A. S. (1978). Auditory streaming: Competition among alternative organization. Perception & Psychophysics, 23, 391-398.

Bregman, A. S. (1990) Auditory scene analysis. Cambridge, MA: MIT Press.

Bregman, A. S. & Doehring, P. (1984.) Fusion of simultaneous tonal glides: The role of parallelness and simple frequency relations. Perception & Psychophysics, 36, 251-256.

Bregman, A. S., & Rudnicky, A. I. (1975). Auditory segregation: stream or streams? Journal of Experimental Psychology: Human Perception and performance, 1, 263-267.

Brown, C. H., Beecher, M. D., Moody, D. B., & Stebbins, W. C. (1978). Localization of primate calls by old world monkeys. Science, 201, 753-754.

Brown, C. H., Beecher, M. D., Moody, D. B., & Stebbins, W. C. (1979). Locatability of vocal signals in old world monkeys: Design features for the communication of position. Journal of Comparative & Physiological Psychology, 93, 806-819.

Brown, C. H., Beecher, M. D., Moody, D. B., & Stebbins, W. C. (1980). Localization of noise bands by old world monkeys. Journal of the Acoustical Society of America, 68, 127-132.

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

Clarkson, M. G. & Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. Journal of the Acoustical Society of America, 77, 1521-1528.

Clopton, B. M. (1972). Detection of increments in noise intensity by monkeys. Journal of the Experimental Analysis of Behavior, 17, 473-481.

Cohen, A. J., Thorpe, L. A., & Trehub, S. E. (1987). Infants' perception of musical relations in short transposed tone sequences. Canadian Journal of Psychology, 41, 33-47.

Colombo, M., & D'Amato, M. R. (1986). A comparison of visual and auditory short-term memory in monkeys (Cebus apella). Quarterly Journal of Experimental Psychology, 38B, 425-448.

Cynx, J. (1993). Auditory frequency generalization and a failure to find octave generalization in a songbird, the European starling (Sturnus vulgaris). Journal of Comparative Psychology, 107, 140-146.

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, 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 starling (Sturnus vulgaris). Journal of the Experimental Psychology: Animal Behavior Processes, 12, 394-402.

Cynx, J., & Shapiro, M. (1986). Perception of missing fundamental by a species of songbird (Sturnus vulgaris). Journal of Comparative Psychology, 100, 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, 453-480.

D'Amato, M. R., & Salmon, D. P. (1982) Tune discrimination in monkeys (Cebus apella) and in rats. Animal Learning & Behavior, 10, 126-134.

D'Amato, M. R., & Salmon, D. P. (1984) Processing of complex auditory stimuli (tunes) by rats and monkeys (Cebus apella). Animal Learning & Behavior, 12, 184-194.

Demany, L. (1982). Auditory stream segregation in infancy. Infant Behavior & Development, 5, 261-276.

Demany, L. & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. Journal of the Acoustical Society of America, 76, 57-66.

Demany, L. & Semal, C. (1990). Harmonic and melodic octave templates. Journal of the Acoustical Society of America, 88, 2126-2135.

Deutsch, D. (1973a). Interference in memory between tones adjacent in the musical scale. Journal of the Experimental Psychology, 100, 228-231.

Deutsch, D. (1973b). Octave generalization of specific interference effects in memory for tonal pitch. Perception & Psychophysics, 13, 271-275.

Divenyi, P. L., & Danner, W. F. (1977). Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. Perception & Psychophysics, 21, 125-142.

Divenyi, P L., & Sachs, R. M. (1978). Discrimination of time intervals bounded by tone bursts. Perception & Psychophysics, 24, 429-436.

Drake, C., & Botte, M .C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. Perception & Psychophysics, 54, 277-286.

Fay, R. R. (1970). Auditory frequency discrimina tion in the goldfish (Carassius auratus). Journal of Comparative and Physiological Psychology, 73, 175-180.

Fitzgibbons, P. L., & Gordon-Salant, S. (1994). Age effects on measures of auditory duration discrimination. Journal of Speech & Hearing Research, 37, 662-670.

Fitzgibbons, P. J., Pollatsek, A., & Thomas, I. B. (1974). Detection of temporal gaps within and between perceptual tonal groups. Perception & Psychophysics, 16, 522-528.

Green, S. (1975). Variation of vocal pattern with social situation in the Japanese monkey (Macaca fuscata): A field study. In L. A. Rosenblum (Ed.), Primate behavior (Vol. 4, pp. 1-102). New York: Academic Press.

Grose, J. H., & Hall, J. W. (1996). Perceptual organization of sequential stimuli in listeners with cochlear hearing loss. Journal of Speech & Hearing Research, 39, 1149-1158.

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

Hopp, S. L., Sinnott, J. M., Owren, M. J., & Peters en, M. R. (1992). Differential sensitivity of Japanese macaques (Macaca fuscata) and humans (Homo sapiens) to peak position along synthetic coo call continuum. Journal of Comparative Psychology, 106, 128-136.

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.

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, 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, 38-54.

Hulse, S. H., MacDougall-Shackleton, S. A., & Wisniewski, A. B. (1997). Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (Sturnus vulgaris). Journal of Comparative Psychology, 111, 3-13.

Izumi, A. (1999). The effect of marker frequency disparity on the discrimination of gap duration in monkeys. Perception, 28, 437-444.

Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. Journal of Experimental Psychology: Human Perception and performance, 7, 1059-1073.

Kameoka, A., & Kuriyagawa, M. (1969a). Consonance theory: Part I. Consonance of dyads. Journal of the Acoustical Society of America, 45, 1451-1459.

Kameoka, A., & Kuriyagawa, M. (1969b). Consonance theory: Part II. Consonance of complex tones and its calculation method. Journal of the Acoustical Society of America, 45, 1460-1469.

Kojima, S. (1985). Auditory short-term memory in the Japanese monkey. International Journal of Neuroscience, 25, 255-262.

Kojima, S. (1990). Comparison of auditory functions in the chimpanzee and human. Folia Primatologica, 55, 62-72.

Lieberman, P. (1968). Primate vocalizations and human linguistic ability. Journal of the Acoustical Society of America, 44, 1157-1164.

Lieberman, P. (1985). On the evolution of human syntactic ability. Its pre-adaptive bases--motor control and speech. Journal of Human Evolution, 14, 657-668.

Lieberman, P., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee. American Anthropologist, 74, 287-307.

Lieberman, P., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. Science, 164, 1185-1187.

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

MacDougall-Shackleton, S. A., Hulse, S. H., Gentner, T. Q., & White, W. (1998). Auditory scene analysis by European starlings (Sturnus vulgaris): Perceptual segregation of tone sequences. Journal of the Acoustical Society of America, 103, 3581-3587.

May, B., Moody, D. B., & Stebbins, W. C. (1988). The significant features of Japanese macaque coo sounds: a psychophysical study. Animal Behavior, 36, 1432-1444.

May, B., Moody, D. B., & Stebbins, W. C. (1989). Categorical perception of conspecific communication sounds by Japanese macaques, Macaca fuscata. Journal of the Acoustical Society of America, 85, 837-847.

Moody, D. B., May, B., Cole, D. M., & Stebbins, W. C. (1986). The role of frequency modulation in the perception of complex stimuli by primates. Experimental Biology ,45, 219-232.

Moody, D. B., & Stebbins, W. C. (1989). Salience of frequency modulation in primate communication. In R. J. Dooling, & S. H. Hulse (Eds.), The comparative psychology of audition: Perceiving complex sounds.(pp. 353-376). Hillsdale NJ: Lawrence Erlbaum Associates.

Neff, D. L., Jesteadt, W., & Brown, E. L. (1982). The relation between gap discrimination and auditory stream segregation. Perception & Psychophysics, 31, 493-501.

Ohgushi, K. (1978). On the role of spatial and temporal cues in the perception of the pitch of complex tones. Journal of the Acoustical Society of America, 64, 764-771.

Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon. Journal of the Acoustical Society of America, 73, 1694-1700.

Owren, M. J. (1990a). Acoustic classification of alarm calls by vervet monkeys (Cercopithecus aetiops) and humans (Homo sapiens): I. Natural calls. Journal of Comparative Psychology, 104, 20-28.

Owren, M. J. (1990b). Acoustic classification of alarm calls by vervet monkeys (Cercopithecus aetiops) and humans (Homo sapiens): II. Synthetic calls. Journal of Comparative Psychology, 104, 29-40.

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 the Experimental Psychology: Animal Behavior Processes, 15, 137-146.

Penner, M. J. (1976). The effect of marker variability on the discrimination of temporal intervals. Perception & Psychophysics, 19, 466-469.

Perrott, D. R., & Williams, K. N. (1971). Auditory temporal resolution: Gap detection as a function of interpulse frequency disparity. Psychonomic Science, 25, 73-74.

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

Ralston, J. V., & Herman, L. M. (1995). Perception and generalization of frequency contours by a bottlenose dolphin (Tursiops truncates). Journal of Comparative Psychology, 109, 268-277.

Richards, D. G., Wolz, J. P., & Herman, L. M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, Tursiops truncatus. Journal of Comparative Psychology, 98, 10-28.

Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. Perception & Psychophysics, 53, 179-189.

Rogers, W. L., & Bregman, A. S. (1998). Cumulation of the tendency to segregate auditory streams: Resetting by changes in location and loudness. Perception & Psychophysics, 60, 1216-1227.

Schellenberg, E. G., & Trainor, L. J. (1996). Sensory consonance and the perceptual similarity of complex-tone harmonic intervals: Tests of adult and infant listeners. Journal of the Acoustical Society of America, 100, 3321-3328.

Schellenberg, E. G., & Trehub, S. E. (1994a). Frequency ratios and the discrimination of pure tone sequences. Perception & Psychophysics, 56, 472-478.

Schellenberg, E. G., & Trehub, S. E. (1994b). Frequency ratios and the perception of tone patterns. Psychonomic Bulletin & Review, 1, 191-201.

Schellenberg, E. G., & Trehub, S. E. (1996). Children's discrimination of melodic intervals. Developmental Psychology, 32, 1039-1050.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980a). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. Animal Behavior, 28, 1070-1094.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980b). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. Science, 210, 801-803.

Sinnott, J. M., & Brown, C. H. (1993a). Effects of varying signal and noise levels on pure-tone frequency discrimination in humans and monkeys. Journal of the Acoustical Society of America, 93, 1535-1540.

Sinnott, J. M., & Brown, C. H (1993b). Effects of varying signal duration on pure-tone frequency discrimination in humans and monkeys. Journal of the Acoustical Society of America, 93, 1541-1546.

Sinnott, J. M., Owren, M, J., & Petersen, M. R. (1987a). Auditory duration discrimination in Old World monkeys (Macaca, Cercopithecus) and humans. Journal of the Acoustical Society of America, 82, 465-470.

Sinnott, J. M., Owren, M. J., & Petersen, M. R. (1987b). Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo). Journal of Comparative Psychology, 101, 126-131.

Sinnott, J. M., Petersen, M. R., & Hopp, S. L. (1985). Frequency and intensity discrimination in humans and monkeys. Journal of the Acoustical Society of America, 78, 1977-1985.

Stebbins, W. C. (1973). Hearing of Old World monkeys (Cercopithecinae). American Journal of Physical Anthropology, 38, 357-364.

Struhsaker, T. T. (1967). Auditory communication among vervet monkeys (Cercopithecus aetiops). In S. A. Altmann (Ed.), Social communication among primates (pp. 281-324). Chicago: University of Chicago Press.

Sugiura, H (1993). Temporal and acoustical correlates in vocal exchange of coo calls in Japanese macaques. Behavior, 124, 207-225.

Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. Animal Behavior, 55, 673-687.

Sugiura, H., & Masataka, N. (1995). Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (Macaca fuscata). In E. Zimmermann, J. D. Newman, & U. Jürgens (Eds.), Current topics in primate vocal communication (pp. 121-140). New York: Plenum Press.

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. Nature, 241, 468-469.

Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy Developmental Psychology, 25, 122-127.

Trainor, L. J. (1997). Effect of frequency ratio on infants' and adults' discrimination of simultaneous intervals. Journal of Experimental Psychology: Human Perception and Performance, 23, 1427-1438.

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. Infant Behavior &

Development, 21, 77-88.

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

Trehub, S. E., Thorpe, L. A., & Morrongiello, B. A. (1987). Organization processes in infants' perception of auditory patterns. Child Development, 58, 741-749.

Williams, K. N., & Perrott, D. R. (1972). Temporal resolution of tonal pulses. Journal of the Acoustical Society of America, 51, 644-647.

Wisniewski, A. B., & Hulse, S. H. (1997). Auditory scene analysis in European starlings (Sturnus vulgaris): Discrimination of song segments, their segmentation from multiple and reversed conspecific songs, and evidence for conspecific song categorization. Journal of Comparative Psychology, 111, 337-350.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. Nature, 387, 176-178.

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 & Development, 21, 483-492.

Zoloth, S. R., Petersen, M. R., Beecher, M. D., Green, S., Marler, P., Moody, D. B., & Stebbins, W. (1979). Species-specific perceptual processing of vocal sounds by monkeys. Science, 204, 870-873.

Tables

	human infants	non-human vertebrates
melody contour (relative pitch perceptoin)	+[1]	+ (European starling[2], zebra finch, pigeon[3], bottlenose dolphin[4]) - (rat, cebus monkey[5])
pitch channel	+[6]	
stream segregation	+[7]	+ (European starling[8])

Table 1. Some auditory perceptual characteristics concerning sequential relations in human infants and non-human vertebrates. Each sign of plus and minus represents positive and negative results, respectively. Pitch channel refers the effect of frequency disparity/proximity on the temporal discrimination. There has been no study concerning pitch channel in non-humans. 1: Demany & Armand, 1984; Trehub, Bull, & Thorpe, 1984; Trehub, Thorpe, & Morrongiello, 1987; Chang & Trehub, 1977. 2: Hulse, Cynx, & Humpal, 1984; Hulse & Cynx, 1985; MacDougall-Shackleton & Hulse, 1996; Cynx, 1995; Page, Hulse, & Cynx, 1989; Braaten, Hulse, & Cynx, 1990; Cynx, Hulse, & Polyzois, 1986. 3: Cynx, 1995. 4: Ralston & Herman, 1995. 5: D'Amato & Salmon, 1982, 1984; D'Amato, 1988. 6: Thorpe & Trehub, 1989. 7: Demany, 1982. 8: Braaten & Hulse, 1993; MacDougall-Shackleton, Hulse, Gentner, & White, 1998.

		human infants	non-human vertebrates
missing fundamental		+[1]	+ (goldfish[2], cat[3], European starling[4], rhesus monkey[5])
octave generalization		+[6]	+ (rat[7], bottlenose dolphin[8]) - (European starling[9], cebus monkey[10])
consonance	sensory	+[11]	+ (European starling[12])
	musical	+[13]	

Table 2. Some auditory perceptual characteristics concerning harmonic relations in human infants and non-human vertebrates. Each sign of plus and minus represents positive and negative results, respectively. The fluctuation resulted from the interaction between harmonic components is perceived as a sensory dissonance. Perception of musical consonance refers to the ability to perceive frequency ratio simplicity without sensory consonance. Perception of musical consonance has not been studied with non-humans. 1: Clarkson & Clifton, 1985. 2: Fay, 1970. 3: Heffner & Whitfield, 1976. 4: Cynx & Shapiro, 1986. 5: Tomlinson & Schwarz, 1988. 6: Demany & Armand, 1984. 7: Blackwell & Schlosberg, 1943. 8: Richards, Wolz, & Herman, 1984. 9: Cynx, 1993. 10: D'Amato & Salmon, 1982, 1984; D'Amato, 1988. 11: Schellenberg & Trainor, 1996. 12: Hulse, Bernard, & Braaten, 1995. 13: Trainor, 1997; Cohen, Thorpe, & Trehub, 1987.

Phase	Target A	Target B	Distracter	Captor
Training	2000	2200		
Condition A	2000	2200	1700	700
Condition B	2000	2200	1700	1700
Condition C	2000	2200	1700	

Table 3. Tone frequency (in Hz) for each stimulus conditions. Hyphens represent the corresponding tone is not exist in the stimulus condition. For example, baseline stimulus sequence composed of only two target tones.

chord name	interval size (semitone)	frequency ratio
unison	0	1:1
minor second	1	15:16
major second	2	8:9
minor third	3	5:6
major third	4	4:5
perfect fourth	5	3:4
tritone	6	32:45
perfect fourth	7	2:3
minor sixth	8	5:8
major sixth	9	3:5
minor seventh	10	9:16
major seventh	11	8:15
octave	12	1:2
minor ninth	13	15:32

Table 4. Chords used in the transfer test of Experiment 1. Although stimulus frequency used in the present study was in equal-tempered scale, frequency ratios presented in this table stand on justly-tuned scale.

chord pair	consonant chord	dissonant chord
pair 1	perfect fifth (2:3; 7 semitones)	major seventh (8:15; 11 semitones)
pair 2	perfect fourth (3:4; 5 semitones)	minor second (15:16; 1 semitones)

Table 5. Chords used in Experiment 2. The frequency ratio and interval size of each chord is shown in parenthesis. There were four types of conditions in a session; both two pairs appeared, and consonant and dissonant chords in each pair was reversed their role as background and target stimuli.

Figure Captions

Figure 1. Frontal panel of the experiment box.

Figure 2. A schematic diagram of the task. In the case shown, the subject performs a correct response and gets a reward.

Figure 3. Schematic representation of the stimuli. (a) Frequency-identical (Fi) conditions. (b) Frequency-different (Fd) conditions.

Figure 4. Psychometric functions of a subject (H) obtained with 500-600 stimulus condition for each number of background repetition. The ordinate represents d', and the abscissa represents the value of ΔT.

Figure 5. Gap detection thresholds for each repetition number of background stimuli. The ordinate represents discrimination threshold of gap duration decrement, and the abscissa represents marker conditions (frequency of pre- and post-gap markers).

Figure 6. Schematic representation of the stimuli. In the case shown, the duration of each tone is 75 ms, and the sequence of target tones is ascending. In transfer test 3, the duration of a target tone was 100 ms and that of distracter and captor tone was 50 ms.

Figure 7. Individual discrimination index for each stimulus condition and transfer tests. Symbols represent the performance in each transfer test.

Figure 8. Schematic representations of stimuli. Each horizontal line represents a 200-ms pure tone.

Figure 9. Response rates for each rising and falling stimulus. The abscissa represents repetition numbers of background stimuli. Filled symbols and open symbols represent rising and falling stimuli, respectively.

Figure 10. Individual response rates for baseline and probe stimuli in probe tests. (a) Probe test 1. (b) Probe test 2. (c) Probe test 3.

Figure 11. Schematic representation of a trial. Each stimulus consisted of two simultaneous tones. Though each tone comprised of six harmonics, only the fundamental component is shown. Background stimuli were repetitively presented with transposition and a target stimulus followed to them. A target stimulus differed to backgrounds in the frequency interval of two tones.

Figure 12. Discrimination index for each subject in Experiment 1. The ordinate represents discrimination index, and the abscissa represents chord intervals (in semitone). Symbols represent the performance for each subject. Each bar represents the consonance index by Schellenberg and Trehub (1994b).

Figure 13. Mean discrimination index for each stimulus conditions in Experiment 2. The ordinate represents discrimination index, and the abscissa represents chords. Pair 1 consisted of perfect fifth and major seventh, and pair 2 was perfect fourth and minor second. Error bars represent standard errors.

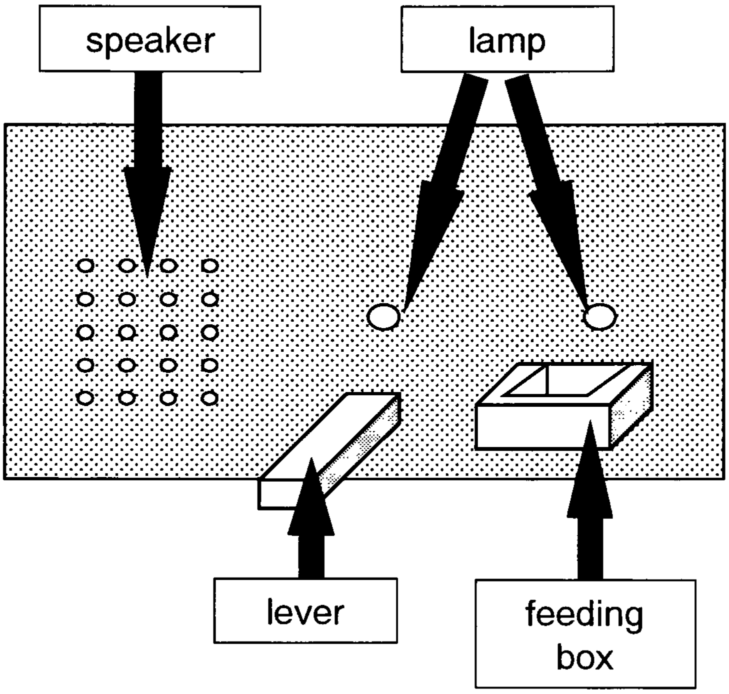


Figure 1

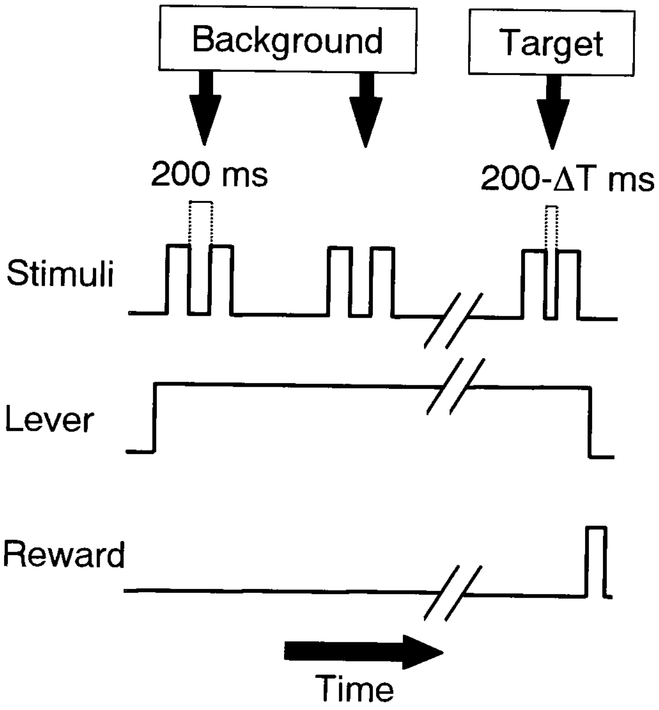

Background
Target
200 ms
200-ΔT ms
Stimuli
Lever
Reward
Time

Figure 2

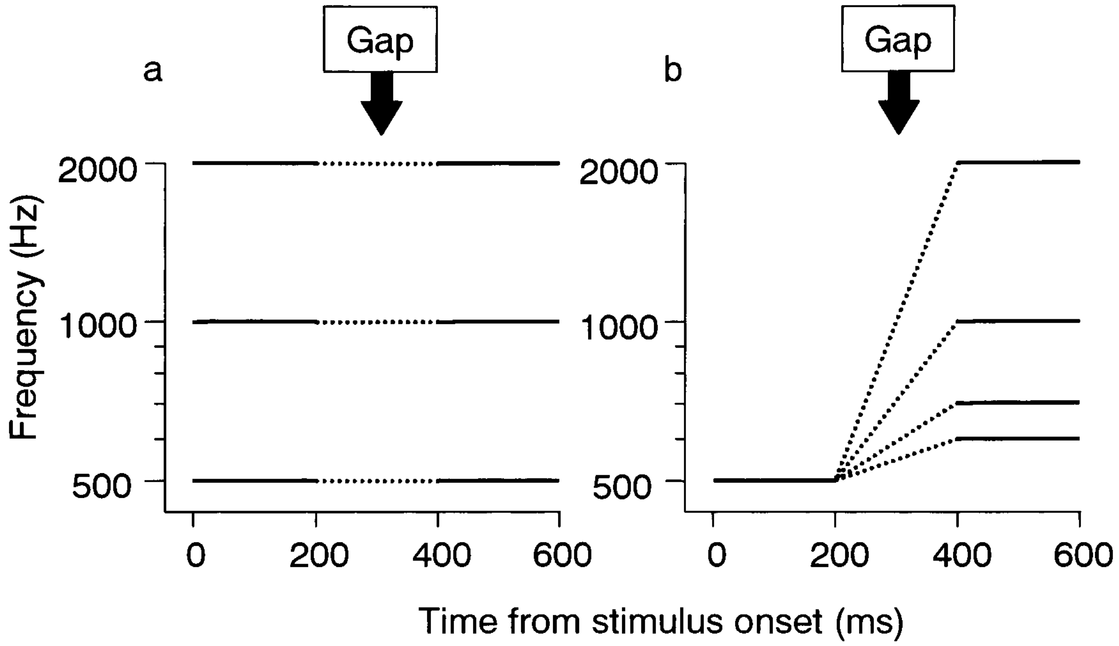

a
Gap
b
Gap
2000
1000
500
Frequency (Hz)
2000
1000
500
0
200
400
600
0
200
400
600
Time from stimulus onset (ms)

Figure 3

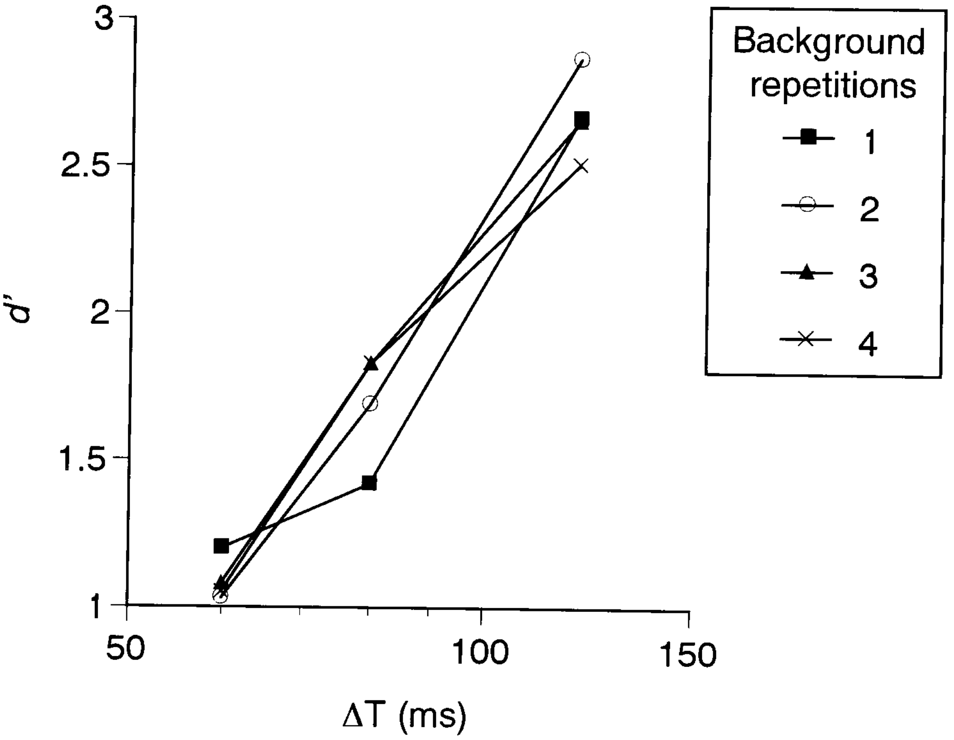


Figure 4

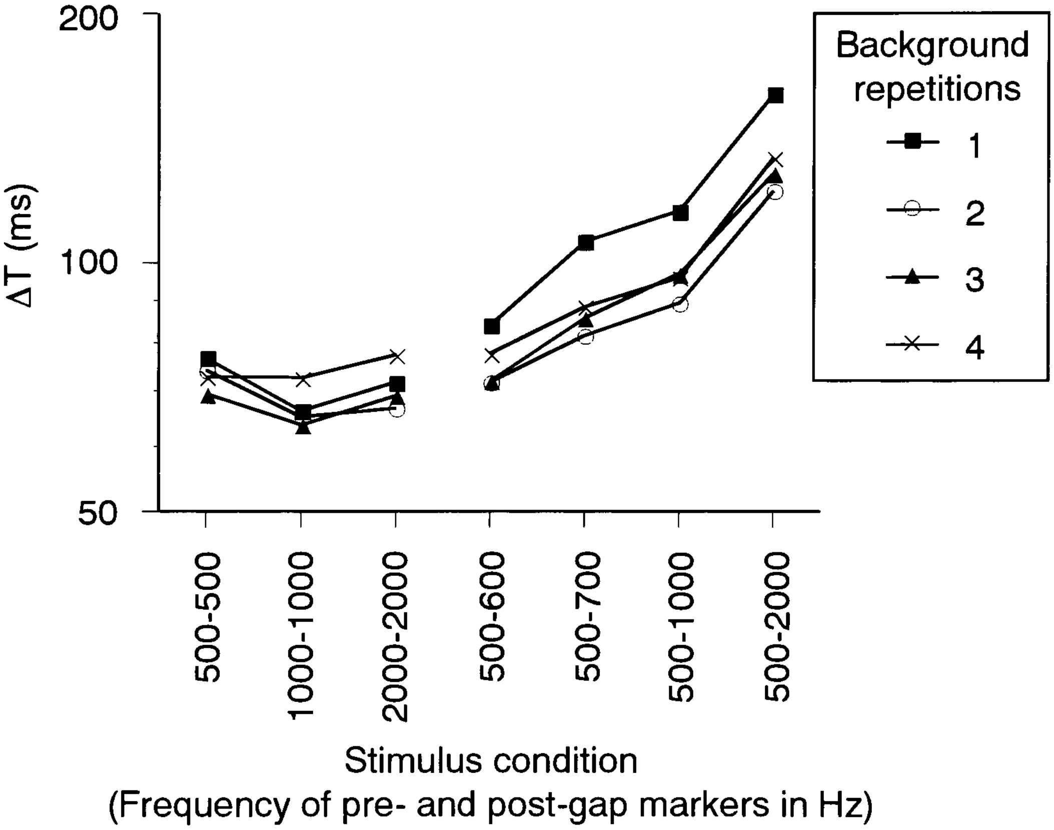

200
100
50
ΔT (ms)
Background repetitions
1
2
3
4
500-500
1000-1000
2000-2000
500-600
500-700
500-1000
500-2000
Stimulus condition
(Frequency of pre- and post-gap markers in Hz)

Figure 5

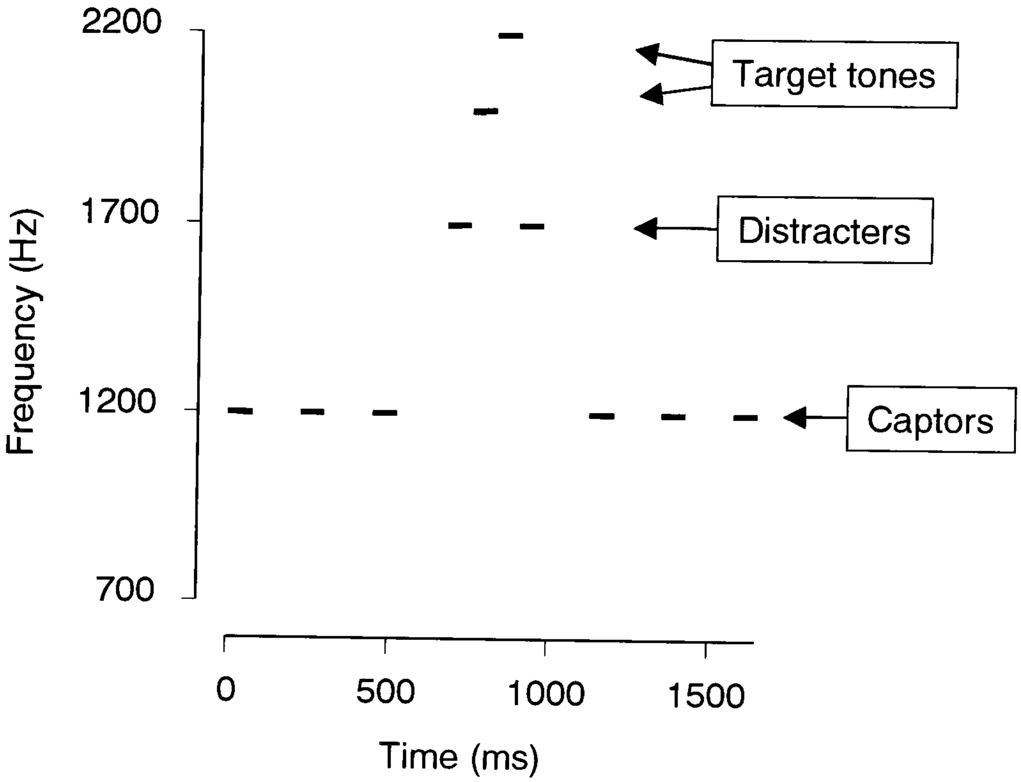

2200
1700
1200
700
Frequency (Hz)
0
500
1000
1500
Time (ms)
Target tones
Distracters
Captors

Figure 6

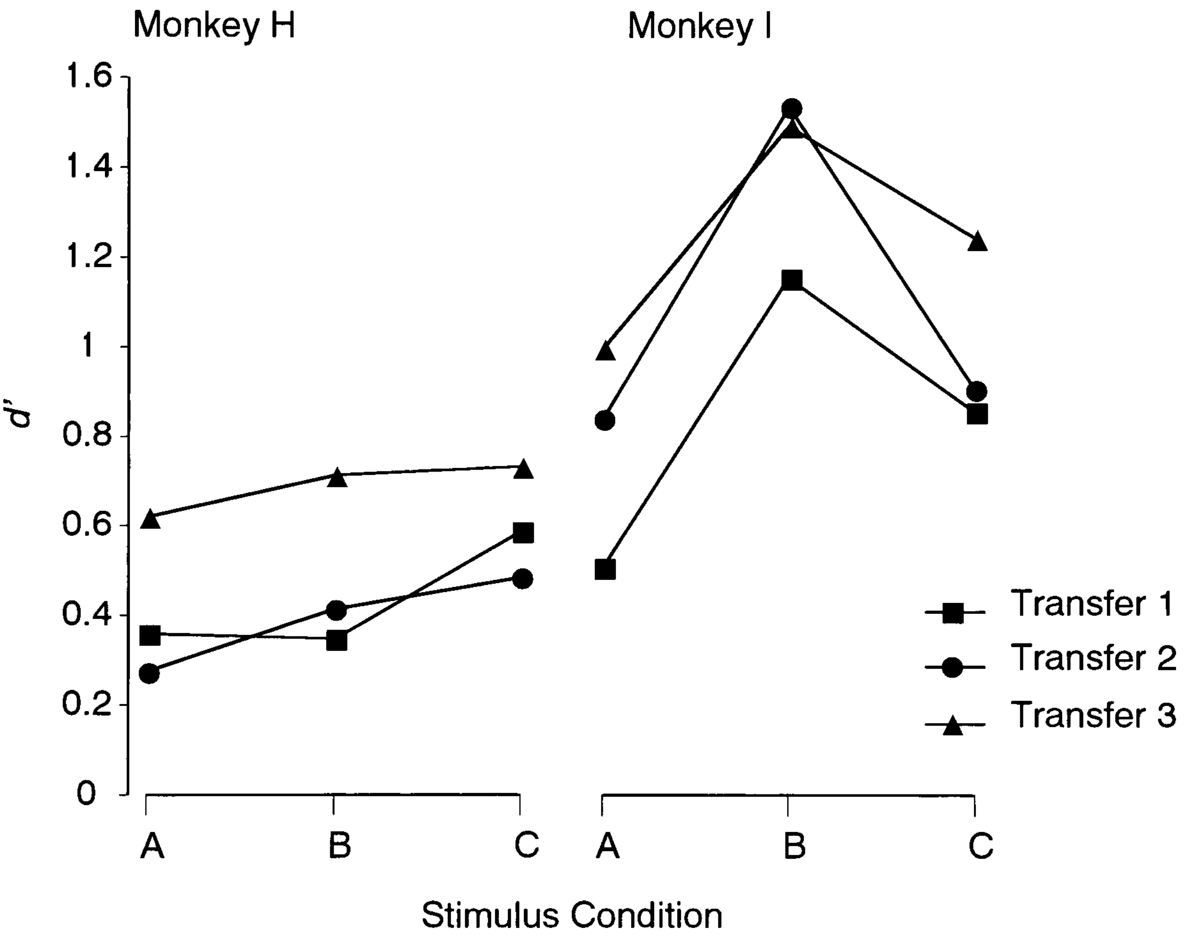

Monkey H
Monkey I
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
d'
A
B
C
A
B
C
Stimulus Condition
Transfer 1
Transfer 2
Transfer 3

Figure 7

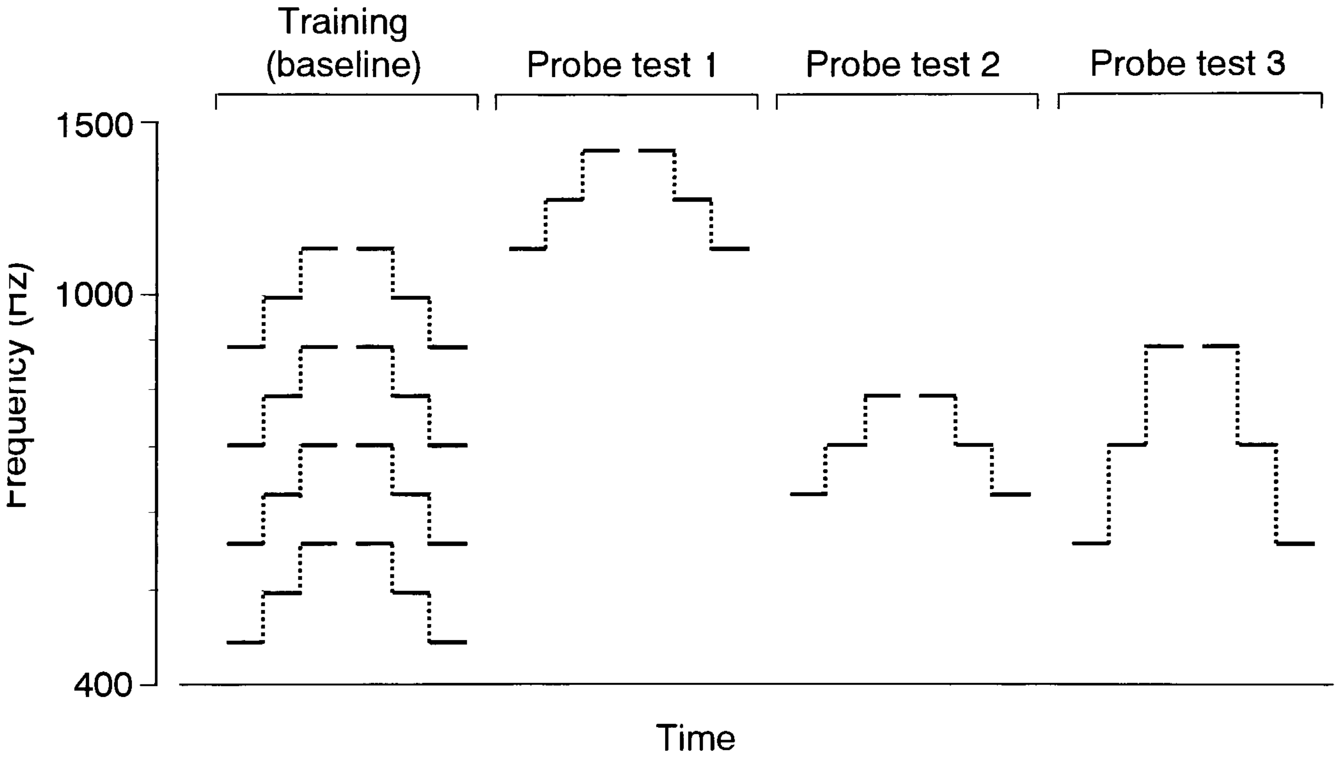
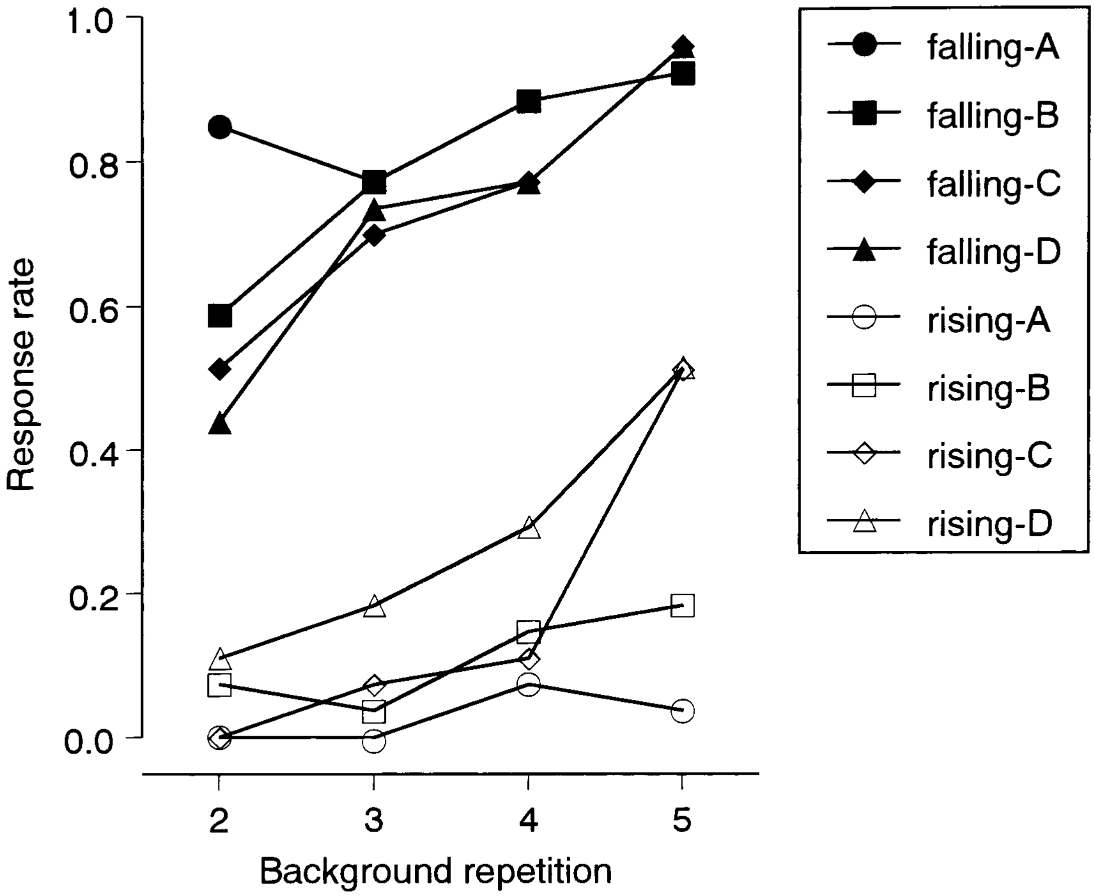

Training
(baseline)
Probe test 1
Probe test 2
Probe test 3
1500
1000
400
Frequency (Hz)
Time

Figure 8



1.0
0.8
0.6
0.4
0.2
0.0
Response rate
2
3
4
5
Background repetition
falling-A
falling-B
falling-C
falling-D
rising-A
rising-B
rising-C
rising-D

Figure 9

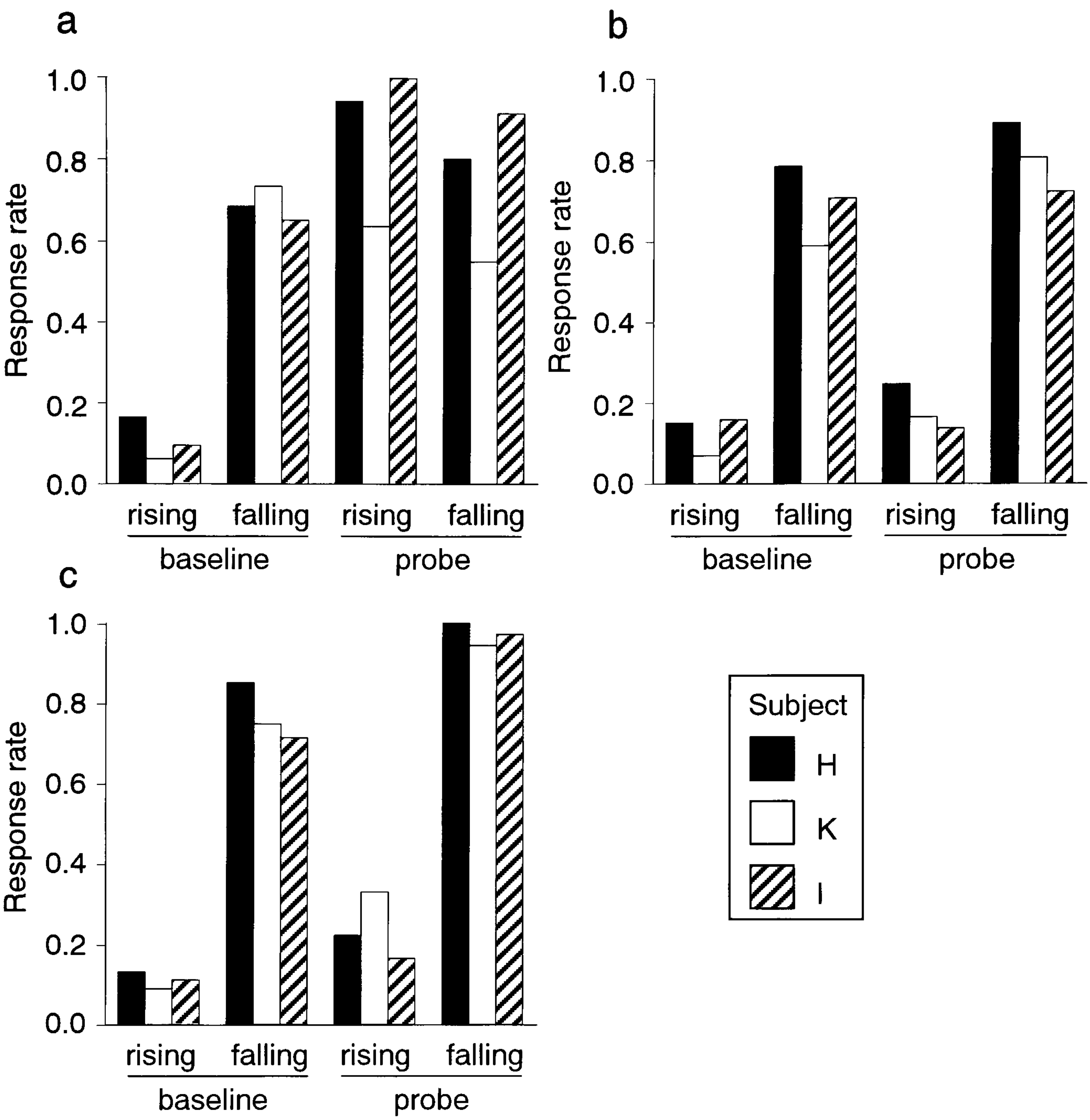

a
b
c
Response rate
1.0
0.8
0.6
0.4
0.2
0.0
rising
falling
rising
falling
baseline
probe
Subject
H
K
I

Figure 10

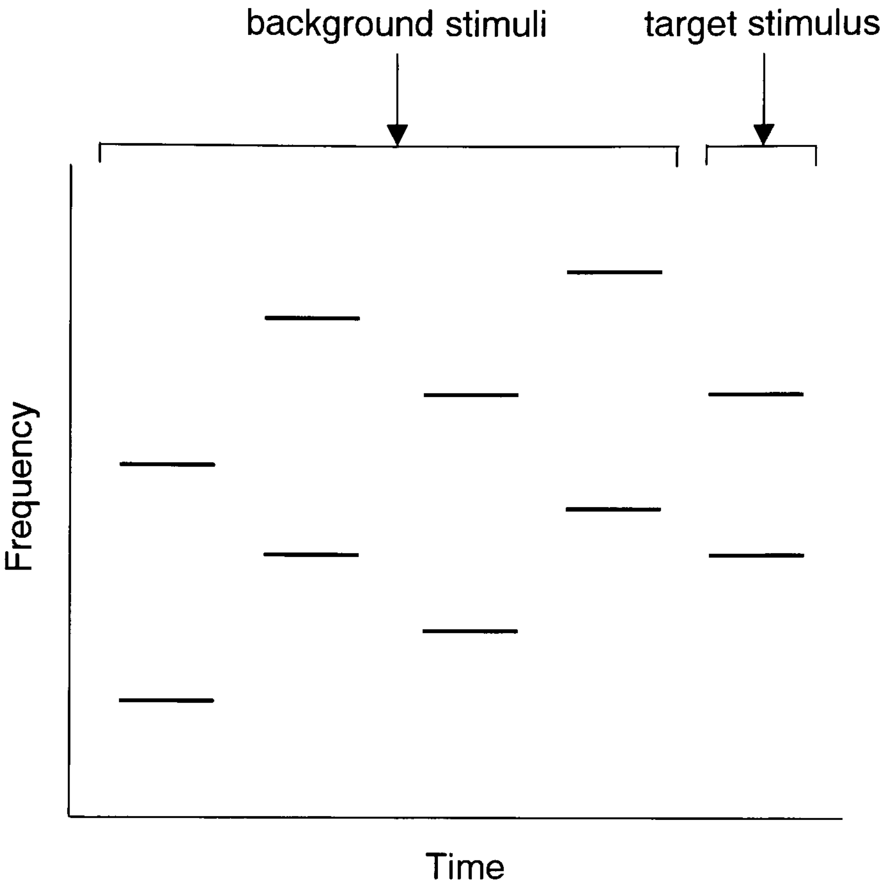
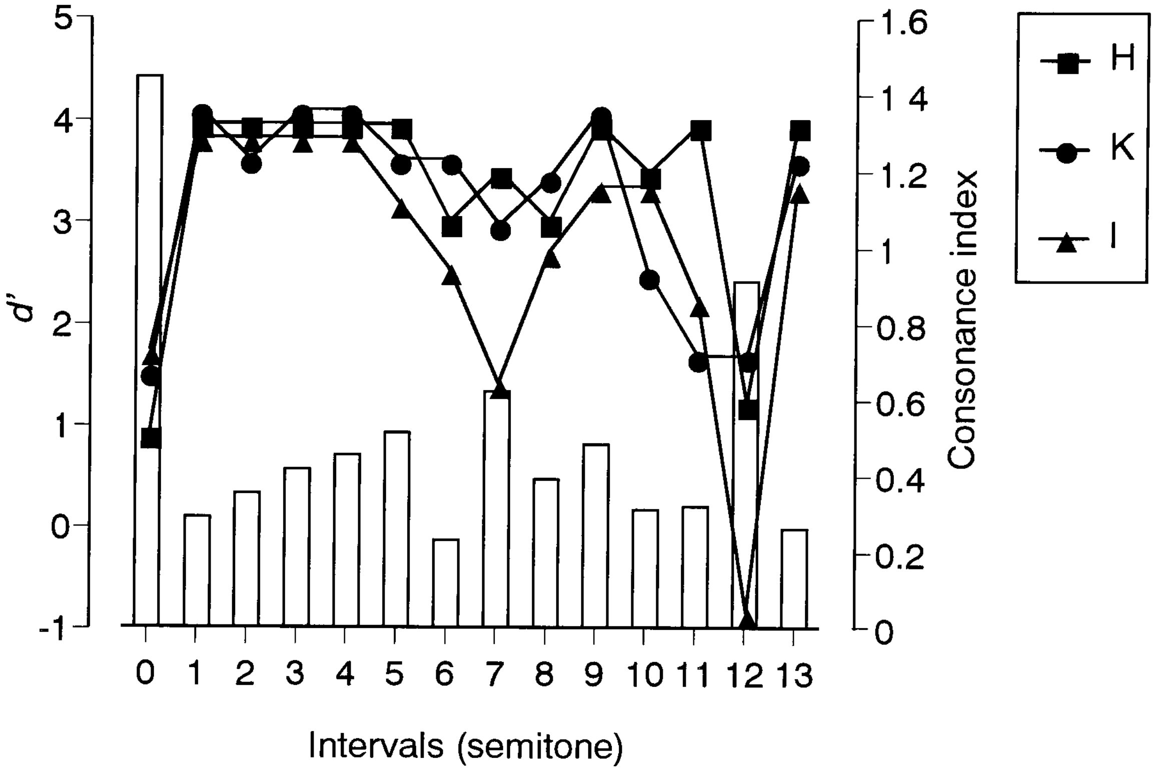


Figure 11



5
4
3
2
1
0
-1
d'
1.6
1 4
1.2
1
0.8
0.6
0.4
0.2
0
Consonance index
H
K
I
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Intervals (semitone)

Figure 12

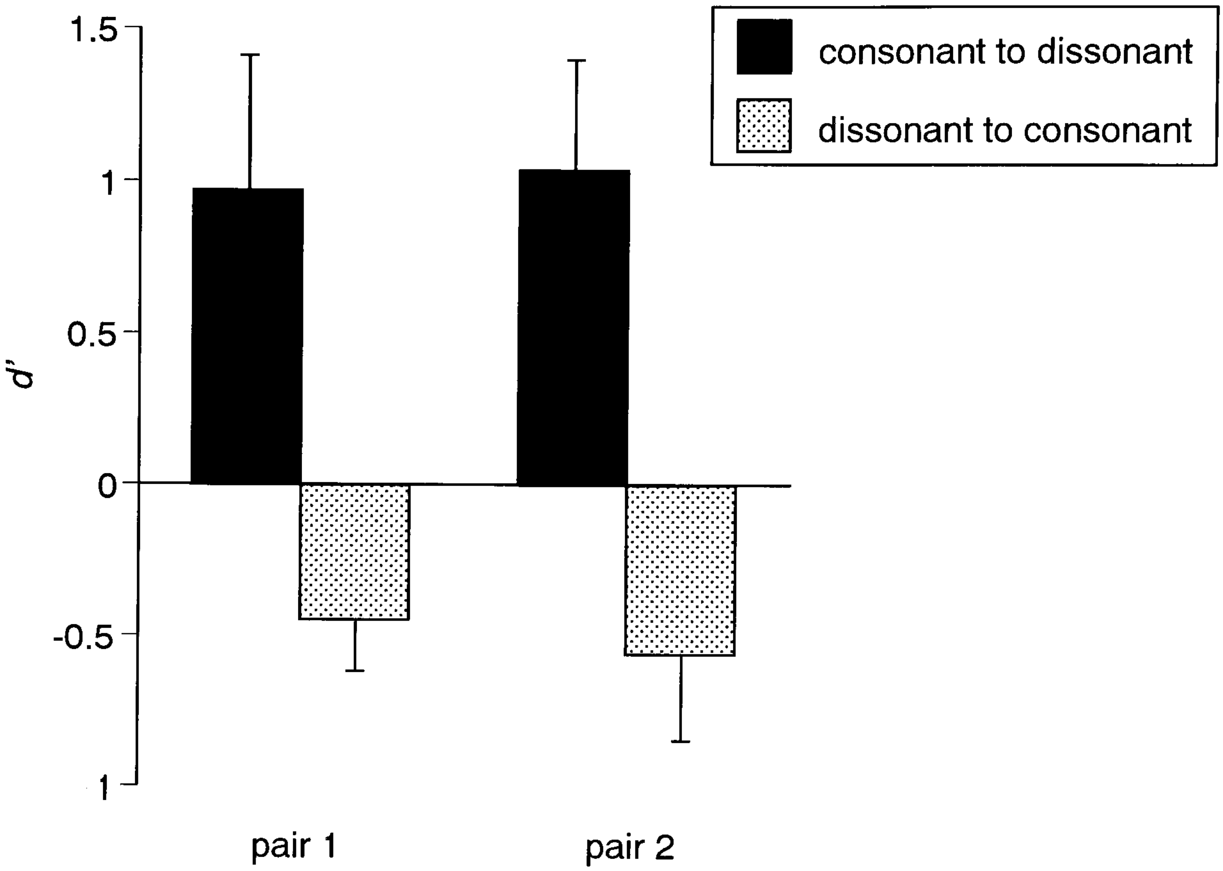

1.5
1
0.5
0
-0.5
1
d'
consonant to dissonant
dissonant to consonant
pair 1
pair 2

Figure 13