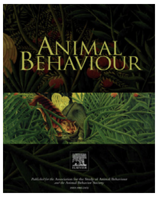
Animal Behaviour xxx (2016) 1–7

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav



Special Issue: Mechanisms & Function

Do we hear what birds hear in birdsong?

Robert J. Dooling*, Nora H. Prior

Department of Psychology, University of Maryland, College Park, MD, U.S.A.

ARTICLE INFO

Article history:
Received 4 July 2016
Initial acceptance 8 August 2016
Final acceptance 10 October 2016
Available online xxx
MS. number: SI-16-00586

Keywords:
audio–vocal communication
budgerigar
call
canary
song
temporal fine structure
zebra finch

Peter Marler's fascination with richness of birdsong included the notion that birds attended to some acoustic features of birdsong, likely in the time domain, which were inaccessible to human listeners. While a considerable amount is known about hearing and vocal communication in birds, how exactly birds perceive their auditory world still remains somewhat of a mystery. For sure, field and laboratory studies suggest that birds hear the spectral, gross temporal features (i.e. envelope) and perhaps syntax of birdsong much like we do. However, there is also ample anecdotal evidence that birds are consistently more sensitive than humans to at least some aspects of their song. Here we review several psychophysical studies supporting Marler's intuitions that birds have both an exquisite sensitivity to temporal fine structure and may be able to focus their auditory attention on critical acoustic details of their vocalizations. Zebra finches, *Taeniopygia guttata*, particularly, seem to be extremely sensitive to temporal fine structure in both synthetic stimuli and natural vocalizations. This finding, together with recent research highlighting the complexity of zebra finch vocalizations across contexts, raises interesting questions about what information zebra finches may be communicating in temporal fine structure. Together these findings show there is an acoustic richness in bird vocalizations that is available to birds but likely out of reach for human listeners. Depending on the universality of these findings, it raises questions about how we approach the study of birdsong and whether potentially significant information is routinely being encoded in the temporal fine structure of avian vocal signals.

Birdsong has served as an extremely productive behavioural and neurobiological model of vertebrate learning in general and as a model of human speech development and acoustic communication specifically. But compared to humans, it is fair to say we know considerably more about vocal production in birds and much less about perception of species-specific vocal signals. Furthermore, while there are many parallels in the learning and production of vocalizations between these two communication systems, there are surprisingly few demonstrations that these parallels extend to the perceptual systems of humans and birds. This raises a simple question: does birdsong sound to birds like it does to humans? What we do know about basic hearing in birds comes mostly from psychophysical studies using simple sounds such as tones and noises. From such studies, we know that birds hear best between about 1 and 5 kHz and show discrimination thresholds for changes in frequency, intensity and temporal envelope generally approaching the values typically reported for humans (reviewed in Fay, 1988; Dooling, Lohr, & Dent, 2000), although species differences in salience sometimes emerge when birds are tested with species-specific vocalizations (Dooling, Brown, Klump, & Okanoya, 1992).

Students of ornithology often describe birdsong in terms of its pitch, tempo, complexity, structural organization and stereotypy. Indeed acoustic correlates of these common perceptual dimensions are how we make judgements about whether a song has been learned or altered in some significant way. While we can describe speech in these terms, we usually do not. Instead, when listening to speech, we typically focus on well-learned acoustic patterns, reflexively attending to critical acoustic features necessary for communication. This combination of well-learned acoustic patterns and sharp attentional focus is part of what leads to the notion that speech perception is special for humans. It is possible that these advantages are available to birds. Anecdotal field observations over the years, coupled with well-known differences between birds and mammals in the anatomy and physiology of the peripheral and central auditory systems, has led to speculations that birds must have extremely fine temporal processing abilities (Carr & Friedman, 1999; Greenewalt, 1968; Konishi, 1969; Pumphrey, 1961; Schwartzkopff, 1968).

---

* Correspondence: R. J. Dooling, Biology/Psychology Building, University of Maryland, College Park, MD 20742, U.S.A.
E-mail address: rdooling@umd.edu (R. J. Dooling).

http://dx.doi.org/10.1016/j.anbehav.2016.10.012
0003-3472/© 2016 Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour.

Please cite this article in press as: Dooling, R. J., & Prior, N. H., Do we hear what birds hear in birdsong?, Animal Behaviour (2016), http://dx.doi.org/10.1016/j.anbehav.2016.10.012

Zebra finches, *Taeniopygia guttata*, have become an extremely popular model for studying song learning, bioacoustics and vocal behaviour and are a good species for investigating these phenomena (Bolhuis & Everaert, 2013; Braaten, Petzoldt, & Colbath, 2006; Brainard & Doupe, 2001; Elie & Theunissen, 2016). These birds are closed-ended learners that have a single sensitive period for song learning, after which new songs cannot be learned. The result of this sensitive period is a single, highly stereotyped song consisting of an ordering of syllables, termed a motif, that is repeated several times throughout the song bout. Motifs are typically composed of five to eight notes or syllables. Each syllable is an acoustically distinct harmonic complex, which contains multiple cues that result in a unique sound (Zann, 1996). The simple and repetitive nature of these songs has allowed for extensive study of the behavioural and neurobiological basis of song development, song learning and song production (see, for example, Brainard & Doupe, 2001; Doupe & Konishi, 1991; Glaze & Troyer, 2006; Margoliash & Fortune, 1992; Troyer, 2016).

Zebra finch contact calls are some of the most obvious and ubiquitous vocalizations given by these birds in captivity (Blaich, Kovacevik, Tansinsin, Van Hoy, & Syud, 1995) and in the wild (Zann, 1996). Male zebra finches learn their songs and perhaps some aspects of calls (Zann, 1984; Simpson & Vicario, 1990). Peter Marler himself identified bird calls as an underutilized model for the neurobiology of acoustic communication (Marler, 2004). Indeed, zebra finch calling behaviour is proving more complex than was previously thought. Importantly, not only do male and female zebra finches produce a wide range of acoustically distinguishable calls (Elie & Theunissen, 2016), but these calls occur in distinct social circumstances (Elie et al., 2010; Elie & Theunissen, 2016; Gill, Goymann, Ter Maat, & Gahr, 2015; Williams, 2001). Furthermore, variation in calling behaviour between birds can have surprising effects on mating and parental behaviour (Boucaud, Mariette, Villain, & Vignal, 2016).

The case to be made here is that a combination of species-specific attentional advantages and acute temporal resolving power in birds could very well put communicatively relevant acoustic details of complex song-like stimuli out of the reach of human hearing. Auditory psychophysical studies in birds that have manipulated attention are rare, but results from one psychophysical study are tantalizing. In this study, brief tonal sequences were concatenated to create a synthetic model of the species-specific contact call of budgerigars, *Melopsittacus undulatus*. Birds were then tested on frequency changes in these sequences, the location of which was both fixed and random from trial to trial. Results showed that budgerigars attended to the tonal complex as a whole while human listeners did not. This suggests top–down attentional processes are available at least to budgerigars, but not to humans, when listening to these call-like tonal complexes (Dent, Dooling, & Pierce, 2000).

Zebra finch vocalizations are extremely complex and provide a multitude of acoustic cues for discrimination, including amplitude envelope modulations, spectral structure and temporal fine structure. In considering these features, it is important to distinguish between two types of temporal cues. Most discussions of the temporal features of birdsong focus on the envelope changes. Temporal envelope is a global timing that occurs over many milliseconds to seconds, and accounts for the global rhythm and timing of song including, motif, syllable, and note durations and intersyllable intervals. Fine structure, on the other hand, is a local timing that occurs over milliseconds, and includes amplitude, spectral and temporal cues within individual harmonic syllables. While both temporal envelope and fine structure cues are present in vocalizations, especially in harmonic zebra finch vocalizations, fine structure has historically been ignored in part because it is not apparent in traditional sonographic analysis that has been the mainstay in birdsong research for decades.

Whether birds can actually hear and discriminate temporal fine structure in complex sounds is another matter. Here we discuss research that addressed this issue using synthetic stimuli and three species of birds with very different vocalizations — two songbird species, one of which was an open-ended learner (canary, *Serinus canaria*) and one of which was a close-ended learner (zebra finch), and a nonsongbird species (budgerigar). The perceptual thresholds for these species were directly compared to those of humans. All three perceptual experiments below used identical standard psychophysical methods: birds were trained by operant conditioning to discriminate a change (a target) against an ongoing, repeating sound (the background). Correct responses were rewarded with food and false alarms were punished with a brief blackout. The fact that birds were all trained and tested using the same behavioural procedures and that psychophysical threshold estimates were obtained using the same method in birds and humans strengthens the comparisons.

DISCRIMINATION OF TEMPORAL FINE STRUCTURE: SCHROEDER HARMONIC COMPLEXES

Some bird vocalizations, like those of the zebra finch, can be described as predominantly complex harmonics (Zann, 1984), making them more difficult to describe and characterize than more tonal or whistled bird vocalizations. Modern signal-processing techniques can be used to manipulate complex harmonic sounds and to create synthetic harmonic models for testing, which can closely mimic some of the natural properties of these harmonic bird sounds. This ability allows for perceptual threshold measurements of the fine details in complex harmonic sounds that escape notice in more traditional sonographic analysis.

Evidence for an extreme sensitivity to temporal fine structure in birds comes from a study looking at the discrimination of Schroeder waveforms (Dooling, Leek, Gleich, & Dent, 2002; Schroeder, 1970). These stimuli were constructed of harmonically related pure tones with the phases of the individual tonal components adjusted so that they 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. Figure 1 shows examples of negative and positive Schroeder-phase waveforms with three different fundamental frequencies. The acoustic differences between members of a pair of these complexes were limited to temporal fine structure: all waveforms had a flat envelope and, within a pair, defined by the fundamental frequency, had identical long-term spectra. These waveforms were 260 ms in duration, including 20 ms rise/fall times. While the envelope and overall spectrum were constant across stimuli to be discriminated, the temporal fine structure was reversed. Test stimuli consisted of seven pairs of these harmonic complexes with fundamental periods ranging from 6.6 ms (fundamental frequency of 150 Hz) to 1 ms (fundamental frequency of 1000 Hz) in duration. Finches, budgerigars, canaries and humans were tested on their ability to discriminate a forward Schroeder complex from a reversed Schroeder complex. Birds were tested on a Go-NOGO task, while humans were tested in a two-alternative, forced-choice task, with values ranging from chance performance at 50% correct to perfect performance (100%), so the bird data were scaled to the range of the human data in Fig. 1 (Dooling et al., 2002).

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

Please cite this article in press as: Dooling, R. J., & Prior, N. H., Do we hear what birds hear in birdsong?, Animal Behaviour (2016), http://dx.doi.org/10.1016/j.anbehav.2016.10.012