Auditory temporal resolution in birds: Discrimination of harmonic complexes

Robert J. Dooling[a)]
Department of Psychology, University of Maryland, College Park, Maryland 20742

Marjorie R. Leek
Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, DC 20307

Otto Gleich
ENT Department, University of Regensburg, 93042 Regensburg, Germany

Micheal L. Dent
Department of Psychology, University of Maryland, College Park, Maryland 20742

(Received 13 November 2001; revised 15 May 2002; accepted 23 May 2002)

The ability of three species of birds to discriminate among selected harmonic complexes with fundamental frequencies varying from 50 to 1000 Hz was examined in behavioral experiments. The stimuli were synthetic harmonic complexes with waveform shapes altered by component phase selection, holding spectral and intensive information constant. Birds were able to discriminate between waveforms with randomly selected component phases and those with all components in cosine phase, as well as between positive and negative Schroeder-phase waveforms with harmonic periods as short as 1–2 ms. By contrast, human listeners are unable to make these discriminations at periods less than about 3–4 ms. Electrophysiological measures, including cochlear microphonic and compound action potential measurements to the same stimuli used in behavioral tests, showed differences between birds and gerbils paralleling, but not completely accounting for, the psychophysical differences observed between birds and humans. It appears from these data that birds can hear the fine temporal structure in complex waveforms over very short periods. These data show birds are capable of more precise temporal resolution for complex sounds than is observed in humans and perhaps other mammals. Physiological data further show that at least part of the mechanisms underlying this high temporal resolving power resides at the peripheral level of the avian auditory system. [DOI: 10.1121/1.1494447]

PACS numbers: 43.80.Lb, 43.66.Gf, 43.64.Tk [WA]

I. INTRODUCTION

Bird vocalizations are known to be extremely complex acoustic signals. This observation, along with the well-known differences between birds and mammals in the anatomy and physiology of the peripheral and central auditory systems, has led some to suggest that birds must have extremely fine temporal processing abilities (Carr and Friedman, 1999; Greenewalt, 1968; Konishi, 1969; Pumphrey, 1961; Schwartzkopff, 1968). Notwithstanding such indications, however, direct psychophysical measures of temporal processing such as detection of gaps in noise, temporal integration, duration discrimination, and temporal modulation transfer functions have all shown that birds are, overall, not more sensitive to the temporal features of acoustic signals than are other vertebrates (Dooling, 1982; Dooling *et al.*, 2000; Dooling and Haskell, 1978; Dooling and Searcy, 1981, 1985; Fay, 1988; Klump and Maier, 1989).

For the most part, the psychophysical tests that have failed to reveal differences between mammals and birds in temporal resolution have used simple stimuli and involved slow, overall changes in amplitude (the waveform envelope) rather than rapid pressure variations that carry acoustic information (the temporal fine structure). It may be that such stimuli do not provide an adequate test of the limits of temporal resolution in the avian ear. Many birds produce and learn complex tonal or harmonic vocalizations that involve rapid modulations in frequency and amplitude (for reviews, see Kroodsma and Miller, 1982, 1996), some of which are inaudible to humans. It is of interest to know whether the kinds of changes that occur in these complex sounds are within the perceptual capabilities of the species of birds that learn and use them as communication signals. As one example, the harmonic vocalizations of the zebra finch (*Taeniopygia guttata*) have very short fundamental periods of about 1.5 ms, shorter than most estimates of temporal resolution in the human auditory system (Viemeister and Plack, 1993). Any acoustic information produced by waveform fine structure within these periods is undoubtedly lost to humans, but may be available to zebra finches and other birds.

Taking human speech as an example, acoustic variability in intensity, frequency, and time, both in steady-state utterances like vowels and in rapidly changing sounds such as consonants and diphthongs, provides information to the listener regarding the speaker's individual identification, emotional state, and intended message. While we are much more familiar with human speech than with bird vocalizations, even the most cursory analysis of vocalizations of various

a)Electronic mail: dooling@psyc.umd.edu

 0001-4966/2002/112(2)/748/12/$19.00

bird species shows that acoustic differences among calls include changes in spectrum, waveform envelope, and temporal fine structure both within and across vocalizations of individuals (as, for example, the fine structure within the periods of harmonic zebra finch vocalizations). To the extent that these changes in acoustic characteristics are distinguishable to the intended communication target, they have the potential of being communicatively relevant. Historically, analyses of bird vocalizations have focused extensively on spectral features, much less on global temporal or envelope modulations occurring over the duration of vocalizations, and not at all on variations in temporal fine structure (Kroodsma and Miller, 1982, 1996). Perception of temporal fine structure may be more relevant to the problems of acoustic communication than previously thought, given the recent findings on the degree of fine motor control in avian vocal production (Brainard and Doupe, 2001; Fee *et al.*, 1998; Tchernichovski *et al.*, 2001; Yu and Margoliash, 1996). In the following series of experiments, the abilities of three species of birds to discriminate temporal envelope and temporal fine structure in complex sounds have been assessed. These sounds have similar spectra but differ systematically in envelope and fine structure. Both behavioral and electrophysiological experiments have been undertaken in a converging approach to defining those temporal limits and to providing preliminary evidence concerning the physiological bases of the perceived differences among stimuli. In experiment 1, birds were required to discriminate harmonic complexes constructed with frequency components all starting in cosine phase or all in randomly selected phases, with fundamental periods varying between 10 and 1 ms. Complexes within pairs of stimuli to be discriminated thus have equivalent long-term spectra, but different temporal waveform structures. Moreover, discrimination must be based on increasingly short segments of the stimuli as the period shrinks. A second experiment retested the birds' ability to use intraperiod structure in the waveforms of harmonic complexes, but envelope information was essentially removed as a cue from the stimuli, and only fine structure remained as a basis for making discriminations. Within a stimulus pair, the members to be discriminated are the time reverse of one another, with harmonic component phases selected according to an algorithm that systematically increases or decreases phase (and instantaneous frequency within the periods). In all, birds of three different species were trained by operant conditioning with food reward to discriminate within pairs of harmonic complexes over several different fundamental frequencies. The results for birds were compared to the results from humans tested on identical stimuli. Finally, cochlear microphonic (CM) and compound action potential (CAP) recordings from the three bird species to the same harmonic complexes used in behavioral tests revealed correlates to the species differences in sensitivity to temporal fine structure obtained behaviorally.

II. EXPERIMENT 1—DISCRIMINATION OF COSINE AND RANDOM PHASE HARMONIC COMPLEXES

In the first behavioral experiment, discrimination of harmonic stimuli with two kinds of phase selection was measured to determine whether, when frequency and amplitude information are held constant, the shape of the temporal waveform alone can provide different perceptions. By using harmonic complexes with a range of fundamental frequencies, we can also determine the limits of the duration of the harmonic periods that will support this discrimination. The discrimination task in each case was between a harmonic complex with each of the components starting in cosine phase (resulting in a highly peaked waveform) with seven different harmonic complexes with all components starting in randomly selected phases.

A. Materials and methods

1. Subjects

Three budgerigars (*Melopsittacus undulatus*) were tested on the discrimination of cosine phase harmonic complexes from random phase harmonic complexes. These birds were either bought commercially or hatched at the University of Maryland and housed in individual cages in a vivarium at the University of Maryland. The birds were kept on a normal day/night cycle correlated with the season at approximately 90% of their free-feeding weights. Animal housing and care met all standards of the University of Maryland Animal Care and Use Committee. All birds had hearing within normal limits for their species, as shown by their audiograms (Dooling *et al.*, 2000). Three humans were also tested. They were experimenters working in the laboratory at the time of the experiment, reported no history of hearing disorders, and had absolute thresholds at audiometric test frequencies better than 20 dB HL (*re* ANSI, 1989).

2. Stimuli and procedures

The stimuli were harmonic complexes consisting of a set of equal-amplitude harmonic components of a given fundamental frequency, with frequencies ranging from 200 (or the fundamental frequency) to 5000 Hz [see Fig. 1(a)]. The actual number of harmonic components varied with the fundamental frequency, which was either 200, 400, 800, or 1000 Hz (fundamental periods of 5, 2.5, 1.25, and 1 ms). For each fundamental frequency, one complex was generated with all components in cosine starting phase, and seven different complexes were produced with each component phase selected randomly from a rectangular distribution ranging from 0 to 2π radians. A different set of random phases was selected for each complex at each of the four fundamental frequencies. Within a given fundamental frequency, each stimulus had identical long-term frequency spectra, but differently shaped temporal waveforms. The waveforms were 260 ms in duration including 20-ms cosine^2 onset and offset ramps. Figure 1(b) shows examples of cosine-phase and random-phase stimuli for two of the fundamental frequencies used in this experiment.

The waveforms were created digitally, at a sampling rate of 40 kHz, using software provided by Tucker Davis Technologies (TDT, Gainesville, FL) to combine frequencies in the correct phases and amplitudes, followed by an inverse

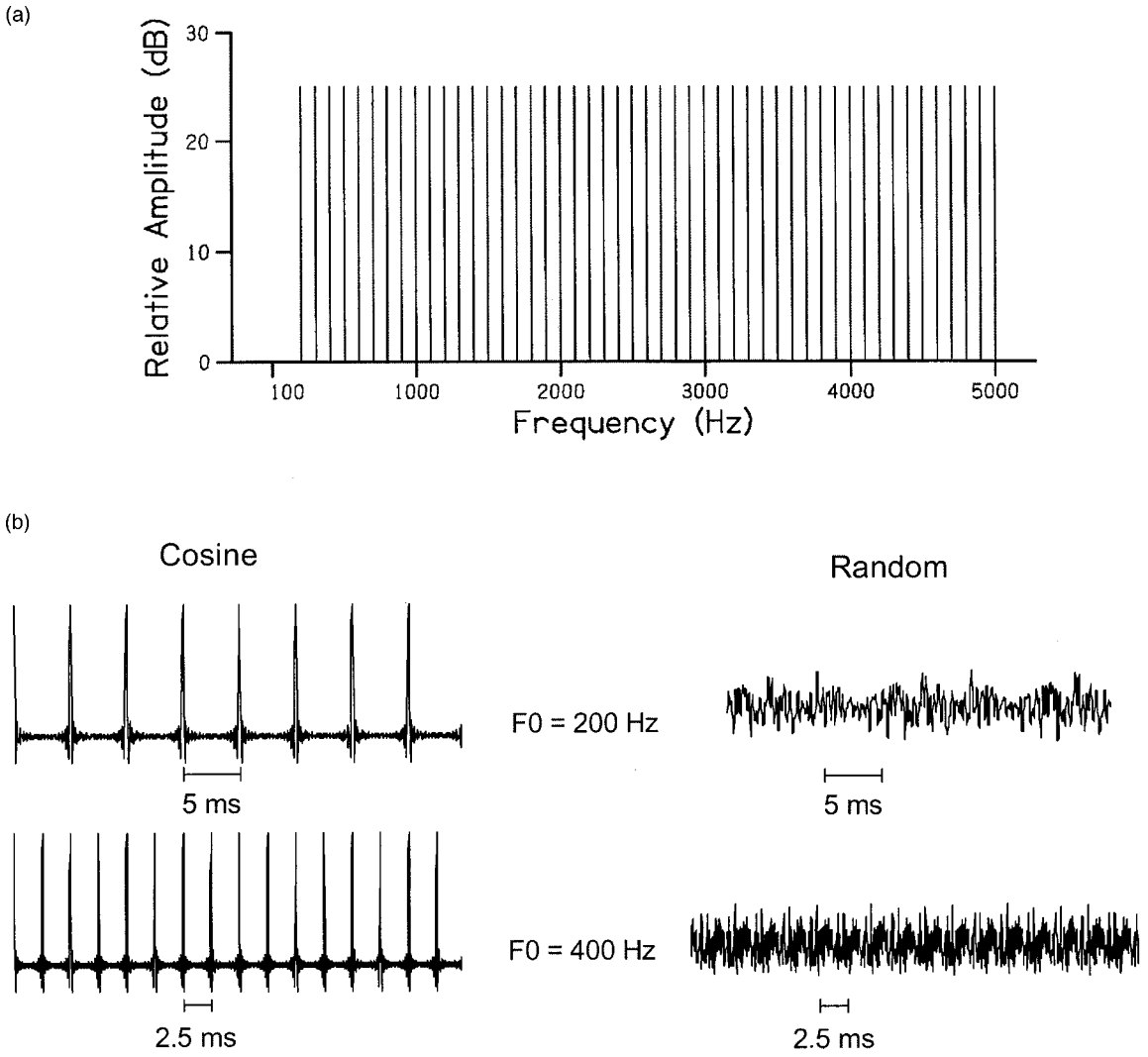

FIG. 1. (a) Schematic of the spectrum of one of the harmonic complexes used in this study. The frequencies in the stimuli ranged from 200 to 5000 Hz. (b) Temporal waveforms of harmonic complexes with the phase of each component starting in either cosine phase (left) or a randomly-selected phase (right).

fast Fourier transform (FFT) to create the waveforms. They were created off-line and stored as files for playback during the experiments.

3. Behavioral testing apparatus

The birds were tested in a wire cage ($23\times25\times16$ cm^3) mounted in a sound-isolation chamber (Industrial Acoustics Company, IAC-3). A response panel consisting of two microswitches with light-emitting diodes (LEDs) was mounted on the wall of the test cage just above a food hopper. The bird could trip the microswitch by pecking at the LED. The left microswitch and LED served as the observation key while the right microswitch and LED served as the report key. The behavior of the animals during test sessions was monitored by a video camera system (Sony HVM-322).

Test sessions were controlled by a Pentium PC computer. The digital stimuli were output to an overhead loudspeaker (KEF Electronics, Holliston, MA, model 80C), located 25 cm above the bird's head. Stimuli were presented through Tucker-Davis modules at a sampling rate of 40 kHz and presented at 80 dB SPL. Stimulus calibration was performed using a Larson Davis (Provo, UT) sound level meter (model 824). Stimulus intensities were measured with a $\frac{1}{2}$-in. microphone attached to the sound level meter via a 3-m extension cable. The microphone was placed in front of the response keys in the approximate position occupied by the bird's head during testing. Stimulus intensities were measured several times during these experiments to ensure that stimulus levels remained constant and that the entire audiometric system remained calibrated.

4. Training and testing procedures

Birds were trained by standard operant auto-shaping procedures to peck at the left LED (observation key) during a repeating background of sound (i.e., the presentation of a cosine phase complex at a rate of 2/s, or an interstimulus interval of 240 ms) until a new stimulus (i.e., the presentation of one of the random phase harmonic complexes with the same fundamental frequency) was presented alternately with the background sound, and then to peck the right LED (report key) when the change was detected. If the bird pecked the report key within 2 s of this alternating pattern, the food hopper was activated for 2 s, allowing the bird to obtain food reinforcement. During each experimental ses-

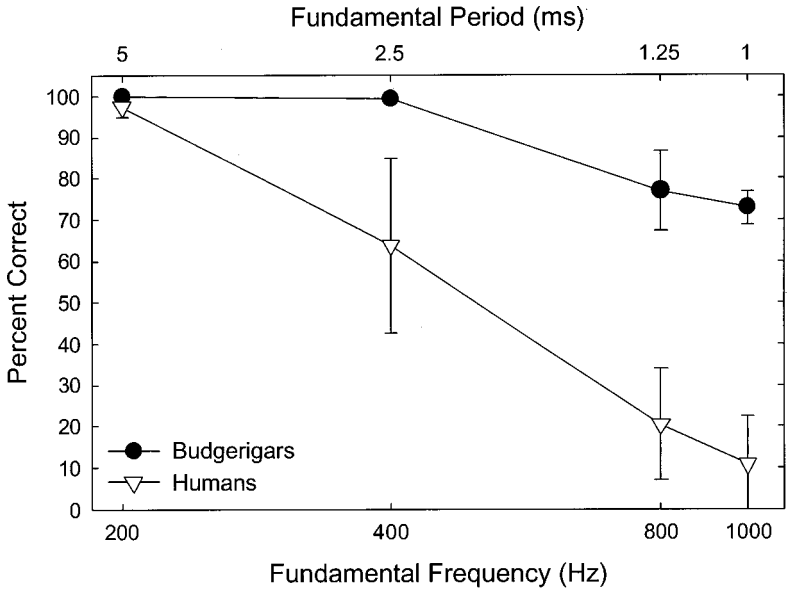

FIG. 2. Results from three budgerigars and three humans tested on the cosine phase versus random phase waveform discrimination at different fundamental frequencies. Performance is shown as percent correct discriminations and error bars are the between-subject standard errors.

sion, subjects listened to only one of the sets of fundamental frequency complexes and were required to discriminate between the cosine phase and the random phase complexes within the set. The harmonic complexes of all four fundamental frequencies (i.e., 200, 400, 800, and 1000 Hz) were tested in a random order and a different random order was used for each bird.

The bird initiated a trial by pecking repeatedly on the observation key. The time between pecking the observation key (i.e., the start of a trial) and the beginning of an alternating sound pattern was randomized between 2 and 7 s. Hit rate was defined as the proportion of target trials (when the background alternated with one of the random phase targets) on which the bird pecked the report key within 2 s of the alternating sound pattern. A failure to peck the report key within 2 s of sound alternation was recorded as a miss for a target trial and a correct rejection in the case of a sham trial. Following a miss, no reinforcement was given and a new trial sequence was initiated. Thirty percent of all trials were sham trials in which the target sound was the same as the repeating background sound. A peck to the report key during a 2-s sham trial was recorded as a false alarm, and the lights in the test chamber were extinguished while the repeating background continued. The length of this time-out period was normally 5 s, but varied according to an individual bird's behavior, with longer time-out periods imposed if birds began developing higher false alarm rates. Sessions with a total false alarm rate of 16% or higher were discarded. Fewer than 20% of all sessions across birds were discarded for this reason. The mean false alarm rate across birds was 5%. The birds were typically tested in one to two daily sessions consisting of about 100 trials each, until percent correct values stabilized, and then testing continued for another 200 trials. Final percent correct discrimination values were taken as the mean percent correct over the last 200 trials.

The human listeners were laboratory staff members, tested with the same stimuli as the birds, using a standard/two-alternative forced-choice procedure. Stimuli were presented over earphones, at a level of 80 dB SPL. For each fundamental frequency, comparisons were made between the cosine-phase stimulus and each of the seven random-phase stimuli in blocks of 40 trials each. On each trial, the cosine-phase complex was presented, followed by the same stimulus and the comparison stimulus, in random order, separated by 300 ms of silence. Subjects were asked to indicate which of the second or third presentation on a trial was different from the standard presentation. After the subject touched a marked area on a touch screen terminal to indicate a response, correct answer feedback was provided, and the next trial was initiated after 500 ms. Percent correct responses were averaged across the seven random-phase stimuli for each fundamental frequency. These values, ranging from chance performance at 50% correct to perfect performance (100%) were subsequently scaled to the range of 0%–100% for comparison with the bird data.

There is always some concern when comparing data for human subjects taken under earphones with animal behavioral data measured in a sound field (see Leek *et al.*, 2000, for a discussion of this problem). To be assured that these differences in measurement were unlikely to materially affect the data reported in this and the following experiments, humans were also tested informally on selected stimuli in the chamber used to test the birds. For these trials, the test cage was removed from the chamber and the human subject stood with their head in the sound chamber in the approximate location of the test cage, with one ear pointed toward the speaker at a distance roughly corresponding to the distance between the speaker and the bird during testing. The same software, procedures, and stimuli were used that were employed in testing the birds. Generally similar results were obtained in these tests when compared with results under earphones. These procedures provided added assurance of the validity of comparing birds tested in free field with humans tested under earphones.

B. Results

Figure 2 shows that budgerigars can discriminate cosine phase harmonic complexes from random phase complexes at

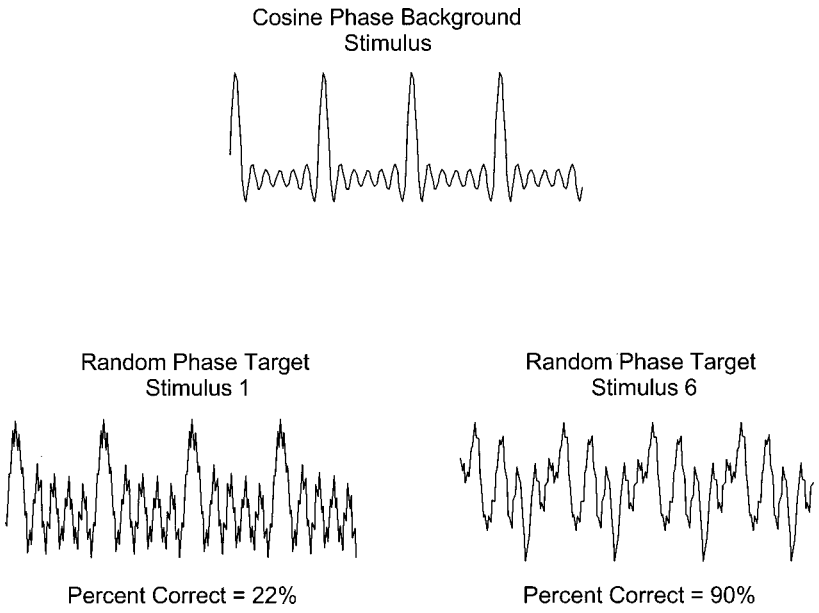

FIG. 3. Several periods of the cosine phase harmonic complex and two of the random phase harmonic complexes, all with a fundamental frequency of 800 Hz. The random phase harmonic complex on the left had a waveform shape more similar to the cosine phase harmonic complex, and overall discrimination was poorer than that for the random phase harmonic complex on the right, which had a shape less similar to the cosine phase harmonic complex.

much higher fundamental frequencies (shorter periods) than can humans. The budgerigars performed at 100% up to 400 Hz, while humans had dropped to about 65% at that fundamental frequency. Even at the highest fundamental frequency of 1000 Hz, the budgerigars' performance was superior to that of humans at 400 Hz. A two-way repeated measures (species×fundamental frequency) ANOVA showed a significant effect of species [$F(1,4)=19.32$, $p<0.05$] and fundamental frequency [$F(3,12)=13.69$, $p<0.001$], but there was no significant interaction between the variables [$F(3,12)=3.41$, $p>0.05$]. *Post-hoc* tests using a Bonferroni *t*-test showed that budgerigars were better than humans at fundamental frequencies ranging from 400 to 1000 Hz ($p<0.05$). The results with humans are consistent with those of Patterson (1987), who found that human subjects could discriminate a cosine-phase waveform from a random-phase waveform as long as the fundamental frequency of the waveforms was below about 400–500 Hz.

By chance, some of the random phase selections produced waveforms with envelopes similar to cosine phase waves. Examples of a cosine phase harmonic complex and two random phase harmonic complexes with fundamental frequencies of 800 Hz are shown in Fig. 3. Performance for the birds was worse on the random phase complexes that were similar to the envelope shape of the cosine phase complex (e.g., random phase target 1 PC=22%) than on random phase complexes that were dissimilar (e.g., random phase target 6 PC=90%). As a measure of similarity between waveforms, a cross-correlation between the cosine-phase waveform and each of the random phase stimuli at fundamental frequencies of 800 and 1000 Hz was calculated (these were the only fundamental frequencies that had sufficient error rates to produce meaningful correlations). In general, as the cosine- and random-phase waveforms increased in similarity, discrimination performance decreased. There was a significant negative correlation between the similarity of the cosine- and random-phase complexes, and the birds' discrimination accuracy ($r=-0.52$, $p<0.05$). The relationship between waveform similarity and discrimination argues that the birds were relying to some extent on characteristics of the temporal waveform in discriminating cosine phase versus random phase complexes.

III. EXPERIMENT 2—DISCRIMINATION BETWEEN POSITIVE AND NEGATIVE SCHROEDER WAVEFORMS

The discrimination of cosine from random phase harmonic complexes was driven, at least in part, by characteristics of the temporal waveform which included both the envelope and the within-period fine structure. Perceptual differences among these stimuli may also include differences in loudness. There are two additional phase selections that can be used to disambiguate the influence of these waveform characteristics. These phase selections essentially hold envelope information constant across stimuli to be discriminated, but reverse the temporal fine structure between stimuli. These Schroeder-phase waveforms have recently been used in a series of behavioral experiments on auditory masking and reveal significant differences in hearing between birds and humans (Dooling *et al.*, 2001; Leek *et al.*, 2000). In humans, but not in birds, Schroeder-phase harmonic complexes constructed with monotonically increasing (positive Schroeder) or decreasing (negative Schroeder) component phases are differentially effective as maskers, even though they have essentially identical temporal envelopes and long-term spectra. One explanation for the similarity in masking effectiveness of these harmonic complexes in birds is that the different masker waveforms were indiscriminable. The present experiment tests this explanation. Further, the temporal limitations on birds' abilities to discriminate harmonic complexes using fine structure alone were tested using pairs of Schroeder-phase waveforms created with different fundamental frequencies.

A. Materials and methods

1. Subjects

Three zebra finches, three budgerigars, and three canaries (*Serinus canaria*) were used as subjects in this experi-

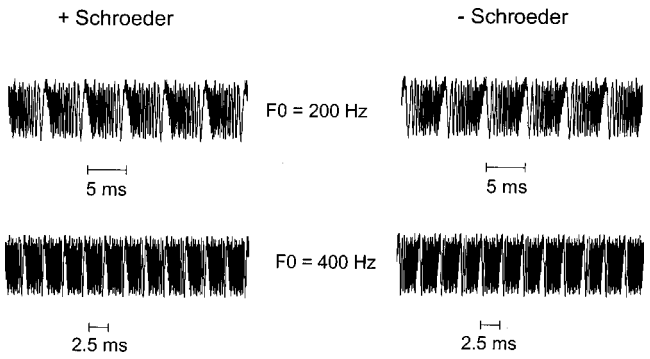

FIG. 4. Temporal waveforms of positive- and negative-Schroeder-phase harmonic complexes with a fundamental frequency of 200 and 400 Hz.

ment. Three humans, researchers in the laboratory, were also tested.

2. Stimuli

Waveforms were constructed in a manner similar to those used in experiment 1, but with component starting phases selected according to an algorithm developed by Schroeder (1970). The component amplitudes were equal, and the frequency range of the stimuli was from 200 (or the fundamental frequency) to 5000 Hz, as was used in experiment 1. Seven pairs of these harmonic complexes were produced, with fundamental periods ranging from 6.6 ms (fundamental frequency of 150 Hz) to 1 ms (fundamental frequency of 1000 Hz) in duration. Figure 4 shows examples of negative and positive Schroeder-phase waveforms for two of the fundamental frequencies used here. The phases of the components were monotonically increasing (positive Schroeder complex) or decreasing (negative Schroeder complex) with harmonic number, resulting in instantaneous frequencies that fell or rose monotonically across each period. The acoustic differences between members of a pair of these complexes are limited to temporal fine structure: all waveforms have a flat envelope and, within a pair defined by the fundamental frequency, have identical long-term spectra. The waveforms were 260 ms in duration including 20-ms cosine^2 onset and offset ramps.

The birds were tested in similar procedures as those described in experiment 1, and used the same psychoacoustic paradigm and experimental chambers. In this experiment, either the positive- or negative-phase waveform was selected as the repeating background, and the other Schroeder-phase wave of the same fundamental frequency was the target sound. Both Schroeder phases were tested as background and target, and the values were averaged for each fundamental frequency. Testing continued for harmonic complexes of all seven fundamental frequencies in random order from 150 to 1000 Hz. The stimuli were generated and output in a similar manner as that described in experiment 1.

Humans were tested under the same conditions (as near as could be) and using the same procedures as for the birds. Humans stood leaning into the small test chamber so that their head was directly under the speaker in the approximate location of the bird's head during testing. A small hand-held button box was used to signal observation and report, analogous to the bird's key pecking. The same stimulus presentation and psychophysical procedure was used as for the birds.

B. Results

Figures 5(a)–(c) show the performance of individual birds of each species tested on the positive/negative Schroeder waveform discriminations at different fundamental frequencies. All birds were able to discriminate between positive and negative Schroeder harmonic complexes at fundamental frequencies up to at least 600 Hz. Budgerigars and canaries showed some difficulty discriminating at the highest fundamental frequencies (800 and 1000 Hz), while the zebra finches discriminated easily between the positive and negative Schroeder waveforms even at the highest fundamental frequency.

Figure 5(d) shows the average of all birds from each species, and the average of three humans. Large differences are evident among the species: human listeners begin having difficulty making these discriminations when the fundamental period becomes shorter than about 3 ms, budgerigars and canaries do much better, and zebra finches have little difficulty even over periods as short as 1 ms. A two-way repeated measures (species×fundamental frequency) ANOVA was conducted. Results showed a significant effect of species $[F(3,8)=38.82,\ p<0.001]$, fundamental frequency $[F(4,32)=53.63,\ p<0.001]$, and a significant interaction between the variables $[F(12,32)=14.12,\ p<0.001]$. *Post-hoc* tests using a Bonferroni *t*-test showed that all three species of birds were better than the humans ($p<0.05$).

Clearly, birds have better resolution of temporal fine structure than humans, notwithstanding some earlier reports of similar performance on other temporal processing tasks. As far as we know, these are the only comparative data available which directly address the question of sensitivity to temporal fine structure in complex sounds and they raise questions about both the perceptual differences among birds and between birds and humans. For this reason, it is of interest to explore the possible mechanisms underlying these species differences by using physiological techniques that may localize responses within the peripheral auditory system.

IV. EXPERIMENT 3—COCHLEAR MICROPHONIC AND COMPOUND ACTION POTENTIALS

The locus of the mechanisms underlying the species' differences in behavioral discrimination thresholds for temporal fine structure are unclear, but there are dramatic differences between birds and mammals at all levels of the auditory system (Carr and Code, 2000; Gleich and Manley, 2000). In fact, much of the past evidence for enhanced temporal resolution in birds has been indirect, with explanations that appeal to the fine structure of vocalizations or invoke either anatomical and physiological aspects of cochlear processing (Greenewalt, 1968; Konishi, 1969; Pumphrey, 1961; Schwartzkopff, 1968). As a first step in exploring the physiological bases of species differences observed psychophysi-

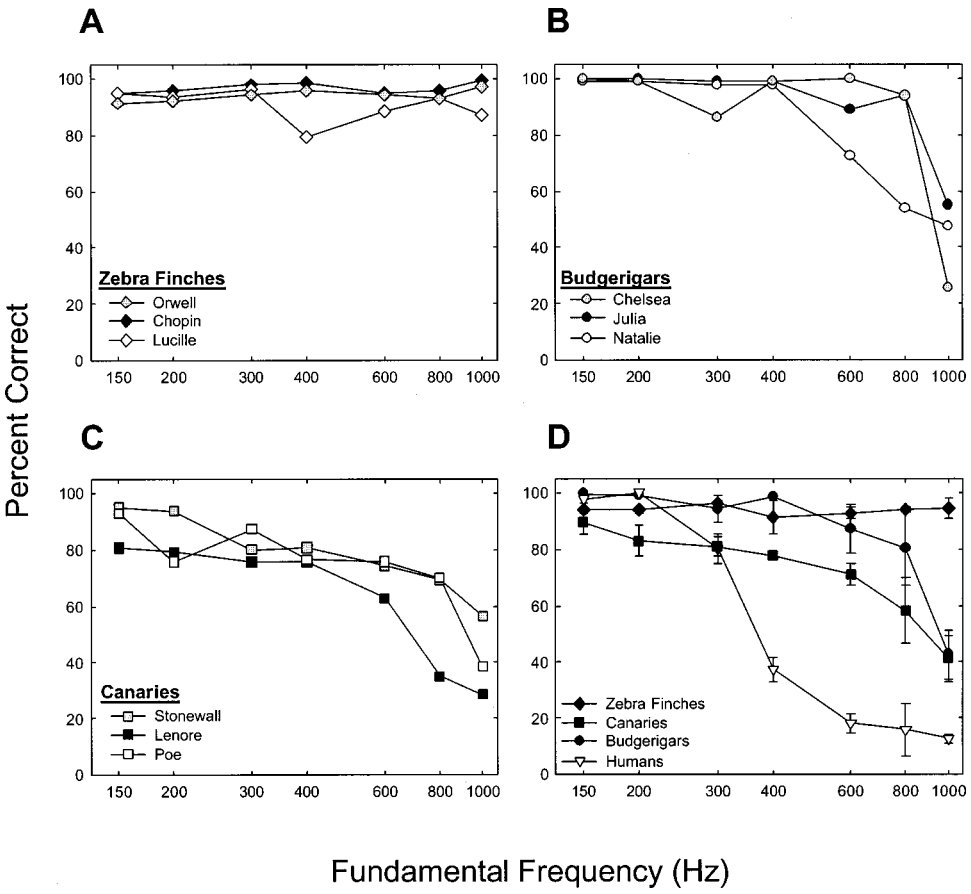

FIG. 5. Results from individual zebra finches (a), budgerigars (b), and canaries (c), tested on the positive/negative Schroeder waveform discrimination at different fundamental frequencies. Average results from the three species of birds and three humans are shown in (d). Error bars are the between-subject standard errors.

cally in experiments 1 and 2, cochlear microphonic potentials (CMs) and compound action potentials (CAPs) from the VIIIth nerve were recorded from the round window in the birds' ears in response to some of the same stimuli used in the behavioral discrimination studies. For a mammalian comparison, similar measures were also taken in the Mongolian gerbil (*Meriones unguiculatus*). If the negative and positive Schroeder complexes that were discriminable at high fundamental frequencies in the behavioral studies with birds do, in fact, generate differential cochlear microphonics or compound action potentials in birds, but not in gerbils, it would suggest that the search for mechanisms accounting for the differences observed behaviorally between birds and humans might begin with consideration of the auditory periphery.

A. Materials and methods

1. Subjects

Two budgerigars, two zebra finches, two canaries, and two Mongolian gerbils were used in these experiments.

2. Procedures

Birds were anesthetized with injections of 20 mg/kg xylazine and 40 mg/kg ketamine into the breast muscle, and gerbils with 15 mg/kg xylazine and 60 mg/kg ketamine injected intraperitoneally. Additional doses of anesthetics (50% of the initial dose) were supplemented as needed, determined either by occasional foot pinch or increased muscle noise in the electrical recordings (generally every 30 to 60 min).

Surgical procedures to gain access to the cochlea have been previously described in detail for birds (Manley *et al.*, 1985) and gerbils (McGuirt *et al.*, 1995). Feathers and hair were removed from the head and around the external ear opening. An incision in the skin along the midline of the skull exposed the bone and it was cleared of connective tissue and dried. A small screw was cemented on the surface of the skull with dental cement to allow precise and stable placement of the head in a holding device and reproducible positioning of the ear canal opening in relation to the speaker.

In birds, the base of the cochlea was exposed by a dorsolateral approach. The core of a thin teflon coated silver wire (0.005 in., WPI) was exposed at the end and inserted through a tiny hole in scala tympani to give direct electrical access to the perilymph. The teflon insulation sealed the hole and prevented leak of perilymph. A subdermal needle inserted into a neighboring neck muscle served as reference and grounding was performed by a needle inserted into the skin of the neck further caudal (standard platinum alloy, subdermal needle electrodes, Grass; E-2B). The bird was placed in a body-shaped styrofoam block with a temperature sensor in contact with the breast muscle. A dc heating pad was used to cover the bird and keep the body temperature close to 40 °C.

In gerbils, an incision behind the external ear and careful dissection of the muscle layer exposed the bulla and a small piece of bone was removed from the bulla to gain access to the niche of the round window. The exposed end of a teflon

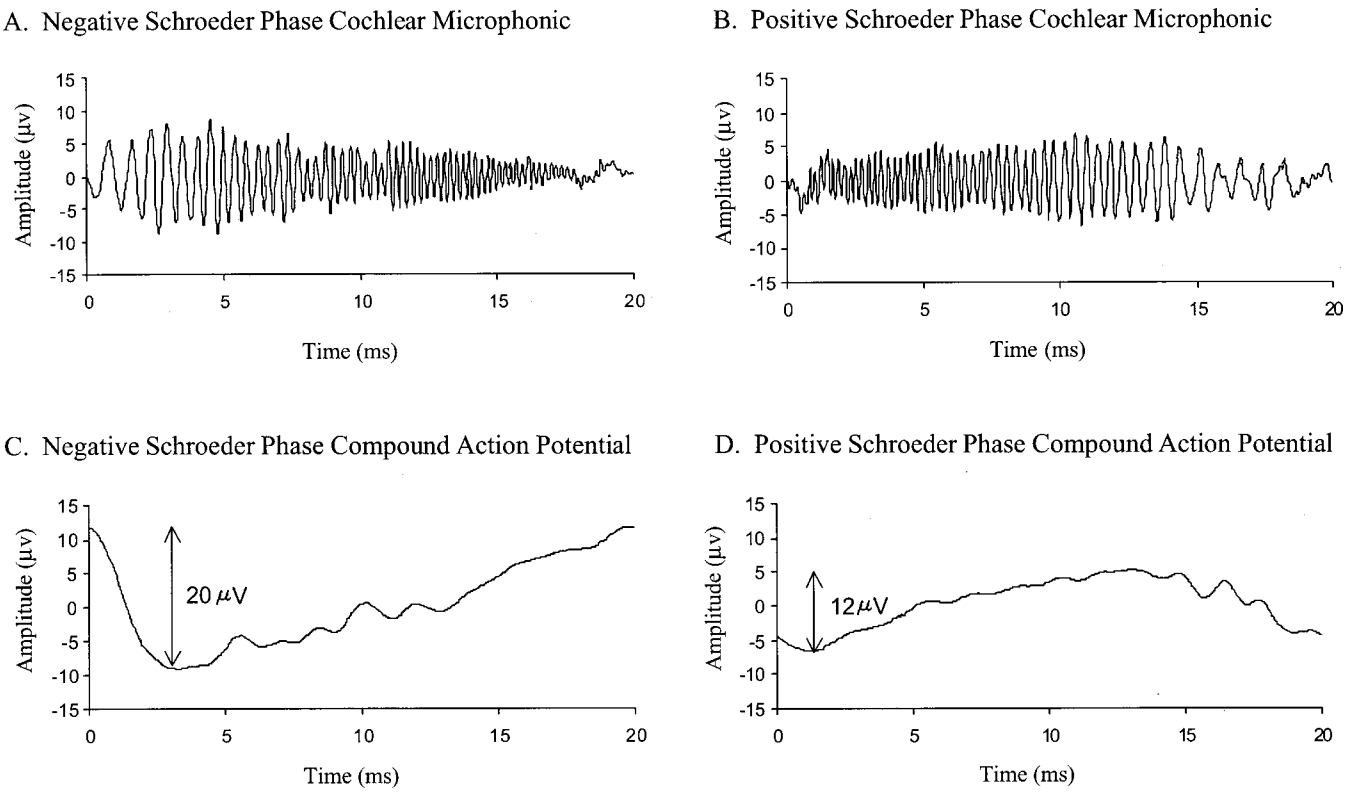

FIG. 6. Examples of CM (a,b) and CAP (c,d) responses to positive and negative Schroeder harmonic complexes, collapsed across periods, for one budgerigar.

insulated silver wire was hooked to the bony ridge of the round window niche and fixed with super glue to the exposed external surface of the bulla. Reference and grounding electrodes were placed as in birds. As with the birds, gerbils were also placed onto a styrofoam block in direct contact with the temperature sensor and covered by the dc heating blanket. The temperature was kept between 37 °C and 38 °C.

Cochlear responses were recorded with a low impedance digital amplifier (TDT, HS4/DB4) using the 60-Hz noise rejection, but no additional filtering. All subjects were placed in the sound field such that their head was at a precise location in the free field 40 cm from the speaker. The stimulus waveforms were fed through a DA1 digital-analog converter, a PA4 programmable attenuator, and a HB6 transducer, which directly drove the speaker (KEF SP 3235, Model 60S, KEF Electronics of America Inc., Holliston, MA). The electrodes were connected to the HS4 Headstage that amplified and digitized the signal before sending it over fiber optic cables to the DB4 Digital Biological Amplifier. A TG6 timing generator was used to synchronize A/D and D/A conversion at a sample rate of 40 kHz.

3. Stimuli

The stimuli were the same as those used in the Schroeder-phase behavioral experiments. The fundamental frequencies of the stimuli ranged from 50 to 1000 Hz. In addition to this set of normal stimuli, a set of inverted stimuli was generated to isolate cochlear microphonic (CM) and compound action potential (CAP) responses, as described below. They were played at 80 dB SPL, and calibrated at the location of the animal's head during the experiment as described for the behavioral experiments. Each stimulus was also recorded using a small microphone placed at that location, and the outputs of the speakers in response to each stimulus waveform were verified off-line.

4. Data collection and reduction

The response to each stimulus was averaged over 124 presentations repeated at a rate of 2/s. After each normal signal presentation, the response to the inverted version of the stimulus was recorded. The CAP response component was derived by adding the response traces obtained for the normal and the inverted stimulus (canceling the CM component) and scaling the resulting response amplitude by half. This derived neural response component was then subtracted from the response trace to the normal stimulus to derive the CM response.

The potentials in this experiment were then further prepared as follows. Taking a response segment, which included the plateau of the stimulus but omitted the rise/fall time (i.e., 220 ms in the middle of the response from 20 to 240 ms within the total 260-ms stimulus), an average CM and CAP period response was obtained by averaging across fundamental periods of the 100-ms segment of the response waveform. From this averaged period response, both the root mean square (rms) of the cochlear microphonic and the peak-to-peak amplitude of the CAP were examined. Examples of a CM and CAP from a budgerigar in response to a negative-Schroeder harmonic complex and a positive-Schroeder harmonic complex are shown in Figs. 6(a)–(d).

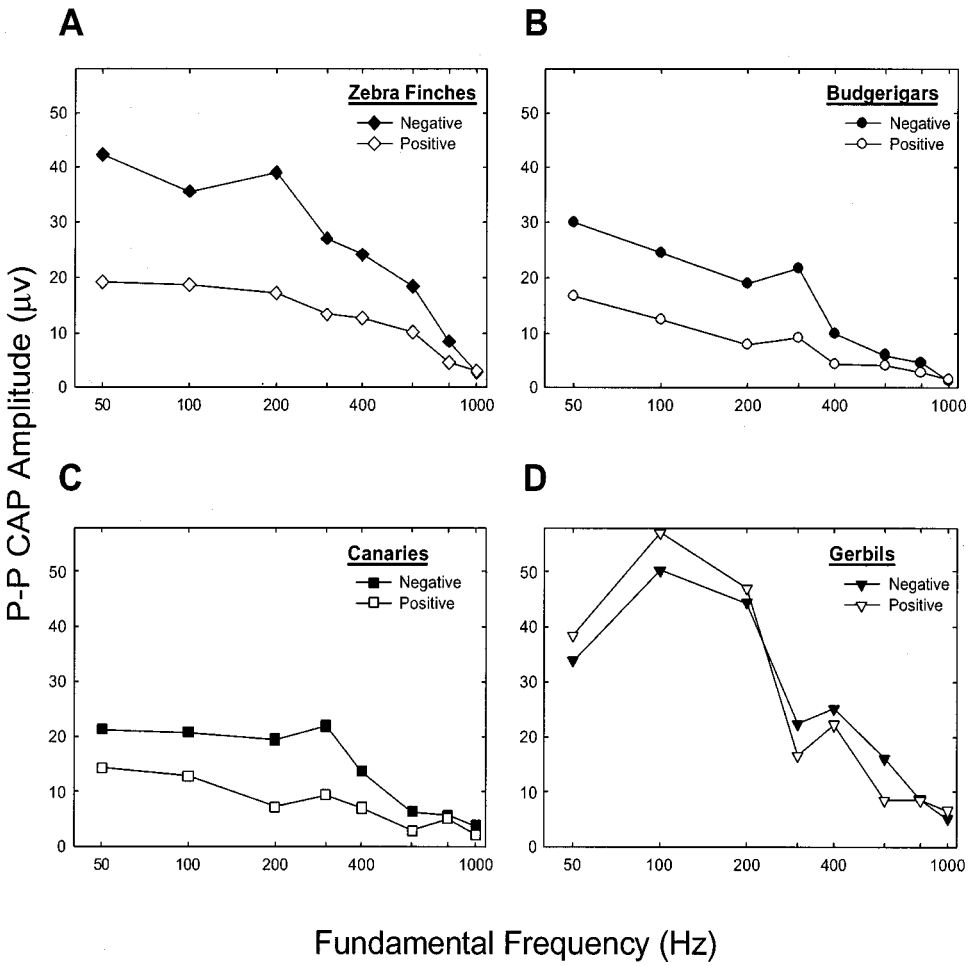

FIG. 7. Average CAP amplitude responses to positive and negative Schroeder complexes as a function of fundamental frequency for zebra finches (a), budgerigars (b), canaries (c), and gerbils (d).

B. Results

In all four species, the CM shape approximately followed the acoustic waveform shape, and the CM amplitude was independent of the fundamental frequency [see Figs. 6(a) and (b)]. In general, gerbils showed much larger CM amplitudes than birds. There was no difference in CM amplitude between negative- and positive-phase waveforms within each fundamental frequency pair for either birds or gerbils. A three-way ANOVA on the rms amplitude values of the cochlear microphonic showed that there was a significant effect of species [$F(3,64)=266.57$, $p<0.001$] but no effect of fundamental frequency [$F(7,64)=0.79$, $p>0.05$] or phase [$F(1,64)=0.97$, $p>0.05$]. No interactions were significant ($p>0.05$). *Post-hoc* tests using a Bonferroni *t*-test showed that rms amplitude of the CM in gerbils was significantly different from that in all species of birds ($p<0.05$), but was not significantly different across bird species ($p>0.05$).

The amplitude of the CAP, however, did vary in several interesting ways. There was a significant decrease for all species in the amplitude of the CAP with increasing fundamental frequency [Figs. 7(a)–(d)]. Gerbils showed no systematic difference in the amplitude of the CAP to positive versus negative Schroeder complexes over the entire range of fundamental frequencies, while all three species of birds showed a much larger CAP to the negative Schroeder complexes than the positive complexes at low fundamental frequencies. A three-way ANOVA showed a significant effect of species [$F(3,64)=21.92$, $p<0.001$], phase [$F(1,64)=35.24$, $p<0.001$], and fundamental frequency [$F(7,64)=27.69$, $p<0.001$]. The interaction between species and fundamental frequency was also significant [$F(3,64)=3.16$, $p<0.001$], but none of the other interactions was significant ($p>0.05$).

The results from the CAP measurements differ from the behavioral measures in that at the highest fundamental frequency tested (1000 Hz), there were no differences between responses to positive and negative Schroeder stimuli for any of the three species. However, paralleling the behavioral responses, as fundamental frequency increased, the differential response to positive and negative Schroeder waveforms persisted longer in finches (800–1000 Hz) than canaries (600–800 Hz), or budgerigars (400–600 Hz). This suggests that CAP responses may be partially related to the discrimination precision observed behaviorally, but that there are contributions from other peripheral or higher auditory processes.

V. DISCUSSION

We have shown that birds can discriminate between synthetic harmonic complexes that differ only in temporal fine structure over extremely short fundamental periods, and that they demonstrate differences in the VIIIth nerve compound action potentials that support this detail of auditory analysis in the synchronization of neural firing. Overall, the three species of birds were able to discriminate between several

types of harmonic complexes with higher fundamental frequencies than humans. These experiments show that birds have an ability to discriminate the temporal fine structure of complex sounds that is two to three times better than the limits shown for humans.

These results are surprising in view of the comparative data showing that birds are similar to mammals in temporal resolving power (see Dooling *et al.*, 2000). Only in a few of these measures is there a hint of consistent differences between birds and mammals. A comparison of temporal modulation transfer functions (TMTFs) of several mammals and several species of birds reveals that humans are more sensitive to modulation-based changes in intensity than birds at low modulation frequencies. Although birds and humans show a similar cutoff of performance at high modulation frequencies, the difference in intensity resolving power at low modulation rates results in a shorter time constant for birds compared to humans (Dooling *et al.*, 2000; Dooling and Searcy, 1981). Gap detection thresholds at very low intensity levels are also generally better in birds than mammals (Dooling *et al.*, 2000), probably reflecting the tendency of mammalian tuning curves to narrow with decreasing sound pressure levels. Taken together, results point to a need for future investigation into the possibility that enhanced sensitivity to temporal fine structure of complex sounds may be a distinguishing feature of the avian auditory system. A more complete evaluation of this hypothesis would require testing, along with more birds, many more mammals, especially those with auditory systems that appear to be specialized for processing auditory temporal information as, for instance, many species of bats (see review in Moss and Simmons, 1996).

Gerbils are becoming popular models of mammalian hearing for both simple and complex sounds (Heffner and Heffner, 1988; Ryan, 1976; Sinnott and Mosteller, 2001), though it is still the case that much less is known about their auditory capabilities than the birds in these experiments. Thus, it is not clear why the physiological data taken from the gerbils do not match the expectations based on behavioral data from the other mammal used in these experiments (humans). Behavioral data from humans suggests easy discrimination between the positive- and negative-Schroeder-phase waveforms at low fundamental frequencies, but the compound action potentials presented here for the gerbil show little difference between the two types of waveforms. Clearly, these results stand in stark contrast to findings for all three bird species that show a difference in CAP amplitude up to nearly 800 Hz paralleling the behavioral performance. The difference in CAP between birds and gerbils clearly demonstrates a difference in cochlear processing. Behavioral experiments are needed to determine whether gerbils truly cannot discriminate sounds differing in only temporal fine structure at the low fundamental frequencies used in our physiological tests. If they can make such discriminations behaviorally, then the coding of temporal fine structure must involve something other than the synchronous firing of the VIIIth nerve in gerbil. Alternatively, the lack of CAP asymmetry in gerbil might have no relation to the gerbil's ability to discriminate temporal fine structure in behavioral experiments.

A. Mechanisms underlying temporal resolution

Cochlear frequency analysis in mammals is often modeled as classical linear filtering with the broader filter bandwidths found in higher frequency regions supporting better temporal analysis and narrower filters in lower frequency regions resulting in poorer temporal analysis (Viemeister and Plack, 1993). Individual frequency components of a harmonic complex that are more widely spaced than the frequency analyzing channels in the auditory system significantly reduce the ability to perceive the temporal characteristics of the complex. When the fundamental frequency, and therefore component spacing, is such that several components fall within an auditory channel, the temporal properties of the sound can be used for discrimination. Discrimination of the random from the cosine phase waves and between the Schroeder-phase complexes demonstrated by the human listeners in this study reflected this limitation, as has been described repeatedly in the literature using tasks such as phase discrimination in three-component complexes (e.g., Goldstein, 1967), perception of pitches generated by spectral edges in harmonic complexes (Kohlrausch and Houtsma, 1992), and other studies of random versus cosine-phase discrimination in harmonic complexes (Patterson, 1987). These reports in the literature, as well as the data presented here, suggest that the ability to discriminate temporal cues in these harmonic stimuli is severely compromised when the fundamental frequency is greater than about 400–500 Hz.

Interestingly, the above relationship between component spacing, filter bandwidths, and temporal resolving power seems not to hold for the birds. By most accounts, birds have cochlear frequency resolution slightly worse than observed in humans (Dooling *et al.*, 2000; Sachs *et al.*, 1978) but tuning curves of some bird auditory nerve fibers are actually more sharply tuned than some mammals like cats and guinea pigs (see review in Gleich and Manley, 2000). Human frequency resolution estimated from suppression of distortion product otoacoustic emissions indicates resolution as good as or slightly worse than avian frequency resolution (see review in Gleich and Manley, 2000). Thus, a frequency domain analysis in the avian cochlea at reported levels conflicts with the ability of birds to make these waveform discriminations at fundamental frequencies exceeding 600 Hz. At such high fundamental frequencies, birds are making discriminations between two sounds that differ in fine structure over time intervals as small as 1 ms—much faster than any estimate of monaural temporal resolution capacity of humans. This basic difference in perception between humans and birds calls into question conventional views of the relationship between frequency resolution and temporal acuity within harmonic complexes, and points to the need for further comparative investigations of pitch, timbre, and frequency perception across species.

B. The relationship between perception and production

One reason that the differences among the species of birds in the ability to discriminate between the waveform shapes of harmonic complexes are intriguing is because of differences among the species in the characteristics of their species-specific vocalizations. Zebra finches are the only one of these three bird species whose calls and songs are strongly and almost exclusively harmonic, with fundamental frequencies approaching 600–700 Hz (see, for example, Zann, 1984). Interestingly, much is also known about the acoustic characteristics of the vocalizations of canaries and budgerigars (see, for example, Farabaugh and Dooling, 1996; Guttinger *et al.*, 1978; Lavenex, 1999). From these studies, as well as from our own casual observation of songs and calls of these species, it is clear that zebra finch calls have the most broadband harmonic structure, and that budgerigar vocalizations in general contain a greater proportion of such acoustic structure than do canary vocalizations. This order parallels the species differences in the ability to discriminate temporal fine structure over exceedingly small time windows, with zebra finches showing the largest perceptual and CAP differences between waveform shapes among the birds.

These apparent parallels between perception and production shown here suggest further comparative and developmental studies may be useful. It might be that these parallels arise through some form of co-evolution as has been suggested for human speech (Liberman *et al.*, 1967). One can imagine various ontogenetic strategies. One possibility is that the avian ear and auditory system in general excel at discrimination of temporal fine structure. Species differences may emerge as developing vocalizations fail to exploit initial sensitivities during critical periods, and these sensitivities are then lost in adulthood. A similar argument has been made regarding the development of human languages in infants by Kuhl and her colleagues (e.g., Kuhl, 2000). Another possibility is that the avian ear and auditory system remain highly plastic throughout life and that some bird species may learn that relevant information is coded in the fine temporal structure of various signals, leading to an enhanced sensitivity to temporal variation for these species.

With the exception of sound localization in birds, these data are the first to demonstrate behaviorally the use of fine temporal processing abilities for perception of acoustic stimuli with some similarities to communication signals. These findings of enhanced temporal resolution may have particular relevance for understanding acoustic communication, song learning, and individual recognition in birds. For zebra finches, in particular, confirmation of precise temporal coding in the auditory system (Janata and Margoliash, 1999; Sen *et al.*, 2001; Theunissen and Doupe, 1998) parallels recent findings of precise coding of temporal information in the motor circuits and syringeal machinery controlling song in this species (Brainard and Doupe, 2001; Fee *et al.*, 1998; Tchernichovski *et al.*, 2001; Vicario, 1991; Yu and Margoliash, 1996).

In summary, these data show temporal resolution in the processing of acoustic communication signals in birds well beyond the limits typically reported for humans and other mammals. Synchronized activity in the VIIIth nerve also reveals a sensitivity to waveform shape of such stimuli that may be unique to birds, suggesting a radical and unexpected difference in the coding of these stimuli in the peripheral auditory system of birds and mammals. At the level of acoustic communication, the analysis of bird vocalizations is usually done in the spectral domain and rarely extends to the level of temporal fine structure. The size of a species' communication repertoire and the degree of vocal learning is judged by these traditional types of analyses. If it turns out that birds typically perceive and make use of fine temporal detail in their complex vocalizations that is beyond the range of human capabilities, they may have much larger vocabularies than previously thought.

ACKNOWLEDGMENTS

This work was supported by NIH Grant No. DC-00198 to RJD and No. DC-00626 to MRL. MLD was supported by NSRAs from NIH (DC-00046 and MH-12698). We are very grateful to J. Cucina for his assistance in analyzing the physiological data. Special thanks also to B. Brittan-Powell, B. Lohr, and M. Mavilia for their help on various aspects of these studies. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

ANSI (**1989**). ANSI S3.6-1989, "Specifications for audiometers" (American National Standards Institute, New York).

Brainard, M. S., and Doupe, A. J. (**2001**). "Postlearning consolidation of birdsong: stabilizing effects of age and anterior forebrain lesions," J. Neurosci. **21**, 2501–2517.

Carr, C. E., and Code, R. A. (**2000**). "The central auditory system of reptiles and birds," in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 197–248.

Carr, C. E., and Friedman, M. A. (**1999**). "Evolution of time coding systems," Neural Comput. **11**, 1–20.

Dooling, R. J. (**1982**). "Auditory perception in birds," in *Acoustic Communication in Birds, Vol. 1*, edited by D. E. Kroodsma and E. H. Miller (Academic, New York), pp. 95–130.

Dooling, R. J., and Haskell, R. J. (**1978**). "Auditory duration discrimination in the parakeet (*Melopsittacus undulatus*)," J. Acoust. Soc. Am. **63**, 1640–1643.

Dooling, R. J., and Searcy, M. H. (**1981**). "Amplitude modulation thresholds for the parakeet (*Melopsittacus undulatus*)," J. Comp. Physiol. **143**, 383–388.

Dooling, R. J., and Searcy, M. H. (**1985**). "Temporal integration of acoustic signals by the budgerigar (*Melopsittacus undulatus*)," J. Acoust. Soc. Am. **77**, 1917–1920.

Dooling, R. J., Lohr, B., and Dent, M. L. (**2000**). "Hearing in birds and reptiles," in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 308–359.

Dooling, R. J., Dent, M. L., Leek, M. R., and Gleich, O. (**2001**). "Masking by harmonic complexes in three species of birds: Psychophysical thresholds and cochlear responses," Hear. Res. **152**, 159–172.

Farabaugh, S. M., and Dooling, R. J. (**1996**). "Acoustic Communication in Parrots: Laboratory and Field Studies of Budgerigars, *Melopsittacus undulatus*," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell U.P., Ithaca, NY), pp. 97–117.

Fay, R. R. (**1988**). *Hearing in Vertebrates: A Psychophysics Databook* (Hill-Fay, Winnetka, IL).

Fee, M. S., Shraiman, B., Pesaran, B., and Mitra, P. P. (**1998**). "The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird," Nature (London) **395**, 67–71.

Gleich, O., and Manley, G. A. (**2000**). ''The hearing organ of birds and crocodilia,'' in *Comparative Hearing: Birds and Reptiles*, edited by R. J. Dooling, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 70–138.
Goldstein, J. L. (**1967**). ''Auditory spectral filtering and monaural phase perception,'' J. Acoust. Soc. Am. **41**, 458–479.
Greenewalt, C. H. (**1968**). *Bird Song: Acoustics and Physiology* (Smithsonian Institute, Washington, DC).
Guttinger, H. R., Wolffgramm, J., and Thimm, F. (**1978**). ''The relationship between species specific song programs and individual learning in songbirds: A study of individual variation in songs of canaries, greenfinches, and hybrids between the two species,'' Behaviour **65**, 241–262.
Heffner, R., and Heffner, H. (**1988**). ''Sound localization and the use of binaural cues in the gerbil,'' Behav. Neurosci. **102**, 422–428.
Janata, P., and Margoliash, D. (**1999**). ''Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system,'' J. Neurosci. **19**, 5108–5118.
Klump, G. M., and Maier, E. H. (**1989**). ''Gap detection in the starling (*Sturnus vulgaris*), I: Psychophysical thresholds,'' J. Comp. Physiol. **164**, 531–539.
Kohlrausch, A., and Houtsma, A. J. M. (**1992**). ''Pitch related to spectral edges of broadband signals,'' Philos. Trans. R. Soc. London, Ser. B **336**, 375–381.
Konishi, M. (**1969**). ''Time resolution by single auditory neurons in birds,'' Nature (London) **222**, 566–567.
Kroodsma, D. E., and Miller, E. H. (**1982**). *Acoustic Communication in Birds, Vol. 2* (Academic, New York).
Kroodsma, D. E., and Miller, E. H. (**1996**). *Ecology and Evolution of Acoustic Communication in Birds* (Cornell U.P., Ithaca, NY).
Kuhl, P. K. (**2000**). ''A new view of language acquisition,'' Proc. Natl. Acad. Sci. U.S.A. **97**, 11850–11857.
Lavenex, P. B. (**1999**). ''Vocal production mechanisms in the budgerigar (*Melopsittacus undulatus*): The presence and implications of amplitude modulation,'' J. Acoust. Soc. Am. **106**, 491–505.
Leek, M. R., Dent, M. L., and Dooling, R. J. (**2000**). ''Masking by harmonic complexes in budgerigars (*Melopsittacus undulatus*),'' J. Acoust. Soc. Am. **107**, 1737–1744.
Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (**1967**). ''Perception of the speech code,'' Psychol. Rev. **74**, 431–461.
Manley, G. A., Gleich, O., Oeckinghaus, H., and Leppelsack, H.-J. (**1985**). ''Activity patterns of cochlear ganglion neurones in the starling,'' J. Comp. Physiol., A **157**, 161–181.
McGuirt, J. P., Schmiedt, R. A., and Schulte, B. A. (**1995**). ''Development of cochlear potentials in the neonatal gerbil,'' Hear. Res. **84**, 52–60.
Moss, C. F., and Simmons, J. A. (**1996**). ''Perception along the axis of target range in the echolocating bat,'' in *Neuroethological Studies of Cognitive and Perceptual Processes*, edited by C. F. Moss and S. J. Shettleworth (Westview, Boulder, CO), pp. 253–279.
Patterson, R. D. (**1987**). ''A pulse ribbon model of monaural phase perception,'' J. Acoust. Soc. Am. **82**, 1560–1586.
Pumphrey, R. J. (**1961**). ''Sensory Organs: Hearing,'' in *Biology and Comparative Anatomy of Birds*, edited by A. J. Marshall (Academic, New York), pp. 69–86.
Ryan, A. (**1976**). ''Hearing sensitivity of the Mongolian gerbil,'' J. Acoust. Soc. Am. **59**, 1222–1226.
Sachs, M. B., Sinnott, J. M., and Hienz, R. D. (**1978**). ''Behavioral and physiological studies of hearing in birds,'' Fed. Proc. **37**, 2329–2335.
Schroeder, M. R. (**1970**). ''Synthesis of low-peak-factor signals and binary sequences with low autocorrelation,'' IEEE Trans. Inf. Theory **16**, 85–89.
Schwartzkopff, J. (**1968**). ''Structure and function of the ear and the auditory brain areas in birds,'' in *Hearing Mechanisms in Vertebrates*, edited by A. V. S. De Reuck and J. Knight (Little, Brown, Boston), pp. 41–59.
Sen, K., Theunissen, F. E., and Doupe, A. J. (**2001**). ''Feature analysis of natural sounds in the songbird auditory forebrain,'' J. Neurophysiol. **86**, 1445–1458.
Sinnott, J. M., and Mosteller, K. W. (**2001**). ''A comparative assessment of speech sound discrimination in the Mongolian gerbil,'' J. Acoust. Soc. Am. **110**, 1729–1732.
Tchernichovski, O., Mitra, P. P., Lints, T., and Nottebohm, F. (**2001**). ''Dynamics of the vocal imitation process: How a zebra finch learns its song,'' Science **291**, 2564–2569.
Theunissen, F. E., and Doupe, A. J. (**1998**). ''Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVc of male zebra finches,'' J. Neurosci. **18**, 3786–3802.
Vicario, D. S. (**1991**). ''Contributions of syringeal muscles to respiration and vocalization in the zebra finch,'' J. Neurobiol. **22**, 63–73.
Viemeister, N. F., and Plack, C. J. (**1993**). ''Time analysis,'' in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer-Verlag, New York), pp. 116–154.
Yu, A. C., and Margoliash, D. (**1996**). ''Temporal hierarchical control of singing in birds,'' Science **273**, 1801–2.
Zann, R. (**1984**). ''Structural variation in the zebra finch distance call,'' Z. Tierpsychol. **66**, 328–345.