

# Phase effects in masking by harmonic complexes in birds<sup>a)</sup>

Amanda M. Lauer<sup>b)</sup> and Robert J. Dooling

*Department of Psychology, University of Maryland, College Park, Maryland 20742 and Center for the Comparative and Evolutionary Biology of Hearing, University of Maryland, College Park, Maryland 20742*

Marjorie R. Leek

*Army Audiology and Speech Center, Walter Reed Army Medical Center, Washington, DC 20307-5001 and Center for the Comparative and Evolutionary Biology of Hearing, University of Maryland, College Park, Maryland 20742*

Jennifer J. Lentz

*Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana 47405*

(Received 3 May 2005; revised 14 October 2005; accepted 23 November 2005)

Masking by harmonic complexes depends on the frequency content of the masker and its phase spectrum. Harmonic complexes created with negative Schroeder phases (component phases decreasing with increasing frequency) produce more masking than those with positive Schroeder phases (increasing phase) in humans, but not in birds. The masking differences in humans have been attributed to interactions between the masker phase spectrum and the phase characteristic of the basilar membrane. In birds, the similarity in masking by positive and negative Schroeder maskers, and reduced masking by cosine-phase maskers (constant phase), suggests a phase characteristic that does not change much along the basilar papilla. To evaluate this possibility, the rate of phase change across masker bandwidth was varied by systematically altering the Schroeder algorithm. Humans and three species of birds detected tones added in phase to a single component of a harmonic complex. As observed in earlier studies, the minimum amount of masking in humans occurred for positive phase gradients. However, minimum masking in birds occurred for a shallow negative phase gradient. These results suggest a cochlear delay in birds that is reduced compared to that found in humans, probably related to the shorter avian basilar epithelia. © 2006 Acoustical Society of America. [DOI: 10.1121/1.2151816]

PACS number(s): 43.80.Lb, 43.66.Gf [JAS]

Pages: 1251–1259

## I. INTRODUCTION

While the representation of the spectral characteristics of sound along the avian basilar papilla is fairly well understood, less is known about how temporal information is processed along the papilla. Yet, the acoustic communication system of birds involves some of the most temporally complex acoustic signals in nature and it is becoming increasingly clear that birds perceive much of this complexity. In mammals, there is considerable evidence of an interaction between the spectral and temporal characteristics of a sound waveform and the response characteristics of the basilar membrane (e.g., Reccio and Rhode 2000). The temporal waveform shape may also influence the internal representation of sound on the bird basilar papilla, but this interaction between stimulus and response characteristics of the papilla has been less well explored. The aim of the present study was to examine how temporal response properties of the avian basilar papilla interact with the waveform shape and spectral characteristics of complex sounds using behavioral masking methods. Masking of a probe tone was measured for harmonic complex maskers with varying phase spectra.

Component phases were selected to provide systematic changes in the temporal waveforms of the maskers without altering the long-term amplitude spectrum.

Variations in the phase spectrum of a harmonic complex sound may produce large differences in waveform shape and in the temporal pattern of instantaneous frequencies within the waveforms. The maskers used in this study were constructed with harmonic component starting phases selected according to a modification of an algorithm developed by Schroeder (1970)

$$\theta_n = C\pi n(n+1)/N, \quad (1)$$

where  $\theta_n$  represents the phase of the  $n$ th harmonic,  $N$  is the total number of harmonics, and  $C$  is a scalar (Lentz and Leek 2001). Phase spectra for several of the complexes are depicted in Fig. 1(a). The Schroeder-phase stimuli feature monotonic changes in phase across frequency that produce upward or downward sweeps in instantaneous frequency within each period of the complex, as illustrated in Fig. 1(b). The direction of frequency sweep may be reversed by altering the sign of the Schroeder algorithm [left versus right columns in Fig. 1(b)]. The speed of frequency sweep may be manipulated by the choice of constant scalar in the phase selection algorithm (Lentz and Leek 2001). Complexes made with the same scalar value, but opposite sign, result in time

<sup>a)</sup>Portions of this work were presented at the 141st Meeting of the Acoustical Society of America, Chicago, IL, June 2001 and the 6th International Congress on Neuroethology, Bonn, Germany, August 2001.

<sup>b)</sup>Electronic mail: alauer@psyc.umd.edu

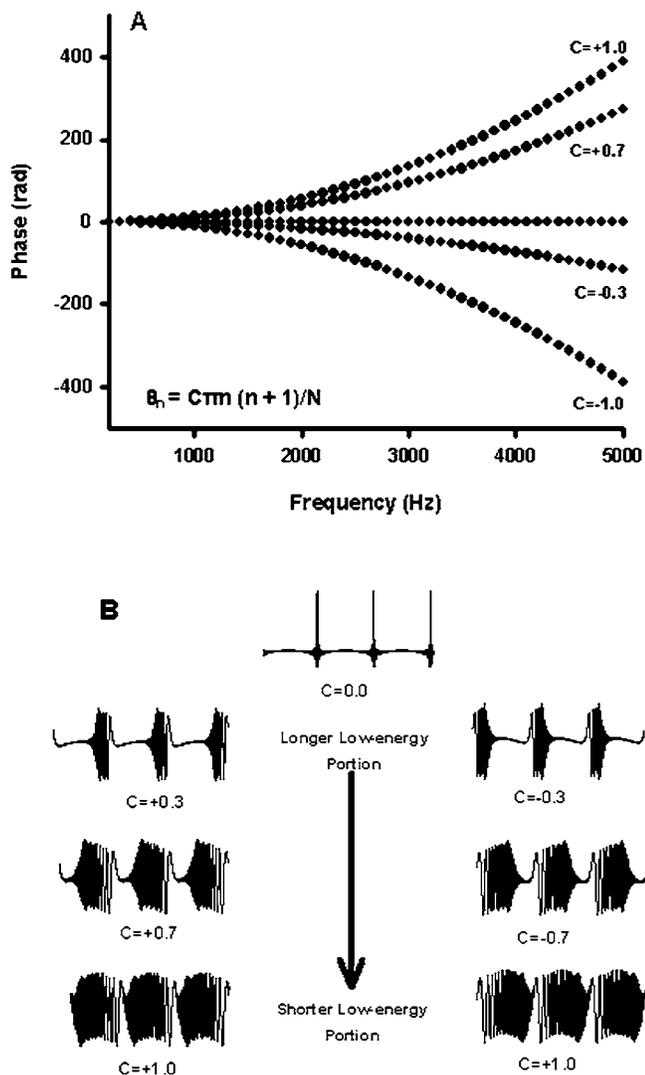


FIG. 1. Phase spectra (a) and temporal waveforms (b) for several of the harmonic complex maskers. Each waveform is constructed of equal-amplitude cosine tones from 200 to 5000 Hz, with a fundamental frequency of 100 Hz. Phases were selected according to scaled modifications of the algorithm shown in the lower left corner of (a) ( $\theta_n$  = the  $n$ th harmonic,  $N$  = total number of harmonics, and  $C$  = a scalar). Three periods (30 ms) are shown for each temporal waveform.

reversed waveforms, with phases increasing (positive Schroeder) or decreasing (negative Schroeder) and instantaneous frequency decreasing or increasing, respectively.

Because none of these phase manipulations has any effect on the long-term frequency spectra, a theory of masking based entirely on the spectrum would predict equal masking from all of these Schroeder complexes. However, the original negative and positive Schroeder-phase maskers (i.e., those with a scalar  $C$  of  $\pm 1.0$ ) can produce large masking differences in humans (Smith *et al.* 1986; Kohlrausch and Sander 1995; Carlyon and Datta 1997a; 1997b; Summers and Leek 1998; Lentz and Leek 2001; Oxenham and Dau 2001). This effect has been attributed to cochlear processing mechanisms that may interact with the stimulus phase spectrum to produce an altered “internal” waveform. Kohlrausch and Sander (1995) proposed that the mammalian basilar membrane alters the shape of the positive Schroeder-phase masker such that the corresponding internal waveform is

more modulated than the internal waveform corresponding to the negative Schroeder-phase masker. This increased modulation creates portions of low energy within each period in the positive Schroeder internal waveform within which a signal tone may be more easily detected in a masking task. Negative Schroeder-phase maskers do not undergo such alteration in internal shape, and therefore masking is greater. Hearing-impaired listeners do not show these large masking differences between the positive and negative waveforms (Summers and Leek 1998). This effect is thought to be related to the loss of the nonlinear active processing mechanism in a damaged cochlea.

Recent studies have demonstrated dramatic differences in masking by Schroeder-phase harmonic complexes between birds and humans when the scalar is  $\pm 1.0$ . In contrast to humans, the negative-phase and positive-phase waveforms produce similar amounts of masking of a 2.8 kHz tone in three species of birds: budgerigars, zebra finches, and canaries (Dooling *et al.* 2001; Leek *et al.* 2000). The differences in the patterns of masking in birds and humans have been attributed to structural differences in the mammalian and avian ears. The length of the basilar papilla in birds is an order of magnitude smaller than the human cochlea (Gleich *et al.* 1994; Manley *et al.* 1993). In addition, the stiffness gradient along the basilar papilla is steeper in the avian ear (von Békésy 1960). Finally, the frequency-dependent cochlear delay is much shorter in starlings and pigeons than in most mammals (Gleich and Narins 1988). These anatomical and physiological differences undoubtedly affect the nature of the traveling wave in birds.

To further investigate differences in phase processing by the inner ears in birds and humans, we used harmonic complex maskers with phase spectra selected according to these scaled modifications of the Schroeder algorithm to produce waveforms with envelopes that vary systematically between highly peaked (cosine phase) and very flat. Both positive-phase and negative-phase maskers were tested. Maskers with more modulated envelopes are characterized by faster repeating frequency sweeps (once each period) and longer low energy portions, and maskers with flatter envelopes contain slower repeating frequency sweeps and very short low energy portions. The internal phase characteristic of the basilar membrane (or basilar papilla) in the region of maximum displacement for a particular frequency may be estimated by finding the least effective scaled Schroeder-phase masker (Lentz and Leek 2001). The phase spectrum of this masker is assumed to approximately cancel the phase characteristic of the cochlea in the frequency region of the signal, thereby creating an internal within-channel waveform with a highly peaked shape (effectively, a cosine-phase internal waveform). In this way, psychophysical measures have been used to estimate the phase characteristic of the human basilar membrane by determining the least effective masker in a set of scaled Schroeder-phase stimuli. In humans, the least effective masker varies somewhat across listeners and across frequency, but is always a positive-phase masker for mid- to high-frequency signals (Lentz and Leek 2001). These results

in humans suggest that the internal waveforms corresponding to the positive-phase maskers are more modulated, enabling a probe tone to be more easily detected.

In this study, the phase response of the basilar papilla in birds was investigated in order to infer temporal characteristics of the traveling wave. Budgerigars, zebra finches, and canaries were tested in a masking paradigm using stimuli similar to those used for testing humans (Lentz and Leek 2001; Oxenham and Dau 2001). In addition, as a control, three humans were tested using the same methodologies used in testing birds. Thresholds for detecting tones embedded in waveforms with systematically varying shape were measured using the Method of Constant Stimuli and operant conditioning techniques.

## II. METHODS

### A. Subjects

Three adult zebra finches, three adult budgerigars, and three adult canaries were used as subjects. The birds were kept on a normal day/night cycle correlated with the season and maintained at approximately 90% of their free-feeding weights. For comparison, three young adult humans (laboratory staff members) also participated in the experiment. All birds and humans had hearing within normal limits for their species, as shown by their audiograms. Animal housing and care met all standards of the University of Maryland Animal Care and Use Committee (ACUC), College Park, MD. All research was approved by the ACUC and the Internal Review Board.

### B. Stimuli

Stimuli were harmonic maskers and maskers plus a signal tone. The masking stimuli were constructed by summing equal-amplitude tones from 200 to 5000 Hz, with a fundamental frequency of 100 Hz. The phases of the tones were selected according to a modification of the Schroeder-phase algorithm [Eq. (1)] described earlier. Maskers were generated for scalars ( $C$ ) ranging between  $-1.0$  and  $+1.0$ . These two end-value scalars result in the original Schroeder-negative and Schroeder-positive phase maskers used in earlier masking studies. When  $C=0.0$ , a highly modulated cosine-phase waveform is produced, characterized by a very high peak once each period with low-amplitude energy during the rest of the period. Negatively valued scalar stimuli have a rising frequency sweep within the masker period, and positively valued scalars have a falling frequency sweep within each period. Changing the scalar changes the rate of the frequency sweep, in that scalars closer to zero produce more rapid frequency sweeps than those close to  $\pm 1.0$ . Seven examples of these masker waveforms are depicted in Figure 1(b). The different scalars generate maskers on a continuum of both frequency sweep rate within each period and relative proportion of low versus high energy within each period.

Twenty-one scalars were tested for zebra finches and 13 scalars were tested for budgerigars and canaries. The maskers were 260 ms in duration with 20 ms raised-cosine rise/fall times. The maskers were presented at a level of 80 dB SPL (63 dB SPL per harmonic component). The sig-

nal was a 2.8 kHz tone added in phase to the 2.8 kHz masker component. The duration of the signal was the same as the masker, including the rise and fall times.

Stimuli were created digitally by summing waveforms at the component frequencies with the appropriate phases and amplitudes. All stimuli were created off line and stored as files for playback during the experiment. The sampling rate was 40 kHz. Each set of stimuli included a masker waveform alone, and a number of masker-plus-signal waveforms, at several signal-to-masker ratios. The signal level at threshold is reported as the level of the signal component added to the masker in decibels (dB) relative to the level of each masker component.

### C. Testing Apparatus

The birds were tested in a wire cage ( $23 \times 25 \times 16$  cm) mounted in a sound-isolation chamber (Industrial Acoustics Company, Bronx, NY, 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 the food hopper. Microswitches were tripped when a bird pecked the attached LED. The left microswitch served as the observation key, and the right microswitch served as the report key. During test sessions, the behavior of the animal was monitored by a video camera system (Sony HVM-322).

Test sessions were controlled by a computer (IBM Pentium III). The digital stimuli were output to a KEF loudspeaker (model 80C) via Tucker-Davis modules at masker levels of 80 dB SPL. Stimulus calibration was performed using a Larson-Davis System 824 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 keys in the approximate position occupied by the bird's head during testing. Masker intensities were measured several times during the experiment to ensure that stimulus levels remained constant and the entire system was calibrated.

### D. Training and testing procedures

Birds were trained by standard operant auto-shaping procedures (Dooling and Okanoya 1995) to peck two keys constructed of LEDs attached to microswitches. Birds pecked at the left key (observation key) during a repeating background for a random amount of time between 2 and 7 s until a target stimulus was alternated with the background sound. If the bird pecked the right key (report key) within 2 s of this alternating pattern, it received access to food from a hopper for 1 or 2 s. The dependent variable was percent correct on trials involving an alternating sound pattern. Failure to peck the report key within 2 s of the alternating pattern was recorded as a miss, 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 background sound. A peck to the report key during a sham trial was recorded as a false alarm, and the test chamber lights were extinguished for 5–15 s. Birds typically exhibited false alarm rates between 3 and 10%. Average false alarm rates were 3.50% for budgerigars, 4.77% for zebra finches, and 2.82% for canaries. Data

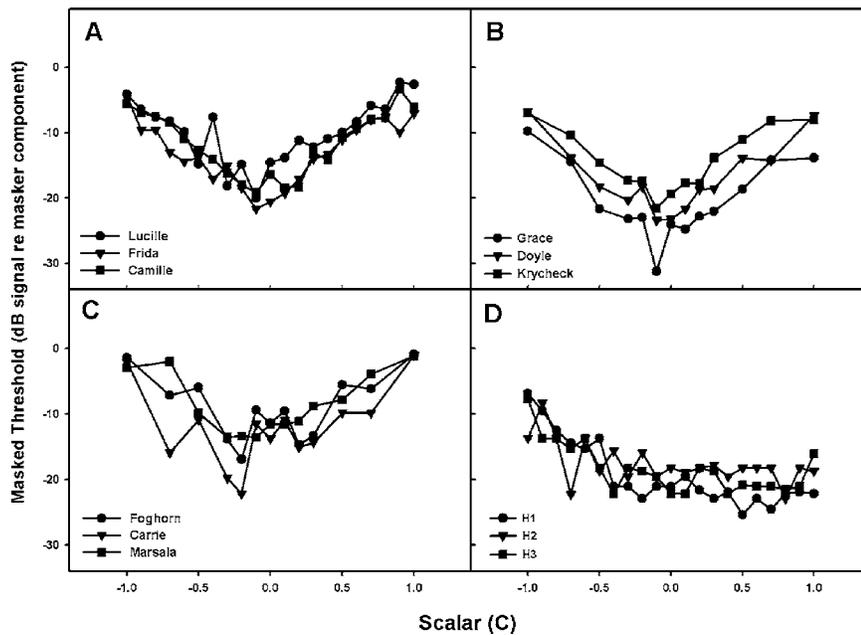


FIG. 2. Thresholds for a 2.8 kHz tone (dB signal re. masker component) embedded in harmonic complex maskers for individual zebra finches (a), budgerigars (b), canaries (c), and humans (d). The abscissa indicates the scalar value used to construct the masker waveform. Values closer to 0.0 correspond to more modulated masker waveforms; values closer to  $\pm 1.0$  correspond to flatter waveforms.

from sessions with false alarm rates higher than 18% were discarded. In all, 0.7% of all sessions for the budgerigars, 5.3% for zebra finches, and 0% for canaries were excluded from analysis.

For each Schroeder-phase scalar masker, signal levels in 0.4, 1, or 2 dB steps were presented using the Method of Constant Stimuli (Dooling and Okanoya 1995). Signal levels within a condition were selected to bracket the presumed threshold, and psychometric functions were developed. Birds ran a minimum of 300 trials for each Schroeder-phase masker, and the last 200 trials once behavior stabilized (threshold did not change more than 1/3 the step size) were used for analysis. Thresholds were defined as the level of the tone detected 50% of the time, adjusted by the false alarm rate [ $Pc^* = (Pc - FA) / (1 - FA)$ ] (Dooling and Okanoya 1995; Gescheider 1985). For comparison, three humans were tested with earphones on the same sounds using similar procedures. In order to estimate a phase response across frequencies, one bird of each species was also tested with signal frequencies of 1.0, 2.0, and 4.0 kHz using several of the maskers.

### III. RESULTS

Figure 2 shows individual masked thresholds for a 2.8 kHz signal in dB (re. the level of each masker component) for zebra finches, budgerigars, canaries, and humans [panels (a) through (d), respectively]. Thresholds for each subject are plotted as a function of scalar value  $C$ . Each bird species shows a general pattern of high thresholds at the scalar extremes (+ or -1.0), with a drop in threshold near the center of the scalar range. The variability of the scalar resulting in minimum masking within bird species is quite small, and across bird species there is a systematic effect of waveform shape produced by different selections of component phase. The pattern shows a release from masking for each bird species for scalars that are just slightly negative (i.e., -0.1 and -0.2). This scalar does not produce the most highly

modulated external waveform. Rather, the most highly modulated waveform is produced by a scalar of 0.0, a cosine-phase wave.

Humans show a different pattern of thresholds across scalar values than do birds. Although there are differences due to phase selection, the minimum masking for humans occurs for maskers with positive scalars. Moreover, the least amounts of masking in humans occur over several positive scalars, resulting in a much broader minimum, on average, than was observed in birds. These results are consistent with the results of an earlier study of humans by Lentz and Leek (2001) who found a minimum masking scalar for humans at a signal frequency of 3 kHz (near the signal frequency tested here) ranging between +0.5 and +1.0, but with considerable variability across subjects.

Figure 3 shows the mean values for each species. Shaded areas indicate the minimum amounts of masking.

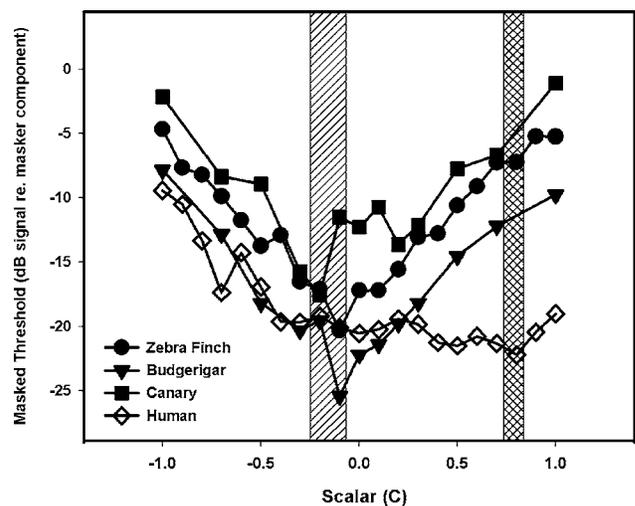


FIG. 3. Average thresholds for a 2.8 kHz tone embedded in harmonic complex maskers for zebra finches, budgerigars, canaries, and humans. The abscissa indicates the scalar value used to construct the masker waveform.

These average functions highlight the species differences in the shape of the masking functions with more sharply defined minimum masking regions for the birds and a relatively shallower and broader minimum for humans. All three species of birds show a similar pattern of masking across scalars and this pattern is different from that observed in humans. Budgerigars have overall levels of masking that match most closely those of human listeners at the negative scalar values but not at extreme positive scalar values. Zebra finches and canaries show patterns that are quite different from humans, and both show similar masked thresholds at negative and positive scalar values. A two-way repeated measures analysis of variance (ANOVA) indicated that there was a significant between-subjects effect of species [ $F(3,8)=13.67, p=0.002$ ] and a significant within-subjects effect of scalar [ $F(12,96)=36.20, p<0.0001$ ]. Furthermore, there was a significant interaction of species and scalar [ $F(36,96)=3.83, p<0.0001$ ]. A post-hoc Bonferroni  $t$  test showed that data from zebra finches and canaries were significantly different from humans ( $p=0.025$  and  $p=0.003$ , respectively), and that budgerigar data were different from canary thresholds ( $p=0.010$ ). In general, the release from masking across scalar values (i.e., the largest difference in threshold across thresholds) is greater in the bird species than in humans, with zebra finches showing the largest release from masking and humans the least.

As a check on the generalizability of the results obtained at a signal frequency of 2.8 kHz, we also tested one bird of each species at three other signal frequencies (i.e., 1.0, 2.0, and 4.0 kHz). Figure 4 shows that while the overall amount of masking at each signal frequency varied somewhat, all three birds still showed similar patterns of masking. The least masking occurred at 1.0 or 2.0 kHz for all species, and the most masking occurred at 4.0 kHz. These overall masking differences probably reflect the critical ratios at these frequencies, with increasing critical bandwidths at the higher frequencies for all these bird species (Okanoya and Dooling 1987).

The release from masking (maximum-minimum amount of masking) that occurs with changes in scalar value is summarized in Fig. 5 for each frequency. All three species showed the smallest release from masking due to a temporal waveform shape for a 1.0 kHz signal. The zebra finch and canary showed the largest release from masking at 4.0 kHz, while the budgerigar showed the largest release from masking at 2.8 kHz. This species difference in masking parallels species differences in other masking phenomena such as critical ratios which are larger for zebra finches and canaries than for budgerigars (Okanoya and Dooling 1987).

Interestingly, the minimum amount of masking occurs at a negative scalar for all birds and does not change considerably across frequencies in birds. These results are plotted in Fig. 6. The relatively small change in the scalar producing the least masking in birds is in contrast to results reported for humans. The data from humans taken from Lentz and Leek (2001) are plotted for comparison and show an inverse relationship between the scalar value resulting in minimum masking and signal frequency. Lentz and Leek suggested that this relationship reflected the curvature in the phase-by-frequency

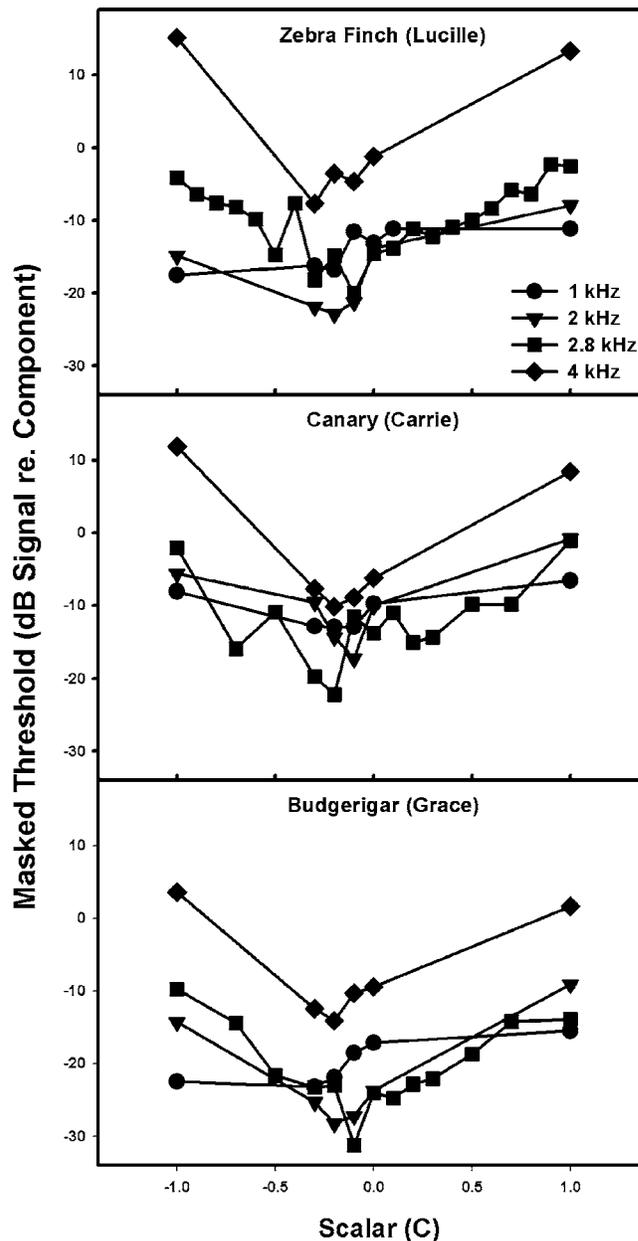


FIG. 4. Masked thresholds for 1.0, 2.0, and 4.0 kHz tones (dB signal re-masquer component) for a zebra finch, a canary, and a budgerigar. The abscissa indicates the scalar value used to construct the masker waveform.

map of the basilar membrane. The similarity in minimum masking scalar at different frequencies of the birds would argue that their phase curvatures are constant across frequency.

#### IV. DISCUSSION

Thresholds were measured in three different species of small birds, as well as in humans, for tones embedded in maskers that varied in waveform shapes from highly peaked to quite flat. Thresholds in all species varied as a function of temporal waveform shape, but different patterns of masking emerged between birds and humans. The results of this study provide further evidence that temporal waveform shape affects masking in birds differently than in humans.

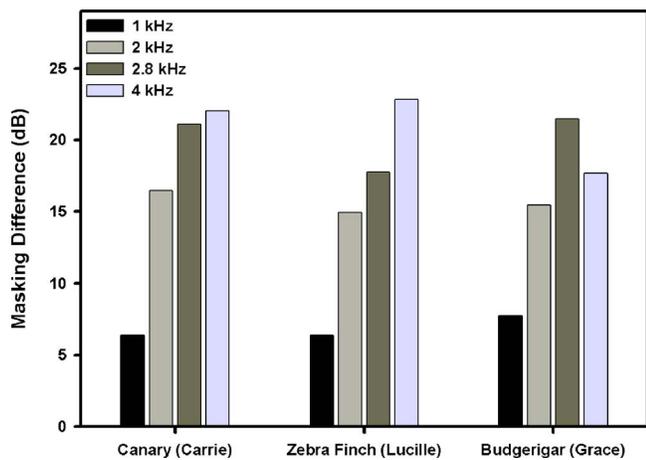


FIG. 5. Average differences in masking effectiveness of the harmonic complexes that produced the maximum and minimum amounts of masking of 1.0, 2.0, 2.8, and 4.0 kHz tones in a zebra finch, a canary, and a budgerigar.

Earlier comparative studies of Schroeder-phase masking in humans and birds have shown that in humans, differences between the original positive and negative Schroeder-phase maskers (i.e., with a scalar of  $\pm 1.0$ ) were on the order of 15 to 20 dB, but birds' thresholds were not more than 3–8 dB apart, depending on the fundamental frequency of the masker waveforms (Leek *et al.* 2000; Dooling *et al.* 2001). Further, the more effective masker in humans was always the negative Schroeder-phase masker, but the positive-phase masker usually produced slightly more masking in birds. These two characteristics, larger differences in Schroeder-phase masking in humans than in birds and the opposite sign of the more effective masker across species, were taken to reflect basic differences in cochlear structure and function between humans and birds. Leek *et al.* (2000), however, also showed that there were some large differences in masking effectiveness in birds for harmonic complexes constructed in cosine phase (here, a scalar of 0.0) and random phase, differences

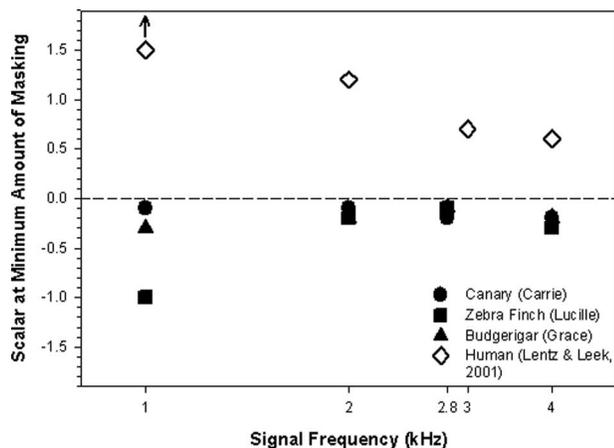


FIG. 6. Scalar values corresponding to the harmonic complexes that produced the minimum amount of masking for different frequencies in a zebra finch, a canary, and a budgerigar are compared to values reported in humans by Lentz and Leek (2001). The arrow above the data point for humans at 1.0 kHz indicates that the true masker producing the minimum amount of masking was probably greater than those reported.

on the order of 15–20 dB. Recall that random-phase waveforms are likely to have relatively flat envelopes, and cosine-phase waveforms are highly peaked.

The results from that earlier study (i.e., Leek *et al.* 2000) indicated that, under some circumstances, the waveform shape could have a large effect on masking in birds, notwithstanding the similar amounts of masking for the positive- and negative Schroeder-phase waves. This actually foreshadowed the finding here, that, given the appropriate scaled Schroeder waveform, reflecting perhaps a “matched” phase curvature in the cochlea of birds, the large Schroeder-phase differences in masking found in humans would be observed in birds. In other words, the lack of a large difference in birds between masking by positive and negative Schroeder-phase waveforms observed in earlier studies may have been because the phase spectra were incorrectly chosen for the avian cochlea. When an appropriate choice is made of monotonic phase change across frequency, and, as a result, an appropriate within-period frequency sweep rate, differences in masking effectiveness may be as large for all three bird species as in humans. This, in turn, suggests that the same mechanisms underlying these large masking differences due to waveform shape may be found in both mammalian and avian auditory processing.

In humans, the release from masking for the positive-phase Schroeder waveforms is thought to result from a cancellation between the phase of the stimulus and the phase of the auditory filter at the frequency place of the signal. This, in turn, is thought to occur because of characteristics of the traveling wave on the basilar membrane. The data reported here for birds also show masking differences in response to systematic phase changes in these harmonic complexes. Notwithstanding large differences in morphology and physiology between avian and mammalian cochleas, a similar explanation might be advanced in birds. Major differences in cochlear anatomy and mechanics include the much shorter length of the papilla, the distribution and types of hair cells, the configuration of hair cells into a matrix of support cells, but without the pillar cells found in mammalian cochleas, and the relatively larger and thicker tectorial membrane in avian ears than in mammalian ears, to name but a few [see Gleich and Manley (2000) for a comprehensive review of bird ear anatomy and physiology]. These differences are all likely to have an impact on the traveling wave and sound processing mechanisms that have been implicated in masking by harmonic complexes in humans.

Harmonic complexes with flatter temporal waveform shapes are generally more effective maskers than complexes with more peaked waveform shapes in birds and in humans. However, none of the species tested demonstrated the lowest thresholds with the peakiest masker waveform ( $C=0$ ), but at less peaked harmonic complexes, in the positive Schroeder-phase maskers for humans and for negative Schroeder-phase complexes in all bird species. In a pattern that is consistent with previous studies that involved the original Schroeder complexes [i.e.,  $C=\pm 1$ ; (Dooling *et al.* 2001; Leek *et al.* 2000)], similar amounts of masking were observed here for the positive and negative stimuli in birds. For nearly all scaled stimuli used here, the symmetry between waveforms

scaled by plus and minus one was notable in all birds. The pattern of results for human listeners was markedly different. As in previous reports, negative- and positive-phase complexes produced very different amounts of masking, and the lowest threshold occurred at a positive-phase scalar in humans (Lentz and Leek 2001; Oxenham and Dau 2001).

Kohlrausch and Sander (1995) and Lentz and Leek (2001) have argued that the patterns of masking by Schroeder-phase complexes may be used to estimate the phase characteristic of the auditory filters centered in regions of the signal frequencies tested. For these scaled versions of the Schroeder complexes, the scalar that produces the least amount of masking at a given signal frequency indirectly indicates the rate of change of phase as a function of the rate of change of frequency across the auditory filter. That is, the second derivative of the phase-by-frequency function in the masker waveform producing the least amount of masking among these stimuli is the phase curvature in the signal frequency region of the basilar membrane (but of opposite sign). As pointed out by Kohlrausch and Sander (1995), the Schroeder-phase waveform has a constant curvature given by the quantity  $\kappa$

$$\kappa = \frac{d^2\phi}{df^2} = C \frac{2\pi}{Nf_0^2}. \quad (2)$$

Applying this equation for  $C = -0.1$ , the scalar producing the minimum masking for budgerigars and finches, and  $-0.2$  for canaries, gives an estimate of the cochlear phase curvature of  $1.28 \times 10^{-6}$  and  $2.56 \times 10^{-6}$ , respectively. Note that the sign of the cochlear phase curvature is opposite to the sign of the phase curvature in the waveform. For humans, the scalar value that provides minimum masking for a signal of 3.0 kHz is  $+0.75$  with the same masker stimuli used in this study (Lentz and Leek 2001), so the curvature of the cochlear phase characteristic in that frequency region is estimated to be  $-9.62 \times 10^{-6}$ . These comparative data suggest that the phase change across frequency in humans is at least four times as rapid as in birds and is in the opposite direction.

The differences in the curvature of the phase characteristic are also reflected in the temporal features of the traveling wave in birds and humans. The species differences in masking patterns observed in this study might be explained by examining the interaction between those features and the phase spectra of the stimuli. Based on the cochlear frequency map provided by Gleich (2000) for the starling, the signal frequency of 2.8 kHz used here would fall at about 0.8 mm from the base of the cochlea. According to traveling wave velocities reported by (Gleich 2000), the time for a stimulus to arrive at the appropriate region of the cochlea would be about 0.4–0.5 ms. In humans, in contrast, the place on the cochlear map that processes 2.8 kHz is about 14 mm from the base (Greenwood 1961), and the time required for the traveling wave to arrive there is some 6.5 to 7.5 ms (Donaldson and Ruth 1993).

The minimum masking produced by a Schroeder-phase complex should occur in the region where the movement of the cochlear partition maximally compensates for the waveform shape of the external stimulus to produce a highly

modulated internal masker. This should occur for a scaled Schroeder-phase masker that matches the time of arrival of the signal component in the stimulus with the time of arrival of the traveling wave in the cochlea, but with a reversed sign to just compensate for the traveling wave. Thus, in birds we need a scaled Schroeder masker in which the 2.8 kHz component arrives at about 0.5 ms into the period of the waveform, and for humans, we require a masker in which the 2.8 kHz component arrives about 7 ms into the period. Recall the stimulus characteristics: frequencies ranging from 200–5000 Hz, with a fundamental period of 10 ms. Recall also that as the absolute values of the scalars approach 0.0, the rate of change of frequency increases. These facts, along with the recognition that the signal frequency of 2.8 kHz is clearly about midway into the frequency range of the masker, indicate that the minimum masker for birds will be a much smaller scalar than that found for humans. In fact, by observing the time course of the frequencies within each period of the masker for each scalar, we may calculate that a scalar of  $-0.1$  will produce a masker that reaches the 2.8 kHz component in about half a millisecond, and, perhaps significantly, it is that scalar that provides the least amount of masking in most of the birds in this study. A scalar of  $+0.5$  to  $+0.7$  results in maskers in which the 2.8 kHz component arrives from 6–7 ms into the period, and it is this region that provides the least amount of masking in humans. Thus, to a first approximation, the time of arrival of the signal frequency in the masker might be thought to just compensate for the time of arrival of the traveling wave in response to the masker-plus signal. This would suggest that a harmonic complex with frequencies arranged to occur in a glide lasting no more than the maximum latency from base to apex of a given cochlea would provide the best estimate of travel time in the cochlea. The complexes used here were 10 ms in glide time (within period). Perhaps something more on the order of 1–3 ms would better suit the shorter cochleas of birds, and in fact the glide times in the scaled stimuli with a scalar of 0.1 to 0.3 requires approximately that much time within each period to extend across the frequency range, regardless of the actual period of the stimulus. In other words, perhaps the bird papillae are simply too short to support different phase alterations across frequency. The papilla of zebra finches and canaries is 1.6 mm, and it is about 2.1 mm for budgerigars (Gleich *et al.* 1994; Manley *et al.* 1993). Thus, an upward-sweeping within-period glide occurring in about 1.03 ms in zebra finches and canaries and about 1.3 ms in budgerigars may be the temporal limits for introducing a phase change, compared to around 10 ms in humans.

Gleich and Manley (2000) proposed that the main stimulus to the hair cells in the avian basilar papilla is likely related to the resonance of the tectorial membrane (TM), and is indirectly influenced by the vibratory motion of the basilar membrane. The frequency-dependent motion of the basilar membrane would first activate those cells over it, the efferently innervated hair cells, producing active movement in the stereociliary bundles. This movement would feed mechanical energy into the TM in phase with the stimulus. Movement of the TM would then activate afferently innervated hair cells. According to this model, altered representations in response

to different input phases must occur subsequent to the motion of the basilar membrane. The near-zero scalars that produce minimum masking in the birds for all three frequencies tested might suggest that there is little phase change imposed by cochlear processing in the bird, which would be consistent with only small changes to basilar membrane motion, in contrast to the large and frequency-dependent phase lags observed in mammalian cochlear processing.

Taken together, the comparative results on masking by harmonic complexes by birds and humans along with previous findings of enhanced discrimination of temporal fine structure in harmonic complexes by birds (Dooling *et al.* 2002) invites speculation about a match between the temporal features of bird song and auditory specializations for perceiving song. The vocalizations of many birds are well characterized by temporal precision, rapid frequency sweeps, and in some cases, complex harmonic patterns (Greenewalt 1968). There is strong evidence for species-specific perceptual specializations enhancing the perception of species-specific calls (Dooling and Searcy 1979; Okanoya and Dooling 1991). More recent investigations show that birds are acutely sensitive to changes in the temporal fine structure of these natural vocalizations (Lohr *et al.* 2000) and evidence from studies of the neuromuscular activity of the syrinx show control of extremely fine temporal and spectral detail (see Suthers and Zollinger 2004 for review). Thus, the curious differences between humans and birds in the processing of these harmonic complexes may be a reflection of auditory processing especially suited for decoding the fine detail in bird vocalizations.

## V. CONCLUSIONS

These data provide a psychophysical description of the phase response of the avian basilar papilla. As in previous reports of masking by Schroeder-phase harmonic complexes in birds, we have shown that temporal waveform shape can affect masking in birds and humans in very different ways. Further, we have demonstrated that the least effective scalar Schroeder-phase masker in birds is different from that of humans at the same frequency channel. In birds, the least effective masker has a scalar value close to zero and has a negative phase. In humans, the least effective masker varies somewhat, but is always a positive phase. These results likely reflect fundamental anatomical and physiological differences in cochlear phase response between birds and humans, in particular differences in lengths of the basilar membrane and basilar papillae, and differences in delay as a function of frequency of the traveling wave. These differences may underlie some demonstrated cases of birds being able to hear fine detail in their vocalizations to which humans are insensitive.

## ACKNOWLEDGMENTS

This work was supported by NIH Grants DC-01372 to R.J.D., DC-005450 to A.M.L., DC-00626 to M.R.L., and DC-04664-01A2 (Core Center grant) to the University of Maryland. Special thanks to Dr. Micheal Dent for 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. The authors thank Otto Gleich for insightful discussions and comments about this research.

- Carlyon, R. P., and Datta, A. J. (1997a). "Excitation produced by Schroeder-phase complexes: Evidence for fast-acting compression in the auditory system," *J. Acoust. Soc. Am.* **101**, 3636–3647.
- Carlyon, R. P., and Datta, A. J. (1997b). "Masking period patterns of Schroeder-phase complexes: Effects of level, number of components, and phase of flanking components," *J. Acoust. Soc. Am.* **101**, 3648–3657.
- Donaldson, G. S., and Ruth, R. A. (1993). "Derived band auditory brainstem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects," *J. Acoust. Soc. Am.* **93**, 940–951.
- Dooling, R. J., Dent, M. L., Leek, M. R., and Gleich, O. (2001). "Masking by harmonic complexes in birds: Behavioral thresholds and cochlear responses," *Hear. Res.* **152**, 159–172.
- Dooling, R. J., Leek, M. R., Gleich, O., and Dent, M. L. (2002). "Auditory temporal resolution in birds: Discrimination of harmonic complexes," *J. Acoust. Soc. Am.* **112**, 748–759.
- Dooling, R. J., and Okanoya, K. (1995). "The Method of Constant Stimuli in testing auditory sensitivity in small birds," in *Methods in Comparative Psychoacoustics*, edited by G. M. Klump, R. J. Dooling, R. R. Fay, and W. C. Stebbins (Birkhauser-Verlag, Basel), pp. 161–169.
- Dooling, R. J., and Searcy, M. (1979). "Early perceptual selectivity in the swamp sparrow," *Dev. Psychobiol.* **13**, 499–506.
- Gescheider, G. A. (1985). *Psychophysics: Method, Theory, and Application* (Lawrence Erlbaum and Associates, New York).
- Gleich, O. (2000). Group delay and the velocity of excitation along the auditory sensory epithelium of starling and guinea pig," *Assoc. for Res. in Otolaryngol.* (ARO, St. Petersburg Beach, FL).
- 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.
- Gleich, O., and Narins, P. M. (1988). "The phase response of the primary auditory afferents in a songbird (*Sturnus vulgaris*)," *Hear. Res.* **32**, 81–92.
- Gleich, O., Manley, G. A., Mandl, A., and Dooling, R. J. (1994). "Basilar papilla of the canary and zebra finch: A quantitative scanning electron microscopical description," *J. Morphol.* **221**, 1–24.
- Greenewalt, C. H. (1968). *Bird Song: Acoustics and Physiology* (Smithsonian Inst. Press: Washington, DC).
- Greenwood, D. (1961). "Critical bandwidth and the frequency coordinates of the basilar membrane," *J. Acoust. Soc. Am.* **33**, 1344–1356.
- Kohlrausch, A., and Sander, A. (1995). "Phase effects in masking related to dispersion in the inner ear II. Masking period patterns of short targets," *J. Acoust. Soc. Am.* **97**, 1817–1829.
- 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.
- Lentz, J. J., and Leek, M. R. (2001). "Psychophysical estimates of cochlear phase response: Masking by harmonic complexes," *J. Assoc. Res. Otolaryngol.* **2**, 408–422.
- Lohr, B., Bartone, S., and Dooling, R. J. (2000). "The discrimination of fine-scale temporal changes in call-like harmonic stimuli by birds," *Association for Research in Otolaryngology* (ARO, St. Petersburg, FL).
- Manley, G. A., Scwabedisen, G., and Gleich, O. (1993). "Morphology of the basilar papilla of the budgerigar (*Melopsittacus undulatus*)," *J. Morphol.* **218**, 153–165.
- Okanoya, K. and Dooling, R. J. (1987). "Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds," *J. Comp. Psychol.* **101**, 7–15.
- Okanoya, K., and Dooling, R. J. (1991). "Perception of distance calls by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*): Assessing species-specific advantages," *J. Comp. Psychol.* **105**, 60–72.
- Oxenham, A. J., and Dau, T. (2001). "Towards a measure of auditory-filter phase response," *J. Acoust. Soc. Am.* **110**, 3169–3178.
- Recio, A., and Rhode, W. S. (2000). "Basilar membrane responses to broadband stimuli," *J. Acoust. Soc. Am.* **108**, 2281–2298.
- Schroeder, M. R. (1970). "Synthesis of low-peak-factor signals and binary

- sequences with low autocorrelation," *IEEE Trans. Inf. Theory* **16**, 85–89.
- Smith, B. K., Sieben, U. K., Kohlrausch, A., and Schroeder, M. R. (1986). "Phase effects in masking related to dispersion in the inner ear," *J. Acoust. Soc. Am.* **80**, 1631–1637.
- Summers, V., and Leek, M. R. (1998). "Masking of tones and speech by Schroeder-phase harmonic complexes in normally-hearing and hearing-impaired listeners," *Hear. Res.* **118**, 139–150.
- Suthers, R. A., and Zollinger, S. A. (2004). "Producing song: the vocal apparatus," *Ann. N.Y. Acad. Sci.* **1016**, 109–129.
- von Békésy, G. (1960). *Experiments in Hearing* (McGraw Hill, New York).