Z. Tierpsychol., 48, 306—330 (1978)

ISSN 0044-3573 / ASTM-Coden: ZETIAG

Duke University and Medical Center Durham, North Carolina

A Simple Method for the Rapid Analysis of Animal Sounds

By J. E. R. STADDON, LEE W. MCGEORGE, R. A. BRUCE, and F. F. KLEIN

With 18 figures

Received: November 7, 1977
Accepted: April 14, 1978

Abstract

A simple, real-time method for displaying the information contained in the zero-crossings of acoustic signals is described. The method can be used even with many signals that have harmonics, and reveals a wealth of fine structure in bird song. Some of this structure may serve a communicatory function.

The visual display provided by the sound spectrograph has greatly aided the study of vocal communication. Nevertheless, the instrument is not without its drawbacks. Among these are its slowness of operation, the small signal sample (typically 2 s or so) that can be processed at one time, the many adjustments that must be made and the consequent difficulty of obtaining standardized outputs, its limited frequency range, and its inability to represent accurately both frequency and modulation rate (CHERRY 1961; DAVIS 1964; GREENEWALT 1968; MARLER 1969; WATKINS 1967). Some of these problems, such as speed of operation and standardized output, are likely to be alleviated by means of the real-time Fourier analysis techniques now coming into wider use (e. g., HOPKINS, ROSSETTO, and LUTJEN 1974). The trade-off between frequency and modulation-rate information is inseparable from any Fourier-analysis technique, however, and cannot be bypassed.

GREENEWALT (1968) has pioneered a promising method that overcomes many of these difficulties. In a tonally pure signal, frequency modulation can be accurately assessed by looking just at the zero crossings of the signal. Variation in a plot of the times between successive positive-going (or negative-going) zero crossings then provides an accurate measure of moment-by-moment variations in the "instantaneous frequency" of the signal. Moreover, modulation-rate information is accurately represented without loss of frequency information. GREENEWALT made use of these characteristics of zero-crossings analysis to demonstrate the existence of rapid frequency modulation (up to 300 Hz or more) in brief fragments of "whistled" (i. e., tonally pure) bird song.

The zero crossings of a complex signal are not as easily interpreted as those from a pure or frequency-modulated tone. For example, the pattern of

U.S. Copyright Clearance Center Code Statement: 0044-3573/78/4803-0306$02.50/0

zero crossings produced by square, triangular, and sine wave-forms of the same fundamental frequency is the same, yet the harmonic structure of the waves is quite different, and the difference would be immediately revealed in a sound spectrogram. This limitation on zero-crossings analysis is not as severe as it might seem. First, almost any signal analysis is subject to ambiguities. The sound spectrograph, for example, is insensitive to phase differences and under some conditions confuses waveforms that are strikingly different both to the ear and to the eye (e. g., WATKINS 1967). Many of the waveforms discussed by WATKINS would be readily discriminated by a zero-crossings analysis, however. Second, little of the information for speech perception, at least, is lost by infinite peak clipping, which preserves only zero-crossings information (LICKLIDER and MILLER 1951). This suggests that much information of communicative significance is contained in zero crossings and merely awaits the proper analysis techniques.

Since no technique is perfect, the question is whether its imperfections are critical to what is being measured. The insensitivity of the sound spectrograph to small phase shifts does not matter for the analysis of speech, where such shifts are not critical. However, the instrument would be a poor choice for the study of sound localization, or for the analysis of bat sonar signals where small Doppler phase shifts may be important. This issue is still open as far as zero-crossings analysis is concerned, for two reasons. First, practical difficulties have meant that the technique has not been widely used. Consequently, there has been little opportunity for its limitations to become apparent. Second, the information-carrying acoustic features in many bioacoustic signals, most notably bird song, have not yet been fully identified. Hence it is not known whether these features are well or poorly represented by a zero-crossings analysis. Recent work, discussed below, suggests that fine structural features of bird song may be critical in communication. We shall demonstrate that such features are uniquely accessible to zero-crossings analysis.

A number of studies of automatic speech recognition have employed zero-crossings analysis (e. g. FLANAGAN 1972; SCARR 1968). The usual approach has been to attempt to predict known phonetic features from the sequence of zero crossings. The analysis of bird song presents a different problem, since neither the information in the signal, nor the information-carrying acoustic features, are known with precision. Hence our objective has been to display zero-crossings information in as unbiased a way as possible. The eye can be relied upon to recognize patterns, whose communicative significance can then be studied experimentally.

Measurement Techniques: GREENEWALT (1968) measured the time between successive zero crossings with the aid of digital counters and timers and then hand plotted each point in his graphs. His method is therefore slow and not practical for scanning large quantities of song. Subsequently, more rapid techniques have been devised. For example, KINDLMANN, BERMAN, JOHNSON, POLLAK, HENSON, and NOVICK (1973) describe a method using a zero-crossing detector in conjunction with a General Radio counter and a digital-to-analogue converter. They used this apparatus to display frequency changes in bat sonar pulses on an oscilloscope. The method has apparently not been applied to bird song. HJORTH (1970) describes a method devised by TOVE, NORMAN, ISAKSSON, and CZEKAJEWSKI (1966) which uses a pen writer to display zero-crossings information from physiological variables such as heart rate. It appears to work quite well for bird song, although octave-band prefiltering is required.

With the exception of GREENEWALT's method, which is unacceptably slow, all the other techniques for displaying zero-crossings information use the time between N successive crossings (where N can be one or more)[1]), or the number of crossings in a fixed time "window",

[1]) It is in fact undesirable to average over more than one period, although this does not appear to be widely recognized. The reason is that if any superimposed frequency modulation approximates N periods of the carrier, spurious interactions (analogous to difference tones) will take place, yielding periodicities in the sequence of averages that are not present in the originating signal. Conversely, if the averaging period is much greater than the modulation period, information about the modulation rate is lost.

to control the Y-axis displacement of an oscilloscope beam (or of a kymograph pen). The record is visible both between crossings and in the absence of a signal. This feature leads to problems in representing signals with multiple repeating waveforms. For example, consider a waveform consisting of tone bursts separated by fixed silent intervals. Clearly there will be just two interzero-crossing times associated with such a signal: the time between crossings within a burst, and the time between bursts. However, any system in which the record is visible between zero crossings will be constrained to represent the signal as either a series of sawtooths (at fast sweep speeds), or a broad band (at low sweep speeds). Since a fairly low sweep speed will usually be desired (so that a signal sample of reasonable length can be encompassed), the user will have to settle for the broad band display, even though such a display is also produced by band-limited noise and fails to show significant characteristics of the wave.

This limitation is serious but unnecessary. Zero crossings are discrete, point events. Hence, in an unbiased display, they should be represented as discrete elements, that is, as points. If the Y-displacement is proportional to inter-crossing time, as before, but each crossing is represented by a single dot, rather than a displacement of a line, the bimodal signal just discussed finds a natural representation: as two rows of dots, at fast sweep speeds, or as two lines, at slow speeds.

This is the principle of the zero-crossings analyzer (ZCA) to be described in this paper. We first describe the mode of operation of the device, then give examples of the displays produced by a sample of animal sounds — chiefly bird song. The displays are contrasted with those produced by means of the sound spectrograph. We give examples to show the utility of the ZCA in analyzing the fine temporal structure of bird song, and in detecting interactions between two voices. The paper concludes with a brief discussion of the research possibilities raised by zero-crossings analysis.

The Zero-Crossings Analyzer (ZCA)

Fig. 1 shows the principle of operation of the zero-crossings analyzer. The Fig. shows a picture taken from the face of a Tektronix type 5103 dual-beam storage oscilloscope. The top trace is the waveform of a 1000 Hz tone. The lower trace shows the operation of the ZCA. At each negative-going zero crossing[2]) the oscilloscope beam is briefly intensified and then reset upwards to a baseline. As soon as the beam is reset it begins to move downwards again at a rate determined by the time-constant of a R-C circuit (see Appendix A). At the next negative-going crossing, the process is repeated. During normal operation the beam is blanked so that only the brief intensifications at each zero crossing are visible: they appear as a row of dots for a pure tone of fixed frequency.

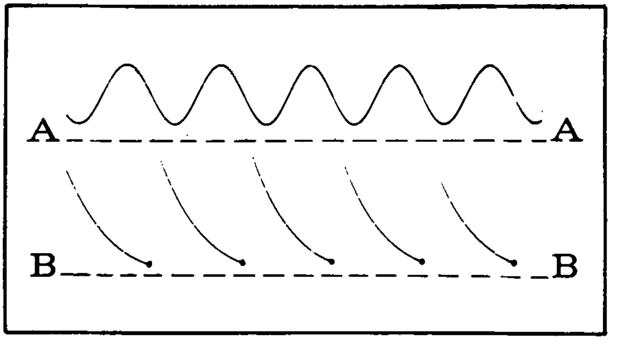

Fig. 1: A. Waveform of a 1000 Hz tone. B. Operation of the ZCA

[2]) In order to facilitate later comparisons with sound spectrograms, the picture is inverted so that low frequencies (i. e., long intercrossing times) appear at the bottom, rather than at the top. Hence it appears as if the ZCA detects negative-going crossings, whereas this and later pictures were made with an instrument that actually responds to only positive-going crossings.

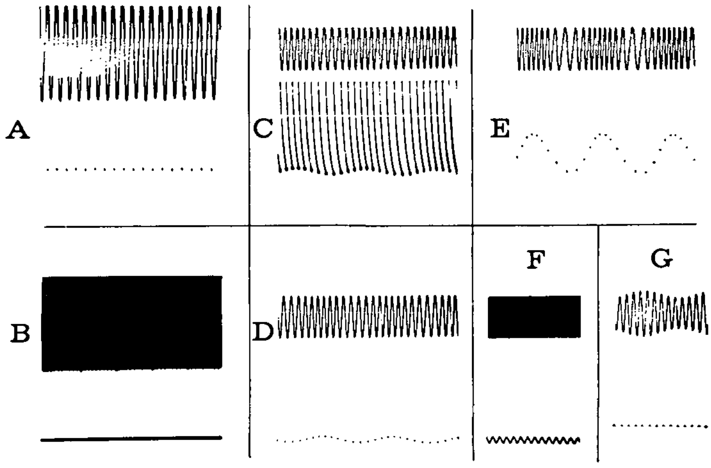

Fig. 2: Sample ZCA displays

Fig. 2 shows displays produced by different waveforms. Panel A shows the display produced by an unmodulated 1850 Hz tone at a fairly rapid sweep speed (2 ms / div, about 20 ms full scale). As in Fig. 1 the amplitude display is at the top, the ZCA display at the bottom. The display produced by a pure tone is simply a row of dots or, at slow sweep speeds (50 ms / div, Panel B), a line, whose downward displacement from the baseline is directly related to inter-crossing time (i. e., inversely related to instantaneous frequency). Panel C shows the effect of 100 Hz frequency modulation imposed on a 1000 Hz carrier, at a 5 ms / div sweep speed; the retrace lines are visible, as in Fig. 1. Panel D shows the ZCA display with retrace lines suppressed, as in normal operation: FM is displayed as a sinusoidally varying row of dots. Panel E shows the effect of deeper FM, at the same 100 Hz frequency. Panel F shows the effect of slowing the sweep speed from 5 ms / div (Panel D) to 50 ms / div (Panel F): the dots visible in Panel D coalesce into a sinusoidally varying line in Panel F. Panel G shows the display produced by a 1000 Hz carrier, amplitude modulated at 100 Hz: the ZCA is totally insensitive to AM and represents the amplitude-modulated tone in the same way as a tone of fixed amplitude.

Frequency Scale

The ZCA circuits we have used allow some flexibility in the choice of frequency scale. For a linear scale, with high frequencies at the top and low at the bottom of the display, the upward displacement of the row of dots representing a simple tone should be proportional to $1/t$, where t is the period of the tone. As explained in Appendix A, the displacement in our device is proportional to $\exp((c-t/\tau))$, which is a decreasing, negatively accelerated function that can be made to approximate $1/t$ over a limited range. Furthermore, if t, the period, is large with respect to c, the fixed write/reset time, and τ, the time constant of the R-C circuit, is also large, then the scale is approximately a linear function of period. Consequently, by means of appropriate adjustments of the R-C time constant, approximations to linear frequency or linear period scales can be obtained. In our experience the identifiability of displays does not depend critically on the linearity or otherwise of the frequency scale.

A convenience of the zero-crossings method is that the frequency scale is independent of the duration of signal that is sampled (which is determined entirely by the X-axis sweep speed). This is not the case with the sound spectrograph, where it has been argued that a logarithmic scale should be employed so that equal frequency ratios (which define equal musical intervals in the well-tempered scale) will appear equal, independently of the playback speed (Marshall 1977). While an approximation to a logarithmic scale can be obtained over a sufficiently limited range with our device, such a scale is not required to enable comparison of samples scanned at different speeds.

All the figures that follow were made using the exponential period scale just described. However, there is nothing in the principle of zero-crossings analysis that requires such a scale, and it may be convenient in future work to construct devices with linear, logarithmic, or some other frequency scale.

Apparatus and Method of Use

The principle of operation illustrated in Fig. 1 can be implemented in a number of ways. Greenewalt used modified commercially available counters and timers, which is expensive and does not lend itself to our method of display. It is possible to use a computer, which has advantages if it is desired to digitize and store the pattern of zero crossings. However, we have used a simpler and cheaper analogue method: an electronic device senses positive-going zero crossings using a conventional comparator circuit. The comparator output activates logic circuits that briefly intensify ("Z-modulate") the oscilloscope beam, then reset it to baseline (AA in Fig. 1), whereupon it begins to move downwards (although not visible) at a rate determined by the resistance setting in a R-C circuit.

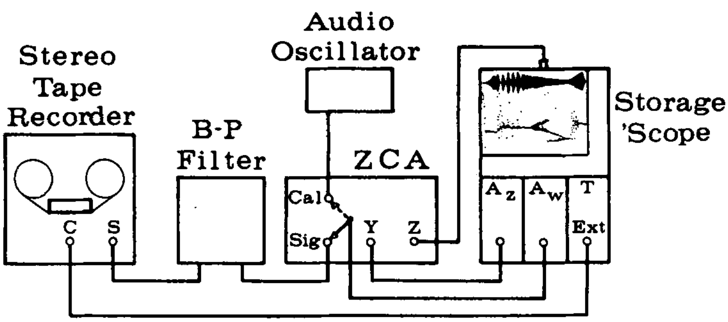

Fig. 3: Arrangement of components for fine-structure sound analysis using the zero-crossings analyzer

The ZCA is incorporated into the arrangement shown in Fig. 3. At the left is a stereo tape recorder (Toshiba model PT862D or Uher model 4200), with the signal to be analyzed recorded on one tape track (S) and a cue signal or signals recorded on the other track (C). The signal output from the tape recorder pre-amplifier goes, via a Krohn-Hite bandpass filter, to the input of the ZCA and also, via an internal switch in the ZCA, to channel A_w of the dual-beam oscilloscope. Thus the waveform of the signal being analyzed by the ZCA is always displayed on channel A_w. There is an alternate input to the ZCA labeled "CAL" (calibrate), which is connected to a standard audio oscillator. With the switch in the CAL position, sine waves of known frequency can be fed into the ZCA, and their waveforms will appear on channel A_w

(upper beam). The ZCA has two outputs: "Z" is connected to the Z-axis input of the oscilloscope and controls the intensity of the lower beam, briefly intensifying it at each positive-going crossing. The other output, "Y," is connected to the second channel, A_z, of the oscilloscope and controls the vertical displacement of the lower beam. The second, cue channel, C, of the stereo tape recorder is connected to the "External trigger" input on the time base, T, of the oscilloscope.

In normal operation the oscilloscope sweep is triggered by a brief pulse previously laid down on the cue track, C, of the stereo tape. We have used Tektronix time base no. 5B12N which has a feature allowing an adjustable delay between trigger onset and the appearance of the sweep. Sweep speed and delay time are independently adjustable, so that any portion of a displayed signal can be examined in as much detail as desired. This feature is indispensible for the analysis of fine structure.

Using a storage oscilloscope, the delay and sweep times can be adjusted until a display of the desired song segment is produced. The stored display can then be photographed. If a nonstorage oscilloscope is used, the cue-trigger feature allows a specific signal segment to be photographed simply by leaving the camera shutter open as the sweep is triggered. Since recorded signals can be scanned very rapidly with this system, it is convenient to be able to take large numbers of pictures in rapid succession, and we have used a specially mounted 35 mm camera for this purpose. The 35 mm negatives are compact and easily stored, and positive (black-on-white) copies in contact size or larger can be simply made using positive printing paper that requires only normal development techniques (Kodagraph Super-K projection positive paper: KPP5).

The signals displayed in the figures that follow were relatively free of noise and required little or no prefiltering. However, if the signals to be analyzed are accompanied by substantial background noise, especially high-frequency noise, some bandpass filtering may be required to obtain a clear display.

ZCA and Sound Spectrograph Compared

Fig. 4 shows an example of simple, "whistled" song, displayed on both ZCA and sound spectrograph. The Fig. shows several repeated phrases from the song of the Prothonotary Warbler (*Protonotaria citrea*). The song is essentially a pure tone, and frequency modulation is relatively slow. Hence ZCA

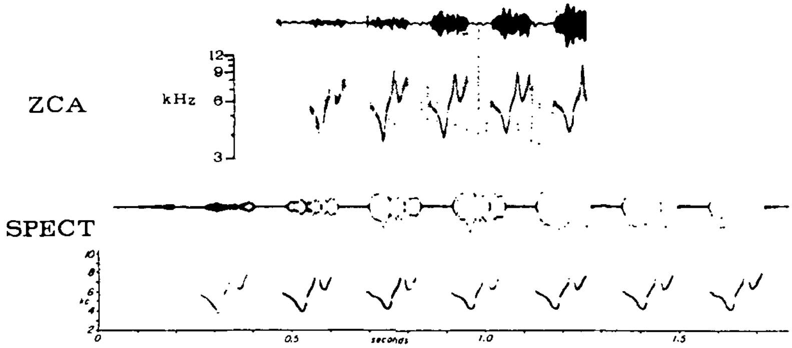

Fig. 4: ZCA and spectrographic displays of the song of the Prothonotary Warbler. Amplitude envelopes appear above each display

and spectrographic representations are closely similar. The spectrogram and associated amplitude plot are taken from Greenewalt (1968; Fig. 32. See Appendix B for other technical details). Small differences between his amplitude plot and ours reflect both filtering (his signal was high-pass filtered above 300 Hz) and degradation in the recording and transcribing process. It is worth noting that a zero-crossings analysis is much less sensitive to these kinds of distortion (which are generally proportional to signal amplitude) than is any kind of harmonic analysis.

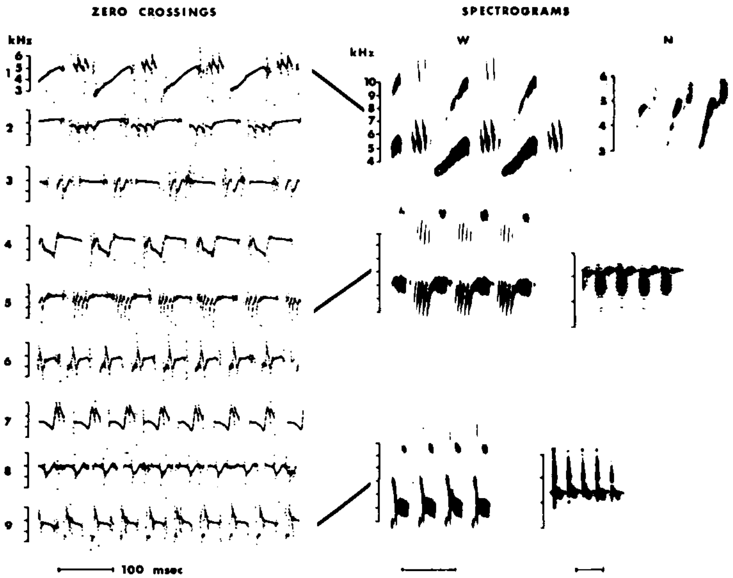

Fig. 5: ZCA and spectrographic displays of songs of the Slate-colored Junco

Fig. 5 shows some more tonally simple examples. The left side shows displays of 9 song variants of the Slate-colored Junco (*Junco hyemalis*) ordered by repetition rate, slowest at the top. Spectrograms of three of the songs, at the narrow (N) and wide (W) filter settings, are shown on the right and center. These songs are essentially pure tones and the harmonics visible in some of the spectrograms may be due to nonlinearities in transcription. The ZCA displays are quite clear and show a distinct family resemblance: all the songs consist of repeated phrases, each phrase being made up of a slowly modulated tone preceded or followed by one to five rapid frequency modulations. The similarities among the calls, and much of their internal structure, are largely lost in the spectrograms. Modulation rate is represented by side bands in the narrow-band spectrograms on the right. These modulations are represented as such by the wide-band spectrograms in the center, but frequency resolution is necessarily poor and this makes it hard to see much continuity between the displays at the top and bottom.

Analysis of Fine Structure

Because of the frequency/modulation-rate tradeoff inherent in the functioning of the sound spectrograph, it is not possible to subject small signal fragments to indefinitely fine analysis, even if slowed playback leaves the signal bandwidth within the useful range of the instrument. For example, a

500-ms song fragment played back at 1/4 speed will yield a 2s spectrogram, with the absolute frequencies reduced by a factor of four. However, no more temporal detail will be revealed than that present in the output obtained at normal playback speed. If the wide-band (300 Hz) setting is used, the frequency-modulation-rate sensitivity is about 330 Hz (CHERRY 1961). At the slowed playback speed, this means that FM as rapid as 1300 Hz in the original could be detected. However, the frequency resolution is degraded in exactly the same proportion: from 300 to 1200 Hz. There is no net gain in information unless the two spectrograms are compared — a time-consuming and not entirely satisfactory procedure.

Zero-crossings analysis is not subject to this limitation since instantaneous frequency is estimated cycle by cycle, with no integration time. The clarity of the display actually improves as sweep rate (hence FM resolution) is increased. This is because dots produced by zero crossings of the signal are superimposed at slow sweep speeds (as in Fig. 2, panel F), whereas "noise" dots, being much less frequent, are not. Hence displays at slow (< 100 ms / div) sweep speeds are sometimes rather "noisy" looking, unless the playback level is adjusted so that the noise between signal segments is below the threshold of the zero-crossing detector. However, as sweep speed is increased, dots due to the signal lose their apparent density only slowly, but "noise" dots are rapidly diluted; hence the display becomes clearer and clearer as sweep speed is increased.

Fig. 6 shows examples of how selected song fragments can be expanded by using high sweep speeds. The Fig. shows the display produced by a 2.5 s song of the Lapland Longspur (*Calcarius lapponicus*), one of the most intricate whistled songs. The signal amplitude envelope is shown at the top and below it is a spectrogram, made with the narrow-band setting. Portions of the spectrogram above the dashed line represent harmonics that are probably due to nonlinearities in the transcription process. Below the spectrogram is a slow-speed ZCA display, and below that five ZCA displays of the indicated segments of the larger display produced at higher sweep speeds.

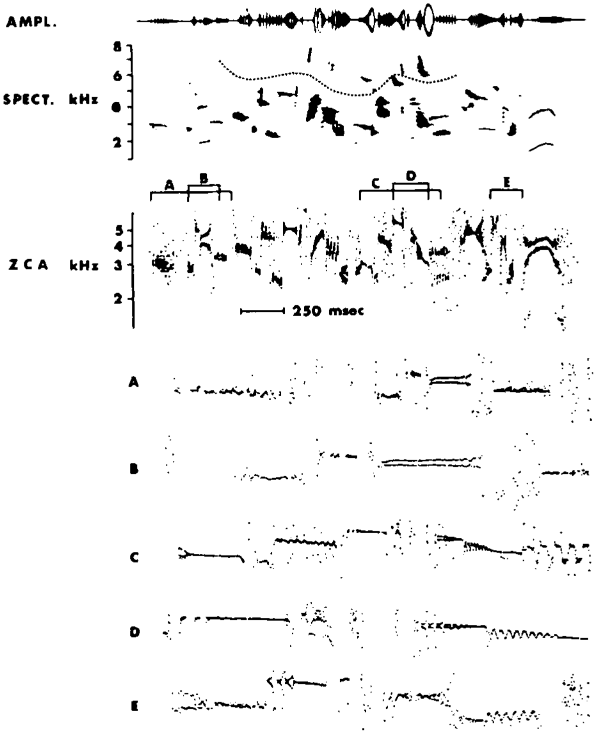

Fig. 6: ZCA and spectrographic representations of a song of the Lapland Longspur. Displays A—E are expanded segments of the indicated portions of the song

Several features of these expanded ZCA displays are of interest:
(a) Rapid FM: Displays C, D, and E all show examples of rapid, repetitive frequency modulation at high rates (200—250 Hz). We have found repetitive FM in many bird songs, although it is also totally absent from some, and we give more examples in a moment. Such rapid modulation cannot be reliably identified spectrographically.
(b) Double-line figures: Examples of this appear in displays A and B, and at the end of the large-scale ZCA display. Comparison with the spectrogram suggests that these figures are associated with the presence of two frequencies, here apparently in harmonic relationship. However, it should be emphasized that the instantaneous frequencies of these two lines are not simply related to the frequencies of the two components as revealed by harmonic analysis. Double lines resolve into an alternating pattern of dots at still higher sweep speeds, as we show in a moment. We return to the interpretation of double-line figures later.
(c) "Splitting" figures: Displays E and C (at the beginning) show figures resembling a letter "Y" or "V" on its side. These "splitting" figures are quite common in songs of the more versatile songsters, such as the Woodthrush, Mockingbird, and Song Sparrow. They may be related to the interaction between two independently modulated voices and are discussed more extensively below.

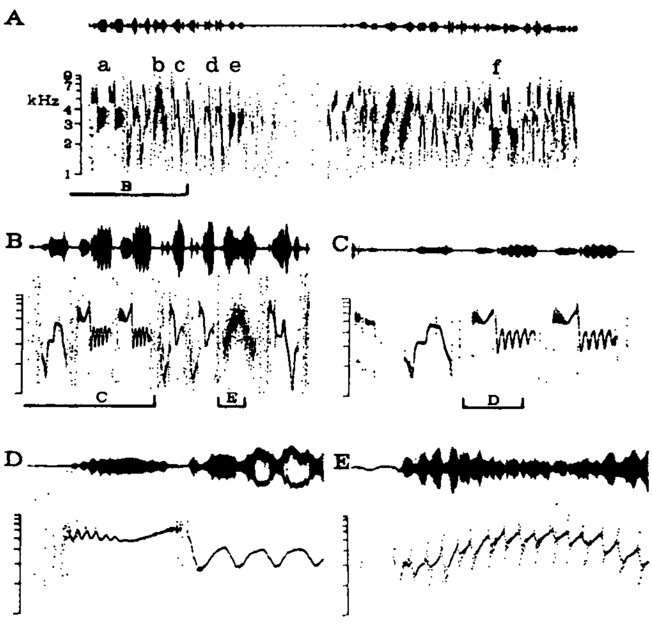

Fig. 7: ZCA analysis of a song of the Indigo Bunting. Whole song sequence appears in Panel A, expanded segments in the other panels, as indicated. Time scales (ms/division): A: 500; B: 100; C: 50; D: 10; E: 20

Fig. 7 shows more examples of fine-structure analysis, using the song of a much-studied bird, the Indigo Bunting (*Passerina cyanea*). This is a whistled song, so that the slow-speed display (e. g., panel B) looks much like a spectrogram of the song (Fig. 8). Panel A shows most of one song sequence and all of

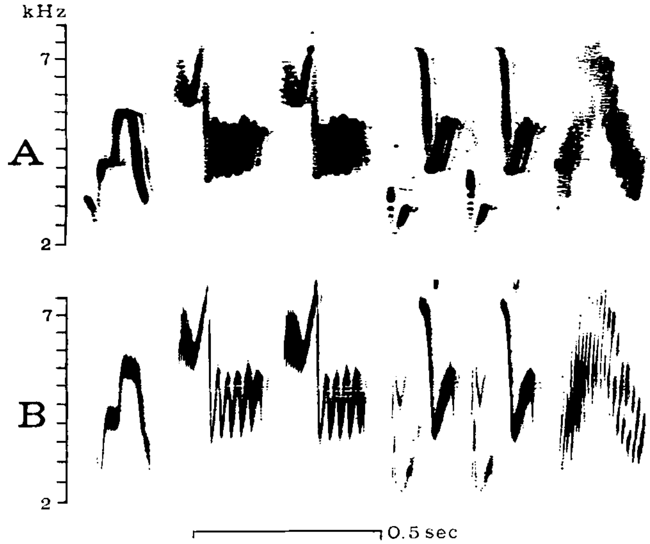

Fig. 8: Sound spectrograms of the Indigo Bunting song segment in Panel B, Fig. 7.
A: Narrow band. B: Wide band

a second sequence, with the amplitude envelope at the top and ZCA displays below. As is usual in this species, most song phrases occur in pairs or triplets. The phrases labelled a and f are similar in form, but the lower note[3]) is at a lower frequency in phrase f than in a. Panel B shows a, b, and c phrases enlarged. Panel C shows the two phrases a and part of the preceding figure at a higher sweep speed. The blurred low-frequency segment in Panel A is now resolved into FM, accompanied by AM, at a rate of about 80 Hz. Panel D

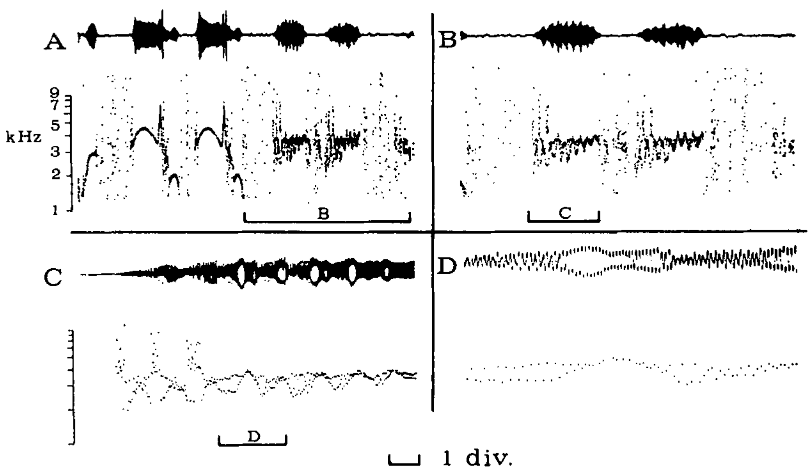

Fig. 9: ZCA analysis of phrases e in Panel A, Fig. 7.
Time scales (ms/div): A: 100; B: 50; C: 10; D: 2

[3]) There is no generally accepted terminology for describing the structure of bird song. Here we use the term *phrase* to denote the largest repeated unit in a song, or if no units are repeated, a song sequence occurring without a break. *Note* is used to denote any identifiable element in a phrase. The term *figure* is used to refer to characteristic patterns seen in many bird songs and reflecting the special properties of the ZCA.

shows a further magnification of one figure a: now the blur at the beginning of the high-frequency phrase is resolved into rapid FM at a rate of about 200 Hz. Panel E shows phrase b in Panels A and B at higher magnification. The repeated "splitting" patterns here suggest two-voice interactions (see below).

The rapid FM at the beginning of phrase a apparent in Panel D, and the slower modulation in phrase b (Panel E), are just visible in the wide-band spectrogram in Fig. 8. However, the modulation in a is lost in the narrow-band picture, and the subtle differences between the modulations of a and b that are apparent in Panels D and E of Fig. 7 are lost even in the wide-band picture in Fig. 8.

Panel A of Fig. 9 shows enlarged displays of phrases d and e in Panel A of Fig. 7. Panels B and C show further enlargements of the first of the two e phrases. Panel D shows an enlargement of an individual signal fragment from Panel C so as to reveal the underlying waveform. The rather complex e phrase can be seen to be made up of two alternating rows of dots, again suggesting interaction between two voices. The appearance of the waveform supports this inference.

Fig. 10 shows how songs that appear quite similar spectrographically may nevertheless show clear differences in fine structure. The column of displays on the left are brief song fragments of the Swamp Sparrow (*Melospiza georgiana*). If one looks just at those portions of the ZCA displays that are associated with substantial signal amplitude, it is clear that these songs have a very simple structure. Frequency shifts are generally monotonic (no repetitive FM); there is little evidence for two-voice interactions, and little amplitude modulation. On the right are brief song fragments of the Song Sparrow (*Melospiza melodia*). In contrast, most of these show repetitive FM as well as AM, and fuzzy or double-line displays, suggesting two-voice interactions. The Song Sparrow phrases differ from those of the Swamp Sparrow in both their greater variety and their richer fine structure. Sound spectrograms of the sets of songs from which these fragments are taken do not show these differences (see examples in Marler and Peters 1977).

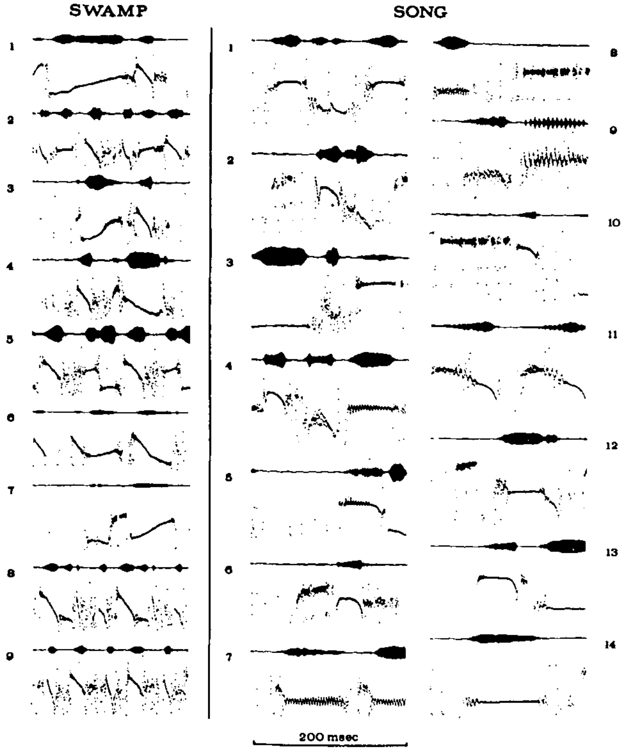

Fig. 10: A selection of 200-ms song fragments of Swamp and Song Sparrows, to illustrate differences in fine structure revealed by ZCA

Two-Voice Phenomenon

It is well known that most birds possess two separate acoustical sources, one in each bronchus, each with its own musculature and innervation. Moreover, there is now ample acoustical evidence that many song birds, and even some nonpasserines, can modulate these two sources independently (BORROR and REESE 1956; GREENEWALT 1968; MILLER in press; STEIN 1968).

Some two-voice interactions can be detected in spectrograms. However, if the two voices are relatively close in frequency, this may not be possible and more elaborate techniques, such as selective filtering and slow-speed playback (permitting aural identification of the two voices), are required.

Perhaps the most extensive analysis of two-voice effects is due to GREENEWALT (1968), who has also published a phonograph record of filtered and unfiltered versions of the songs he has analyzed. This has provided us with an excellent opportunity to compare ZCA displays of these songs with GREENEWALT's analyses to see if two-voice interactions can be easily identified in ZCA displays. The results of these comparisons show that two-voice interactions often produce quite distinctive ZCA figures.

Fig. 11 shows a ZCA analysis of the call of the Naked-throated Bellbird (*Procnias nudicollis*), taken from GREENEWALT's record. (This Fig. should be

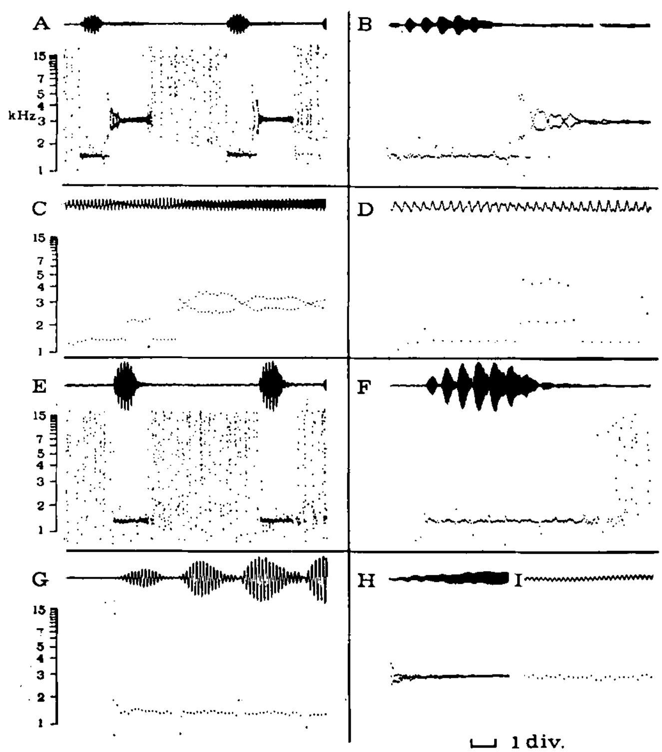

Fig. 11: Two-voice interactions in the Naked-throated Bellbird, from GREENEWALT's record. A: Two songs (100 ms/div); B: First song, magnified (20 ms/div); C: Same (5 ms/div); D: Same (2 ms/div); E: Low voice, isolated by filtration (100 ms/div); F: Same (20 ms/div); G: Same (5 ms/div); H: High voice, isolated by filtration (20 ms/div); I: Same (5 ms/div)

compared with Fig. 49 in GREENEWALT's book.)[4]) Panel A shows two successive unfiltered calls at a relatively slow sweep speed. Panel B is a magnified display of the transition between the low, high-amplitude voice and the high, low-amplitude voice in the first call. Panels C and D are further magnifications of the transition. The gradual emergence of the higher frequency component is quite apparent in the waveform display in Panel D. Panel E shows the two calls low-pass filtered at 2.5 kHz to show the low voice in isolation. Panel F shows the first call at higher magnification. Both panels show little variation in frequency for the low voice, but strong amplitude modulation. This is even more apparent in Panel G, which is at high magnification to show the waveform. Panels H and I show the high voice (extracted by high-pass filtering at 2.5 kHz) at two different sweep speeds. Its amplitude is low, and there is no amplitude or frequency modulation.

The double-line figure in Panels A—C is evidently characteristic of two-tone mixtures. "Splitting" patterns, as at the onset of the high voice in Panel A, may also be characteristic of the introduction of a second tone. Fig. 12 shows another example that supports these conjectures. Panel A shows a 1 s segment of the song of a Mockingbird (*Mimus polyglottos*) taken from GREENEWALT's record (compare with his Fig. 42). The rather striking "barred" display is produced by the interplay of two voices, each of which is essentially a pure tone. Panels B and C show segments of the display at successively higher magnifications. The onset of the higher voice is quite apparent in the waveform, shown at the highest magnification in Panel C. It is clear from Panel B that the onset of the high tone is signalled in the ZCA display by a "splitting" pattern; the fact that the barred pattern appears on the high limb of the split, while the low limb fades out, shows that it is the high tone which is increasing in relative amplitude.

The right-hand panels of Fig. 12 show the two voices separately. Panel D is the unfiltered transition (seen in Panel C, but repeated here for ease of

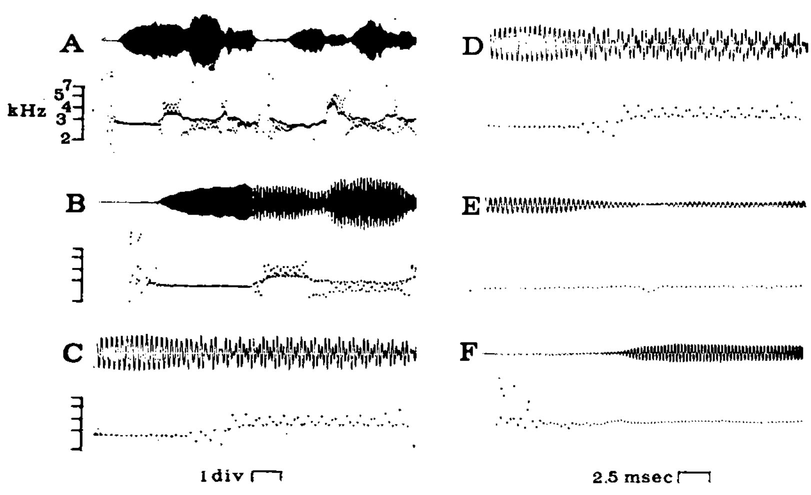

Fig. 12: Two-voice interactions in the Mockingbird, from GREENEWALT's record. A: Whole sequence (50 ms/div); B: First transition (10 ms/div); C: Same, magnified (2.5 ms/div); D: Unfiltered first transition; E: Low voice, isolated by filtration; F: High voice, isolated by filtration. D, E, and F at 2.5 ms/div

[4]) Note that the high voice appears at the bottom of GREENEWALT's Fig. 49, and the low voice at the top.

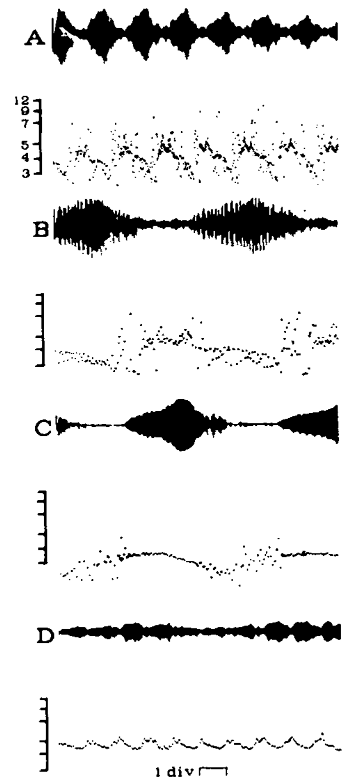

Fig. 13: Two-voice interactions in the Wood Thrush, from GREENEWALT's record. A: Unfiltered sequence (25 ms/div); B: Same, magnified (6.25 ms/div); C: Low voice, isolated by filtration (6.25 ms/div); D: High voice, isolated by filtration (6.25 ms/div)

comparison). Panel E shows the low voice (separated by filtration above 3.2 kHz) as it decreases in amplitude. Below it is the high voice (separated by filtration below 3.2 kHz) gradually increasing in intensity.

Fig. 13 shows a more complex example, in which both voices are independently modulated. Panel A shows a 250 ms fragment of Woodthrush (*Hylocichla mustelina*) song from GREENEWALT's record (compare his Figs. 53—56). Panel B shows a 62.5 ms fragment of the song in Panel A, and while there is a hint of the kind of splitting seen in the other two-voice interactions, the display is rather confused. Panels C and D show that the confusion is the result of simultaneous independent modulation of the two voices. Panel C shows the low voice (separated by filtration above 3.8 kHz), which has the higher amplitude and is modulated (in both frequency and amplitude) at the slower rate. Panel D shows the higher voice which shows only irregular amplitude modulation, but clear frequency modulation at quite a high rate (ca. 144 Hz).

Finally, Fig. 14 shows a rather striking example of the interaction between two voices in the song of the Lapland Longspur (GREENEWALT, Fig. 130). Panel A is a 250 ms fragment in which the barred splitting display is caused largely by a slow decrease in the amplitude of the high voice relative to the low, succeeded by a rapid increase in the amplitude of the high voice. The amplitude changes in the high voice are shown separately in Panel B, and the low voice is shown in Panel C.

We have noticed that the bird songs that are most interesting musically are often associated with the most intricately patterned and aesthetically pleasing ZCA displays. Figs. 12—14 are good examples of this correlation that the reader can test for himself by listening to GREENEWALT's slow-speed recordings of these songs.

These examples, and others not shown here, show that "splitting" patterns in the ZCA display are a fairly reliable sign of the interaction between two voices. Absence of such a pattern signifies little, however, since confused and "noisy" displays may just be the result of two voices that are both varying in frequency (cf. Fig. 13), or of multiple harmonics (see below). "Barred" displays seem to be produced when the two voices are of constant frequency and similar amplitudes; such displays can be produced by mixing equal-amplitude sinusoids of two different frequencies. However, more research is needed to map out the quantitative details of such two-tone interactions and

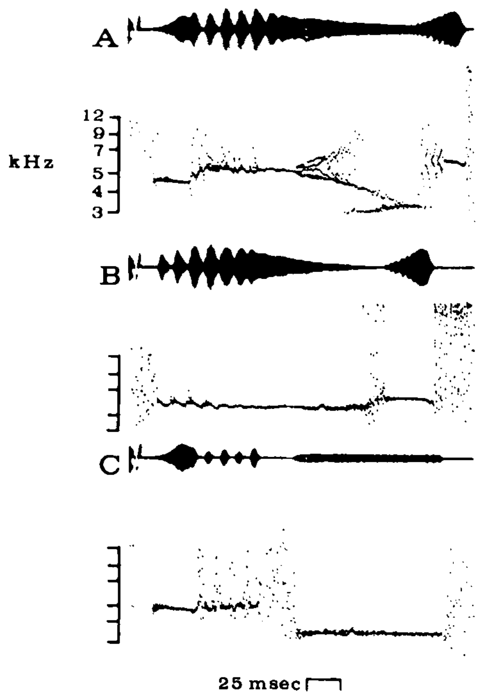

Fig. 14: Two-voice interactions in the Lapland Longspur, from Greenewalt's record. A: Unfiltered sequence; B: High voice, isolated by filtration; C: Low voice, isolated by filtration

the dependence of displays on the relative amplitudes, frequencies, and phase relations of the interacting tones.

Complex Sounds and Harmonics

One reason that the ZCA works so well with "whistled" bird song is the almost complete absence of harmonics. There are two difficulties in applying the technique to complex sounds. The first is that some information is necessarily lost by looking just at zero crossings and amplitude envelope—recall the confusion among triangle, square, and sine waveforms. However, both triangle and square waves are artificial creations, approximated by few natural sounds because they contain very high frequency components. Thus, although there is an uncertainty involved in going from zero-crossings (and amplitude envelope) information back to the originating waveform, in practice the ambiguity may be quite small. Empirical support for this belief comes from elegant work on synthetic animal sounds by Dörrscheidt (1973, and in press). Dörrscheidt notes that a signal sample can be considered as the product of a frequency-modulated carrier, c(t), and an amplitude function, a(t), where a(t) is always positive and $-1 \leqq c(t) \leqq 1$.

This formulation implies that the zero crossings of the signal are the same as those of the carrier; hence c(t) can be estimated from the sequence of zero crossings actually obtained: a(t) is then obtained by division, using the postulated relation $s(t) = a(t)c(t)$, where s(t) is the original signal. The amplitude function, a(t), must carry a lot of information about harmonics. In practice, however, its frequency range is limited by the exigencies of analogue-to-digital conversion. Nevertheless, Dörrscheidt reports that acceptable resyntheses of insect and bird calls are produced by this procedure, even if the original calls have substantial harmonic content.

It is likely, therefore, that much of what needs to be known about animal sounds is contained in the sequence of zero crossings and the peak-amplitude envelope. The critical task is to extract the information. This poses two related problems: one is the informational value of features revealed by zero-crossings analysis. This is a behavioral question that is yet to be explored. The second is the relation between ZCA displays and the spectral properties of the generating signal. This is a mathematical problem for which no general solution is available. However, we have carried out some preliminary work, both theoretical and experimental, involving pure-tone mixtures with various

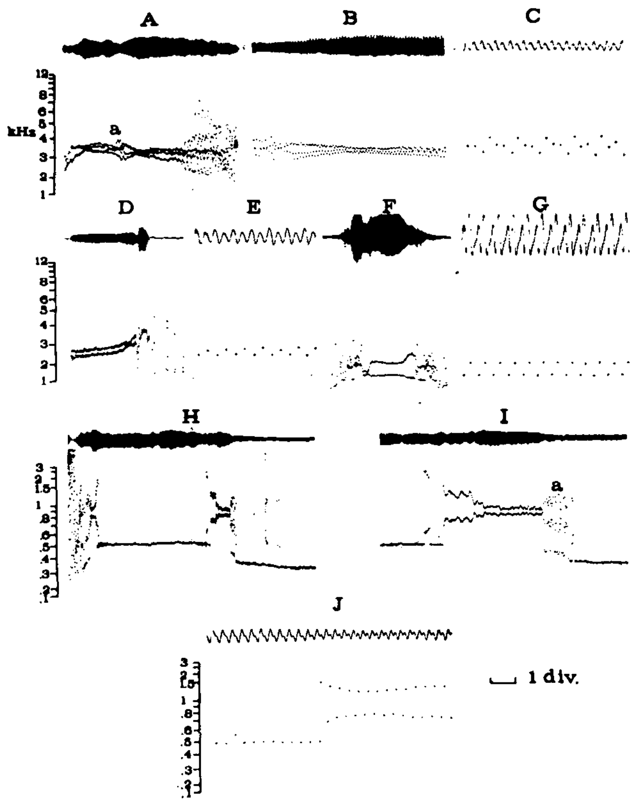

Fig. 15: Multiple-line figures. A: "Scold" call of a Bluejay (50 ms/div); B: Same, magnified from a in Panel A (10 ms/div); C: Same, magnified (1 ms/div); D: Call of Holboell's Grebe (50 ms/div); E: Same, magnified (1 ms/div); F: Call of the Whistling Swan (50 ms/div); G: Same, magnified (2 ms/div); H: Wolf call with two "yodels" (500 ms/div); I: Second yodel, magnified (100 ms/div); J: Same, magnified (5 ms/div)

amplitudes and phase relations. On this basis it is possible to make some qualitative generalizations about strictly periodic signals: (a) The presence of more than one line (row of dots) in the ZCA display indicates the presence of more than one frequency component. However, just two components may give rise to more than two lines under some conditions. (b) The instantaneous frequency of each line in a multiple-line display is not simply related to the frequencies of the components. (c) For two-tone mixtures, the location of the lines depends on all three parameters of the tone complex: the relative frequencies, amplitudes, and phases of the components. This result is in sharp contrast to the sound spectrograph, where the location of a line depends solely on frequency (although, of course, lines may appear artifactually, as sidebands caused modulations too rapid for the time resolution of the filter). (d) For tone pairs, graded changes in relative amplitude generally produce graded changes in the location of lines; however, graded changes in phase can produce sudden changes in the location of lines.

Complex sounds sometimes give rise to confused, "noisy" displays, and sometimes to patterned displays. Confused displays are more likely if the components are varying in amplitude or frequency. However, if the components are of relatively constant amplitude and phase, and there are not too many of them, a display consisting of multiple lines may be seen. Fig. 15 shows examples of such displays, all associated with calls having a harsh or plaintive quality. Panel A is the "scold" call of a Bluejay (*Cyanocitta cristata*). Panel B shows the same call at a higher magnification, and Panel C shows the waveform of the first 7 ms of Panel B. Panel D shows a call of Holboell's Grebe (*Colymbus grisegena holböllii*) and Panel E shows a segment of the waveform. Panels F and G show the call of the Whistling Swan (*Olor columbianus*) and a waveform sample. The last three panels show that these multiple-line displays are not restricted to birds. Panel H shows the howl of a Timber Wolf with a "yodel" phrase in the middle. The yodel is magnified in Panel I, and its waveform is shown in Panel J. The "jump" shown by the lines in the Wolf call is associated with an increase in the relative amplitude of the higher frequency component and with a shift in its phase, as can be seen from the waveform in Panel J.

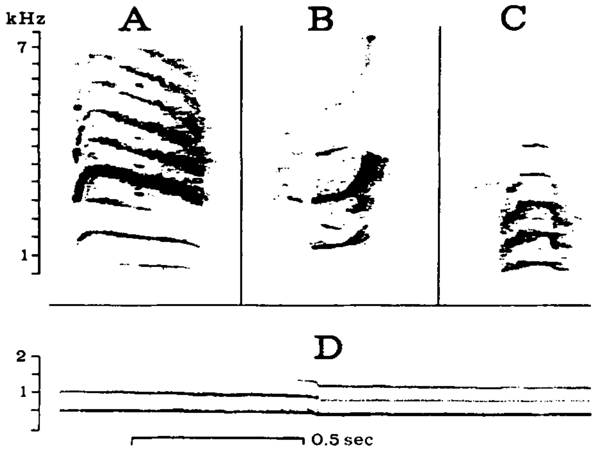

Fig. 16: Sound spectrograms of the calls shown in Fig. 15. A: Bluejay; B: Holboell's Grebe; C: Whistling Swan; D: Wolf

Fig. 16 shows spectrograms of the sounds whose ZCA displays appear in Fig. 15. All consist of a number of components that are apparently harmonically related. The Bluejay call (Panel A) contains 9 recognizable components, all but the lowest following a similar temporal course: an initial rise, followed by a slow decline in frequency. The strongest component begins its decline at about 3 kHz, corresponding to the dominant instantaneous frequency in the ZCA plot. The call of Holboell's Grebe (Panel B) contains 4 recognizable components (some are barely visible in this high-contrast reproduction), the second beginning at about 2.5 kHz—the dominant ZCA frequency (Fig. 15, Panel D)—is much the strongest. The call of the Whistling Swan (Panel C) contains 6 recognizable components, the first three, ranging from 700 Hz to about 2300 Hz, are substantially stronger than the rest. This compares with the two lines at 1200 and 2000 Hz in the ZCA plot (Fig. 15, Panel F). The Wolf call (Panel D) contains three main components and the complex pattern shown in the ZCA plot (Fig. 15, Panels I and J) seems to correspond to a sharp decrease in the frequency of all components (at a in Panel I), accompanied by changes in relative amplitude and phase that cannot be estimated from the spectrogram. The upper and lower frequencies on either side of the transition in Panel I of Fig. 15, 350 and 480 Hz, seem to correspond to the frequencies of the lowest spectral component in Fig. 16, Panel D. It is worth noting that the subjective impression produced by the Wolf "yodel" is quite striking, and in this respect the large changes shown in the ZCA plot are closer to the phenomenal impression than the relatively modest frequency shifts apparent in the spectrogram.

The only general conclusion that can be drawn from these comparisons is that the ZCA display preferentially represents the strongest component in a complex signal. However, if more than one line is seen, the instantaneous frequencies do not bear a simple relation to the spectral components.

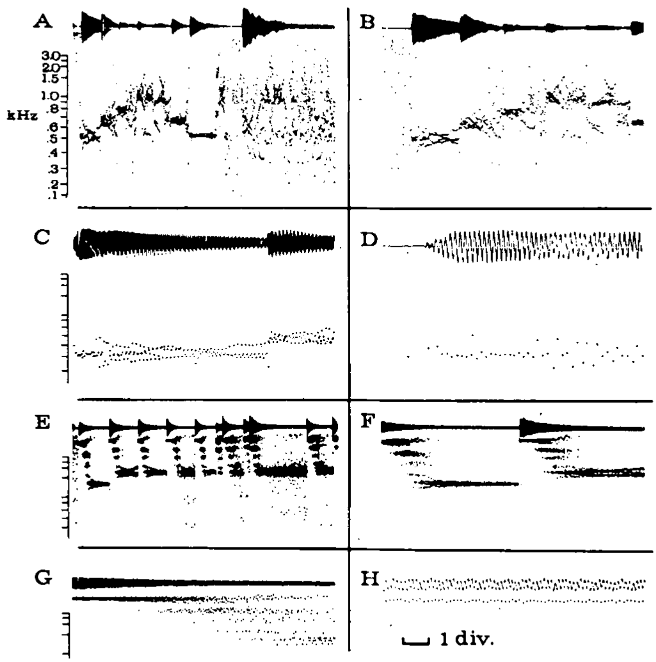

Fig. 17: Musical sounds rich in harmonics. A: Piano (500 ms/div); B: Same, magnified (200 ms/div); C: Same, magnified (50 ms/div); D: Same, magnified (10 ms/div); E: Bells (500 ms/div); F: Same, first two notes (100 ms/div); G: First note (10 ms/div); H: Same, magnified (1 ms/div)

Apart from the sudden shifts owing to the appearance of a new component, the sounds depicted in Figs. 15 and 16 are sustained and show few abrupt changes in amplitude. However, many natural sounds show a decay in strength over time, with different harmonics dying out at different rates. The ZCA displays produced by such sounds can be quite complex, as illustrated in Fig. 17. Panel A shows an 8-note sequence, ending in a chord, played on a piano. The amplitude envelope shows the exponential decay characteristic of this percussive instrument. The intricate pattern of lines in the ZCA display is caused by the constantly shifting amplitude relations among the components as the sound dies away. The tonal quality of the chord is not at all obvious from the display, because of the complexity of these interactions. Panels B, C and D show the single-note sequence at increasing magnifications, showing that the complex display in Panel A is in fact highly structured. The last five notes of the piano sequence are also shown as spectrograms in Panel A of Fig. 18 (chord at right). The differential rate of decay of the numerous harmonics is easily seen.

Panels E—H of Fig. 17 show ZCA displays of a sequence of four single notes played on bells. The display here bears a qualitative similarity to the wide-band spectrogram in Panel C of Fig. 18. Panel H in Fig. 17 suggests that

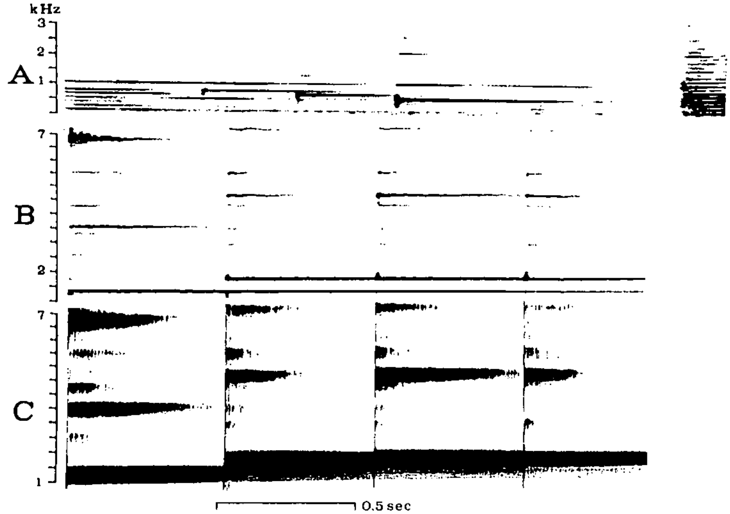

Fig. 18: Sound spectrograms of the musical sounds in Fig. 17. A: Piano, chord at right (narrow band); B: Bells, first four notes (narrow band); C: Bells (wide band)

this is owing to the presence of a relatively low-frequency component (ca. 1400 Hz) of substantial amplitude that has the effect of "sweeping" the waveform up and down across the zero axis. In this way even low-amplitude higher frequency components can make some contribution to the pattern of zero crossings, and the effect of phase differences is largely eliminated.

The practical implications of these comparisons are straightforward. If the signals of interest consistently produce single-line ZCA displays and sound tonally pure when played at the appropriate speed (i. e., not too fast for higher harmonics to be heard), then the ZCA display can be relied upon as a faithful record. Most passerine song falls into this category. If multiple-line or broad-band displays are produced, the ZCA is probably still useful, but spectrographic comparisons should be made to see that important features are not being missed. Very complex signals (such as the piano chord in Panel A of Figs. 17 and 18) may give rise to complex ZCA displays that cannot be distinguished from noise. In such cases the ZCA cannot be used unless the information of interest lies in a restricted frequency band that can be isolated by filtering.

Research Possibilities

The ZCA displays suggest a number of possible questions for future behavioral research. The largest number arise from the fine structure that the technique immediately reveals in most bird song. Such questions could have been raised some years ago, of course, since GREENEWALT's book contains many examples of rapid temporal variations in song. However, the method he employed in addition to being very slow also requires formidable technical resources, and no doubt the lack of these prevented many from following his lead. Moreover, at that time the more obvious features of bird song, such as its intensity and frequency range, and the timing and ordering of phrases, had not been extensively studied. In recent years, however, the work of EMLEN (1972), FALLS (1969), GOTTLIEB (1974), SHIOVITZ (1975) and others has pointed increasingly to the importance of features at the level of the individual phrase and note. For example, in summarizing his study of Indigo Bunting song, EMLEN writes: "The structure or morphology of individual notes is also important in song identification. I hypothesise that notes must have a particular acoustical quality characterised by their covering a wide frequency range ... in a short time interval ... and by their containing abrupt changes and reversals of pitch. A wide diversity of Bunting notes fit this description, and differences in details of their fine structure probably provide the variability necessary for permitting individual identification" (1972, p. 165). In a follow-up study, SHIOVITZ (1975) found that " ... territorial display and attack behaviors are dependent on the fine structure of the Indigo Bunting figures" (p. 171).

The ZCA does not suffer from the disadvantages of GREENEWALT's method, since it is both rapid and relatively inexpensive. It provides a technique ideally suited to the study of rapid frequency and amplitude modulation at a time when the importance of fine structure is becoming widely recognized. The method suggests a number of behavioral questions. We briefly discuss four relating to fine structure out of a number of possibilities.

SHIOVITZ (1975) in an extensive study of song recognition did a number of playback experiments to explore the communicative significance of features of the song of the Indigo Bunting. In one experiment (p. 152, Fig. 27) Bunting song phrases were imitated by a human whistler, and then played back to the birds at a speed appropriate to the model phrases. Apparently the imitations sounded quite similar to the originals to most human ears, although highly skilled listeners could detect differences. However, the spectrographic displays of the imitations do not differ in any obvious way from those of the models. One of the phrases in the imitation is very similar to phrases a and f in Fig. 7 (Panel A); others are similar to phrases b and g. However, none of the imitations contains the rapid FM characteristic of these phrases, as displayed by the ZCA: at the beginning of phrases a and f (Panel D of Fig. 7), and throughout phrases b and g (which probably also involve two-voice interactions). Imitating such rapid modulations would be difficult for a human whistler, even with the allowed speed reduction; but since the modulations are not visible in SHIOVITZ's spectrograms, presumably no imitation was attempted. SHIOVITZ reports that the whistled imitations were less effective in eliciting agonistic behavior than the natural songs. An obvious possibility, therefore, is that the fine structure revealed by zero-crossings analysis may be necessary for the song to have its full effect.

A subsidiary point concerns the covariation of amplitude and frequency modulation (cf. GREENEWALT 1968; STEIN 1968). The terminal note in phrase a (e.g., in Panel D of Fig. 7) involves both AM and FM. Possibly both play a role in the communicatory function of this phrase. Since the usual spectrographic display does not allow for simultaneous recording of amplitude and frequency information, the role of AM has probably often been neglected (cf. BEER 1970). These conclusions about the possible role of fine structure are consistent with SHIOVITZ's other results, on the basis of which he writes that: "The facts suggest that components within the song figure may be supplying recognition cues ..." (1975, p. 135).

A second example is provided by Fig. 10, which shows a collection of fragments from the songs of Swamp and Song Sparrows, two closely related, sympatric species. The Song Sparrow examples nearly all show rich fine structure: rapid FM and AM, and evidence for intricate two-voice interactions. The Swamp Sparrow examples, in contrast, are all simple and rather stereotyped in form. MARLER and PETERS (1977) have recently shown that young Swamp Sparrows are able to learn selectively Swamp Sparrow song phrases when exposed to phrases from both Swamp and Song Sparrows arranged in various temporal patterns. It is likely, therefore, that differences in fine structure may play a role in this discrimination: for example, the simple, monotonic frequency modulation of the Swamp Sparrow phrases may facilitate learning by the young Swamp Sparrows, or the elaborate modulation and two-voice structure of the Song Sparrow songs may exert a depressive effect.

A third example relates to the evidence for two-voice interactions in displays of many songs, especially those of "virtuoso" songsters such as the Song Sparrow, Woodthrush, and Mockingbird. In the study of human speech, emphasis is often placed on the role of the production mechanism in perception (FLANAGAN 1972). The idea is that much of speech perception is devoted to the detection of acoustic features associated with particular muscular operations involved in speech production: this is "analysis by synthesis." Extension of this line of thought to bird song suggests that two-voice interactions of different sorts may play an important role in communication. Since many two-voice effects are readily detectable in ZCA displays, the method may prove useful in uncovering their communicative significance.

A final example comes from recent work on species identification in the Brown-headed Cowbird (*Molothrus ater*). During the breeding season female Cowbirds show a rapid and distinctive copulatory response to male song (KING and WEST 1977). Using this response as a bioassay, WEST, KING, STADDON, and EASTZER (unpubl.) have shown that a particular brief (ca. 50 ms) fragment of the approximately 500 ms song is critical to its effectiveness. Zero-crossings analysis of the filtered song has identified two features that appear to determine the effect of this brief feature: its amplitude relative to the rest of the song, and the presence of exceedingly rapid (ca. 500 Hz) frequency modulation.

A second class of problems where the ZCA may prove useful is the analysis of song variation, an essential component in studies of individual variation (e.g. BEER 1970; MUNDINGER 1970), geographical variation (e.g. BAKER 1975; LEMON 1975; KROODSMA 1974), and communication (e.g. GREEN 1975; MAURUS and PRUSCHA 1973).

The sound spectrograph requires so many individual adjustments that it is difficult to obtain a standard record of a given signal. Moreover, the information in the spectrogram cannot easily be converted to digital form for quantitative analysis, although some attempts have been made, using largely manual methods (FIELD 1976). In contrast, zero-crossings information is easily digitized and highly reproducible. The display is not affected by small variations in input amplitude (equivalent to variation in the threshold for detection of zero-crossings), and even large changes have mild effects; simply eliminating parts of the display, leaving the remaining figures unaffected. Thus, once the threshold level is even approximately set, and the frequency and time scales chosen, perfectly reproducible displays are easily produced.

A measure may be highly consistent, yet fail to measure features that are related to the function of a vocalization. The probable importance of the information revealed by ZCA displays rests on three kinds of test: (a) Aural impressions. The displays fare quite well on this score: when songs are suitably slowed down to allow for the lesser time resolution of human ears (cf. KONISHI 1969; PUMPHREY 1961; SCHWARTZKOPFF 1963), the rapid modulations visible in the display are heard as such, and double-line displays are associated with the sound of two separate voices. GREENEWALT's $^1/_8$ speed recording of the song of the Lapland Longspur shown in Fig. 6 provides a good demonstration of these correlations. (b) Rendering of similarities. A good display technique should show as similar songs that sound similar and come from similar sources. The 9 song variants of the Slate-colored Junco shown in Fig. 5 illustrate this point, since the family resemblance among them is quite apparent. Many other examples could be added. Capturing these resemblances quantitatively is obviously not a simple matter. However, the relative simplicity of the ZCA displays, the ease with which zero-crossings information can be digitized, and the possibility the displays offer of incorporating amplitude information, open up possibilities not easily available with the sound spectrograph. (c) Behavioral tests. The crucial test for any system that represents an information-carrying signal is obviously a behavioral one: does it faithfully depict those features of the signal known to have communicative significance, each with its proper weight in relation to the others and to the noninformation-carrying features? Obviously, the ZCA display awaits tests of this sort.

Summary

A simple, real-time method for displaying the information contained in the zero crossings of a bioacoustic signal is described. The technique yields reliable, highly-structured displays of most bird songs and is superior to the sound spectrograph in speed and reliability of operation, in the detection of rapid frequency modulation, and in permitting easy comparison of frequency and amplitude modulation. Structured displays are produced even by many signals rich in harmonics, but it is not yet clear how some of these should be interpreted. Zero-crossings analysis has obvious application in research on the communicatory function of fine structure, including two-voice interactions, and in the quantitative analysis of song variation.

Zusammenfassung

Eine einfache Echtzeitmethode zur Darstellung von der im Nulldurchgang eines bioakustischen Signals gegebenen Information wird beschrieben.

Dieses Verfahren erlaubt eine zuverlässige Feinstrukturdarstellung der meisten Vogelgesänge. Es hat sich in der Registrierung schneller Frequenzmodulation durch Schnelligkeit und Zuverlässigkeit dem Klangspektrographen überlegen erwiesen und gestattet einen einfachen Vergleich von Frequenz- und Amplitudenmodulation. Strukturdarstellungen lassen sich auch von solchen Signalen geben, die reich an harmonischen Obertönen sind; doch ist noch nicht geklärt, wie einige dieser Darstellungen zu interpretieren sind.

Nulldurchgangsanalysen bieten sich an zur Erforschung der kommunikativen Funktion der Feinstruktur von Lautäußerungen, einschließlich zweistimmiger Gesänge, sowie für das Gebiet der quantitativen Analyse von Gesangsvariationen.

Acknowledgements

We thank Joachim Dörrscheidt, Andrew King, and Meredith West for comments on the manuscript, and Peter Marler for supplying us with tapes of the Song and Swamp Sparrow songs shown in Fig. 10. Thanks are also due to the Anatomy Department at Duke for use of photographic facilities. We are grateful to Christine Varley, who prepared the Zusammenfassung.

Appendix A

Frequency Scale

The oscilloscope beam moves between an upper bound, when t, the intercrossing time, is 0 (line AA in Fig. 1), and a lower bound, the asymptote of Equation A. 1 (below), when t is large (line BB in Fig. 1). The displacement of the beam from the lower bound at intercrossing time t is given by:

$$D(t) = a. \exp(-(1/n - c)/\tau) \quad (A. 1)$$

where a is a scale factor, τ is the time constant of the R—C circuit, n is the instantaneous frequency ($= 1/t$), and c is the constant time taken up in intensifying the trace and resetting it at each positive-going crossing.

For the range of frequencies considered here, $1/n = t \gg c$, hence c can be neglected and the displacement function becomes:

$$D(t) = a. \exp(-t/\tau) \quad (A. 2)$$

In series expansion form, Equation A. 2 is:

$$D(t) = a(1 - t/\tau + t^2/2\tau^2 - \ldots) \quad (A. 3)$$

If τ is set to be large relative to t, terms in τ^2 and above can be neglected and D(t) becomes:

$$D(t) = a - at/\tau \quad (A. 4)$$

which is a linear function of t, the period of a simple sinusoid. Hence a good approximation to a linear period scale is easily obtained by appropriate adjustment of the R—C time constant.

A linear frequency scale requires that $D(t) = k/t$, where k is a scale factor. This obviously cannot be realized in practice by a device that measures period, since it implies that $D(0) = \infty$. However, it is reasonable to consider as an alternative the function:

$$D(t) = k/(1 + t) \quad (A. 5)$$

which equals k when $t = 0$, but is close to k/t when t is large (i.e., at relatively low instantaneous frequencies). Equation A. 5 is negatively accelerated, has an asymptote at $D(t) = 0$, intersects the D(t) axis, and (as can be seen from its series expansion) is linear when $t \rightarrow 0$. These characteristics are shared by Equation A. 2, so that the exponential period scale can provide an approximation to a linear frequency scale over a restricted range of an octave or so.

Appendix B

Sound Sources

Fig. 4. Prothonotary Warbler: Greenewalt (1968) record. Unfiltered.

Fig. 5. Slate-colored Junco: 9 song variants, from record by Borror (1967). Band-pass filtered from 70 Hz—10 kHz (Krohn-Hite filter at 24 dB/octave attenuation). Spectrograms made on Kay Sonograph at FL-1 setting. Wide-band records are at the 0.16—16 kHz range, narrow band records are at the 0.08—8 kHz range.

Fig. 6. Lapland Longspur: Greenewalt record. Band-pass filtered from 50 Hz—10 kHz. Spectrogram is narrow band; other details as in Fig. 5.

Fig. 7. Indigo Bunting: Peterson (1971) record. Unfiltered.

Fig. 10. A selection of songs from Swamp and Song Sparrows recorded in the vicinity of Millbrook, N.Y. The songs were selected only for diversity and recognizability for use in playback experiments (Marler and Peters 1977). They are otherwise quite representative of the songs of these two species.

Fig. 11. Naked-throated Bellbird: GREENEWALT record. Unfiltered.
Fig. 12. Mockingbird: GREENEWALT record. Unfiltered.
Fig. 13. Wood Thrush: GREENEWALT record. Unfiltered.
Fig. 14. Lapland Longspur: GREENEWALT record. Unfiltered.
Fig. 15. Bluejay, Holboell's Grebe, Whistling Swan: PETERSON record. Unfiltered. Wolf call: From record: The Language and Music of the Wolves. Natural History Magazine, New York 1971. Unfiltered.
Fig. 17. Piano, bells: From record: The Shure Trackability Test Record. Shure Inc., Evanston, Ill., No. TTR-101. Unfiltered.

Literature Cited

BAKER, M. (1975): Song dialects and genetic differences in White-crowned Sparrows (*Zonotrichia leucophrys*). Evolution **29**, 226—241 • BEER, C. (1970): Individual recognition of voice in the social behavior of birds. In: Advances in the Study of Behavior. (LEHRMAN, D., R. HINDE and E. SHAW, eds.) Acad. Press, New York, **3**, 27—74 • BORROR, D. J. (1967): Common Bird Songs. Dover, New York (Booklet and phonograph record) • BORROR, D. J., and C. R. REESE (1956): Vocal gymnastics in Wood Thrush songs. Ohio J. Sci. **56**, 177—182.

CHERRY, C. (1961): On Human Communication. Sci. Ed., New York.

DAVIS, L. I. (1964): Biological acoustics and the use of the sound spectrograph. S. W. Naturalist **9**, 118—145 • DÖRRSCHEIDT, G. J. (1973): Statistical modulation analysis of animal sounds as a basis for computer-synthesized quasi-natural stimuli in bioacoustical research. Proc. of the 2nd Seminar on Exp. Modeling and Solution of Probability Problems, Liblice (CSSR), Czechoslov. Acad. of Sci. Praha, p. 32 • DÖRRSCHEIDT, G. J. (in press): A reversible method for sound analysis adapted to bioacoustical experiments. Int. J. Bio-M Computing.

EMLEN, S. T. (1972): An experimental analysis of the parameters of bird song eliciting species recognition. Behaviour **41**, 130—171.

FALLS, J. B. (1969): Functions of territorial song in the White-throated Sparrow. In: Bird Vocalizations. (HINDE, R. A., ed.) Cambridge Univ. Press, Cambridge, 207—232 • FIELD, R. (1976): Application of a digitizer for measuring sound spectrograms. Behav. Biol. **17**, 579—583 • FLANAGAN, J. L. (1972): Speech Analysis, Synthesis, and Perception. Springer-Verlag (Orig. publ. in 1965).

GOTTLIEB, G. (1974): On the acoustic basis of species identification in Wood Ducklings (*Aix sponsa*). J. Comp. Physiol. Psychol. **87**, 1038—1048 • GREEN, S. (1975): Variation of vocal pattern with social situation in the Japanese Monkey (*Macaca fuscata*): A field study. In: Primate Behavior. (ROSENBLUM, R. A., ed.) Acad. Press, New York, **4**, 2—102 • GREENEWALT, C. H. (1968): Bird Song: Acoustics and Physiology. Smithsonian Inst. Press, Washington, D. C. (Book and two phonograph records).

HJORTH, I. (1970): A comment on graphic displays of bird sounds and analyses with a new device, the Melograph Mona. J. Theor. Bio. **26**, 1—10 • HOPKINS, C. D., M. ROSSETTO, and A. LUTJEN (1974): A continuous sound spectrum analyzer for animal sounds. Z. Tierpsychol. **34**, 313—320.

KINDLMANN, P. J., L. B. BERMAN, R. A. JOHNSON, G. D. POLLAK, O. W. HENSON, Jr., and A. NOVICK (1973): Measurement of "instantaneous" carrier frequency of bat pulses. J. Acoust. Soc. Am. **54**, 1380—1382 • KING, A. P., and M. J. WEST (1977): Species identification in the North American Cowbird: Appropriate responses to abnormal song. Science **195**, 1002—1004 • KONISHI, M. (1969): Time resolution by single auditory neurons in birds. Nature **222**, 566—569 • KROODSMA, D. (1974): Song learning, dialects, and dispersal in the Bewick's Wren. Z. Tierpsychol. **35**, 352—380.

LEMON, R. E. (1975): How birds develop song dialects. Condor **77**, 385—406 • LICKLIDER, J. C. R., and G. A. MILLER (1951): The perception of speech. In: Handbook of Experimental Psychology. (STEVENS, S. S., ed.) John Wiley, New York, 1040—1074.

MARLER, P. (1969): Tonal quality of bird sounds. In: Bird Vocalizations. (HINDE, R. A., ed.) Cambridge Univ. Press, Cambridge, 5—18 • MARLER, P., and S. PETERS (1977): Selective vocal learning in a Sparrow. Science **198**, 519—521 • MARSHALL, J. T. (1977): Audiospectrograms with pitch scale: A universal "language" for representing bird songs graphically. Auk **94**, 150—152 • MAURUS, M., and H. PRUSCHA (1973): Classification of social signals in squirrel monkeys by means of cluster analysis. Behaviour **47**, 1—2, 106—128 • MILLER, D. B. (in press): Two-voice phenomenon in birds: Further evidence. Auk • MUNDINGER, P. (1970): Vocal imitation and individual recognition of Finch calls. Science **168**, 480—482.

PETERSON, R. T. (1971): A Field Guide to Bird Songs of Central and N. America. Houghton Mifflin, New York (Orig. publ. in 1959) (Book and two phonograph records) • PUMPHREY, R. J. (1961): Sensory organs: Hearing. In: Biology and Comparative Physiology of Birds. (MARSHALL, A. J., ed.) Acad. Press, New York, **2**, 69—86.

SCARR, R. W. A. (1968): Zero crossings as a means of obtaining spectral information in speech analysis. IEEE Trans. Audio Electroacoust. AU-**16** (2), 247—255 • SCHWARTZKOPFF, J. (1963): Morphological and physiological properties of the auditory system in birds. Proc. of the XIII-th Internat. Ornithol. Congress **2**, 1059—1068 • SHIOVITZ, K. A. (1975): The process of species-specific song recognition by the Indigo Bunting, *Passerina cyanea*, and its relationship to the organization of avian acoustical behavior. Behaviour **55**, 128—175 • STEIN, R. C. (1968): Modulation in bird sounds. Auk **85**, 229—243.

TOVE, P. A., B. NORMAN, L. ISAKSSON and J. CZEKAJEWSKI (1966): Direct-recording frequency and amplitude meter for analysis of musical and other waveforms. J. Acoust. Soc. Am. **39**, 362—371.

WATKINS, W. A. (1967): The harmonic interval: Fact or artifact in spectral analysis of pulse trains? In: Marine Bioacoustics. (TAVOLGA, W. N., ed.) Pergamon Press, New York, 15—43.

Authors' addresses: J. E. R. STADDON, Department of Psychology, Duke University, Durham, North Carolina 27706, U.S.A. L. W. MCGEORGE, Department of Zoology, Duke University, Durham, North Carolina 27706, U.S.A. R. A. BRUCE, Department of Biomedical Engineering, Duke University, Durham, North Carolina 27706, U.S.A. F. F. KLEIN, Department of Anesthesiology, Duke Medical Center, Durham, North Carolina 27710, U.S.A.