Condor 85:259–261
© The Cooper Ornithological Society 1983

COMMENTARY

ON SONOGRAMS, HARMONICS, AND ASSUMPTIONS

The ability of sonograms to present acoustic data in a format that is easily comprehended, compared, and printed has made them an important tool in studies of avian behavior and systematics. The 10 volumes of *The Condor* from 1972–1981 contained 51 papers, involving 75% of all issues, that used sonographic data and two papers discussing the presentation and use of such data. This comment follows in the latter series by indicating some possible pitfalls in the interpretation of sonograms.

Birds can utter a wide variety of phonations, but these can be divided into two categories depending on how the sound is produced. The most common kind of avian phonation is a simple tone, or whistle, in which all the energy at any given instant is concentrated in a single frequency. Sonograms of unmodulated, simple tones show a single horizontal band. The production of a whistled sound does not depend on the oscillation of a membrane or any other mechanical element of the syrinx (Gaunt et al. 1982). A second kind of sound is that in which, at any moment, energy is distributed into more than one frequency. The generation of many of these sounds is supposed to involve either mechanical oscillators or coupled resonators. I am concerned here with multi-frequency sounds in which the energy is distributed into distinct, discreet frequencies, i.e., not broad-band noises. Davis (1964) included such sounds among those he called "complex tones." However, I find it useful to restrict "complex tones" to those sounds that are produced by a single generator. A sonogram of an unmodulated, complex tone contains a series of parallel bands. Such sounds are popularly termed "harmonic tones," and the distribution of frequencies is called a "harmonic spectrum." In a true harmonic tone, such as is produced by most stringed and wind instruments, the higher frequencies (overtones) are integer multiples of the fundamental frequency (or first harmonic). This simple relationship is familiar to many researchers, but some investigators assume that *all* multi-frequency sounds are harmonic tones. Therein lies a set of problems because (1) not all multi-banded figures on sonograms represent complex tones, (2) not all complex tones are composed of harmonics as here defined, and (3) the mechanism(s) whereby a syrinx generates complex tones is (are) not clearly understood.

First, an apparent harmonic may be an artifact. Sound spectrographs will produce spurious harmonics if the signal is introduced at too high a gain (Davis 1964, Greenewalt 1968:9). This type of error is usually overcome with experience, but that experience should not be assumed in reviewers, editors, or readers.

Multi-frequency sounds may also be the product of the two-voice phenomenon. If both sides of the syrinx should produce complex tones with slightly different fundamentals, very puzzling patterns may result (S. Nowicki, pers. comm.). The two-voice phenomenon is surely responsible for instances in which frequency bands cross over each other, diverge with falling frequencies, or converge with rising frequencies. (Overtones, whether partial or harmonic, must diverge with rising frequencies and converge with falling ones.) In a confusion of the concepts of two-voices and harmonics, Anderson (1978) described "peculiar . . . harmonic bands" exhibiting several of the criteria for two voices, that may be an example of "dual control of sound production."

Second, if a tone is sufficiently modulated in either frequency or amplitude, then its sonogram will show side bands to both sides of and parallel to the tone's frequency (carrier frequency). The side bands will be separated from the carrier frequency and, if there are several, from each other by the frequency of repetition rate, or modulating frequency. Hence, if the carrier frequency can be evenly divided by the modulating frequency, the sonographic pattern will resemble a harmonic spectrum of a tone with a fundamental frequency equivalent to the modulating frequency. The carrier frequency usually appears as an emphasized harmonic. Because the modulating frequency is usually much lower than the carrier frequency, the apparent "fundamental frequency" is low, the "harmonic spectrum" is composed of closely spaced bands, and, if the carrier is high, the "lower harmonics" are usually absent. The exact sonographic pattern depends both on the nature of the modulation and the setting of the spectrograph; the problem is acute with narrow band-pass filter settings. Appropriate sounds for producing such patterns appear to be common in avian phonations (Davis 1964, Watkins 1967, Marler 1969).

Third, a freely oscillating, edge-clamped membrane, such as a syringeal membrane, behaves quite differently from a freely oscillating, end-clamped string. A wave propagates in one dimension along the length of the string. If, as is probable, the wave encounters an impedance mismatch at the end of the string, some of it will reflect back in the opposite phase, so that it seems to have changed to the opposite side of the string. If the vibration is repeated, succeeding waves will interact with the reflected waves to form a series of standing waves, the number of which is equivalent to the number of harmonics. In a membrane, however, the wave radiates as a series of arcs from the point of stimulation. Hence, it propagates and is reflected in two dimensions, deforming a surface rather than a line. The resulting interaction is far more complex than in the one-dimensional string, and the resulting overtones are not normally harmonic. Rather, they will occur at varying fractional (partial) multiples of the fundamental (Rossing 1982). Such partial overtones account for the characteristic, pitchless sounds of many drums. Tunable drums with distinct pitch, such as kettledrums, represent special classes of membrane instruments in which various factors force the partial frequencies closer to harmonic values of an apparently missing fundamental. Casey (1981) showed that partials will be generated by membranes of a variety of shapes subjected to various degrees of damping. Hence, it is unlikely that a freely oscillating syringeal membrane of any shape, even if constrained like the head of a kettledrum, will produce a classic harmonic spectrum. Yet some birds, e.g., Phainopepla (*Phainopepla nitens*; Leger and Carroll 1981) and Whimbrel (*Numenius phaeopus*; Skeel 1978) do seem to utter sounds with true harmonic spectra. This suggests that something is faulty, or at least lacking, in our present understanding of syringeal mechanics.

Problems in the interpretation of complex tones may be exacerbated by improper use of a sonogram. For instance, the analysis of an harmonic spectrum depends on a rather precise determination of frequencies, but many sound spectrographs, including the commonly used Kay Elemetrics "Sona-Graph," do not measure frequency well, especially when used with a wide-band filter. The trigger frequency is generally assumed to lie at the center of the inscribed band, but that may not be so (Davis 1964), and even if it does, the center must be estimated with some degree of error. Further, the width of the band varies with the gain setting of the spectrograph. Even when using a narrow-band pass filter, e.g., 40 Hz, for high resolution, one should expect a production error of up to ±20 Hz. To this must be added the investigator's interpolation error. Calibration marks are, at best, at 500 Hz, more usually 1,000 Hz intervals, and intervening values must be interpolated. On a sonogram with a frequency range of 80–8,000 Hz, which is common in avian studies, each millimeter on the ordinate represents 70 Hz. Hence, a mistake of 10 Hz will result from an interpolation error of only 0.14 mm!

For most behavioral or systematic investigations, failure to recognize these problems is of little import, but for those of us interested in syringeal function, such errors can be serious. The reason for this is that we have yet to devise a technique to observe a functioning syrinx in vivo. True, syringeal structures are well-known, but most of our notions of how those structures work are based on extrapolations from the sounds they produce. Hence, misanalyzed sonograms constitute noise in our data base. Let me provide an example.

I have chosen this example for several reasons. First, the faults in acoustic analysis are irrelevant to the author's main interest. Thus, the paper is typical of many papers that use acoustic data but are not concerned primarily with acoustic analysis. Second, the author recognizes the limitations of sonograms and has taken more pains than most to insure the accuracy of her frequency estimates. Third, and this is most unusual, she has provided sufficient details of her methods and sufficient numerical data that I was able to confirm a problem rather than simply suspect it. Finally, the author informs me that she was specifically trying to avoid any acoustic implications, in which case her failure to do so provides a splendid example of how insidiously a hidden assumption can force us to seek a preconceived pattern where it does not exist.

In a paper on the vocalizations of gulls, Hand (1981: 290) stated:

> The mean harmonic band intervals were *estimated* [author's italics] by placing a grid calibrated in 50-Hz intervals over the sound figures at the highest point. Frequencies of visible harmonics at this point were estimated to the nearest 10 Hz and compared to values on a numbers table, these values being integral multiples of possible intervals. For example, a choking sound had visible bands at the following Hz: 180, 320, 450, and 650. The value from the numbers table producing the best fit to the observed bands corresponds to a harmonic interval of 160 Hz. A calculated mean interband interval for this call would be 156.6 Hz, which closely fits the 160 Hz estimate determined by my method.

I am not sure what "closely" means in the context of a supposedly precise mathematical relationship, but will assume it means "within an acceptable margin of error." Hand mentioned an error of ±7 Hz. This value was calculated as the variation between the estimated value of the supposed fundamental frequency and its frequency calculated from the mean interband interval for a sample of calls (Hand, pers. comm.). With this error, 156 Hz is a good approximation of 160 Hz, and we cannot quarrel with 320 Hz for the second harmonic of 160 Hz, but none of the other values are close. Still, even with an error of ±20 Hz, the values of 180 and 659 Hz for the fundamental and third harmonic are only marginal, and 450 Hz is significantly different from the expected 480 Hz of the third harmonic. None of the higher frequencies is within 30 Hz of the expected harmonic values of Hand's lowest estimated frequency of 180 Hz, if that were considered the fundamental. Consider also that the mean interband interval is an average of 130, 140, and 200(!) Hz. Whatever this series may represent, it is not a simple, harmonic series.

Hand's treatment suffers from three flaws. The first, use of the inappropriate term "harmonic," is trivial because it fits a tradition that has not previously been challenged. The second is choice of an analytic procedure that infers a certain kind of result, for the use of a numbers table assumes that the observed frequencies *should* be distributed in a pattern of integer multiples. Ironically, some of Hand's data probably did conform to that pattern. Her Figure 8 suggests that the Long Call of some gulls is composed of harmonics. The third, and in my view most serious flaw, is failure to state explicitly that at least some of her data are not easily fitted to the expected pattern. On such discrepancies scientific theories can hang or fall, and failure to recognize or proclaim them only preserves the present paradigm.

Even when performed with great care, acoustic analyses may not lead to unambiguous solutions. In developing his ideas of how a syrinx might produce harmonic sounds, Greenewalt (1968, Chap. 10) considered the fact that many birds produce sounds in which the apparent fundamental and several lower harmonics are missing. He hypothesized that these were suppressed by proximity of the vibrating membrane to the opposite tracheo-bronchial wall. As one example, he presented a detailed analysis of the Scold call of a Black-capped Chickadee (*Parus atricapillus*). A sonogram, a "harmonic spectrum," a chart of instantaneous frequencies (taken from a period counter), and a table of measured "harmonic" frequencies were included as data. The sonogram shows five parallel bands between 2,900 and 4,000 Hz. Greenewalt interpreted these as the 7th–11th harmonics of a fundamental of 415 Hz with the darker, ninth harmonic (3,739 Hz) "dominant." Frequencies of the harmonics were determined by isolating each by filtration and determining the time interval for 100 successive periods with a period counter. This technique provides far more precision than any estimate from a sonogram.

On the basis of the sonogram alone, we might entertain three hypotheses for the nature of this sound: (1) a fundamental frequency of about 3,320 Hz (Greenewalt's 8th harmonic) with four overtones and a weak undertone (as sometimes occurs for kettledrums), (2) a strongly modulated tone, or (3) Greenewalt's explanation of the higher harmonics of a suppressed fundamental. His other data allow us to extend the analysis. The table of frequencies clearly shows that the bands occur at regular intervals. That fact alone dismisses the first hypothesis, because partial overtones do not occur at regular intervals. Moreover, the predicted partials of either a round or oval membrane, with fundamentals of 3,320 Hz would be higher than the observed frequencies (Casey 1981). Both the oscillogram and instantaneous frequency chart show strong modulations of both amplitude and frequency. The mean modulating frequency is given as 417 Hz; the interval between bands can be calculated to range from 414 to 420 Hz. The amplitude modulation is sufficient that I would consider the sound a series of pulses. Greenewalt rejected the interpretation of a pulsed tone on the basis that a coupled tracheal resonance was not present. However, a pulsed tone need not couple to a resonator to produce a harmonic spectrum (Watkins 1967).

Both of these analyses assume that the sound is generated by a single source. However, S. Nowicki (pers. comm.) has data suggesting that a chickadee's "dee" contains an effect of the two-voice phenomenon. Each side appears to produce a different complex tone. How the interaction of these produces the sonogram and wave form analyzed by Greenewalt requires further study. If Nowicki is correct, then an already perplexing situation has become more so. Further, the alternatives to Greenewalt's hypothesis still beg the question of how harmonics are formed in those avian phonations in which they are certainly present. Until recently, most investigators of syringeal function evidently *assumed* that harmonic oscillation was a natural mode for membranes. As that is not so, then many avian phonations constitute a puzzle—the key to which may lie in anyone's data.

In summary, a sonogram can be an extremely useful tool for investigations of behavior and systematics. It is not a good tool for detailed acoustic analysis. Investigators in behavior and systematics should be aware of the weaknesses of the technique, even though those weaknesses will seldom directly affect their own work. On a more general level, we should all remember that apparently aberrant data may represent a reality rather than an expression of some error in the system. That which is puzzling to us may be exactly the information someone else seeks. Hence, drawing attention to such situations may be a service. Finally, we should constantly remind ourselves that all of our theories are underlain by sets of assumptions, some not at all obvious. A periodic review of the assumptions in one's field, and of their likely effects on one's procedures, will amost always be profitable.

I am extremely grateful to Judith Latta Hand for comments on an earlier draft, and for her willingness to let her work be used in this way. Stephen Nowicki very graciously gave permission to use some of his unpublished ideas. Lincoln Fairchild provided valuable discussions and commented on the draft, which was also critiqued by Richard M. Casey and Sandra L. L. Gaunt. Reviewers Eugene Morton and William Barklow also provided valuable comments. This work was supported by grant number DEB-7911774 from the National Science Foundation.

LITERATURE CITED

Anderson, W. L. 1978. Vocalizations of Scaled Quail. Condor 80:49–63.

Casey, R. M. 1981. Theoretical analysis of tympanic membranes in avian syrinx. M.Sc. thesis, Ohio State Univ., Columbus.

Davis, L. I. 1964. Biological acoustics and the use of the sound spectrograph. Southwest. Nat. 9:118–145.

Gaunt, A. S., S. L. L. Gaunt, and R. M. Casey. 1982. Syringeal mechanics reassessed: evidence from *Streptopelia*. Auk 99:474–494.

Greenewalt, C. H. 1968. Bird song: acoustics and physiology. Smithsonian Inst. Press, Washington, DC.

Hand, J. L. 1981. A comparison of vocalizations of Western Gulls (*Larus occidentalis occidentalis* and *L. o. livens*). Condor 83:289–301.

Leger, D. W., and L. F. Carroll. 1981. Mobbing calls of Phainopepla. Condor 83:377–380.

Marler, P. 1969. Tonal quality of bird sounds, p. 5–18. *In* R. A. Hinde [ed.], Bird vocalizations. Cambridge Univ. Press, New York.

Rossing, T. D. 1982. The science of sound. Addison-Wesley, Reading, MA.

Skeel, M. A. 1978. Vocalizations of the Whimbrel on its breeding grounds. Condor 80:194–202.

Watkins, W. A. 1967. The harmonic interval: fact or artifact in spectral analysis of pulse trains, p. 15–41. *In* W. N. Tavolga [ed.], Marine bioacoustics. Vol. 2. Pergamon Press, New York.

ABBOT S. GAUNT, *Department of Zoology, The Ohio State University, 1735 Neil Avenue, Columbus, Ohio 43210.*