FEATURE WEIGHTING IN SPECIES SONG RECOGNITION BY THE FIELD SPARROW (SPIZELLA PUSILLA)

by

DOUGLAS A. NELSON[1])
(The Rockefeller University, Field Research Center, Box 38B, RR 2,
Millbrook, NY 12545, U.S.A.)

(With 9 Figures)
(Acc. 22-XII-1987)

Introduction

Experimentation on species song recognition by birds has been active over the past three decades (reviewed in BECKER, 1982). Much of the research has focussed on determining the importance of different acoustic features in the recognition process. Two hypotheses have been discussed most frequently in the literature.

One hypothesis predicts that song features that vary little within a species ('invariant' features) will be the most reliable song recognition cues (MARLER, 1960; FALLS, 1963; EMLEN, 1972; BECKER, 1982). EMLEN (1972) and BECKER (1982) in their reviews of the literature concluded that only a limited subset of song features, the invariant ones, are used by birds in song recognition; but they stressed that not all invariant features within a species' song are used as recognition cues. However, a recent study employing parametric manipulations of song features demonstrated that relatively variable features are also used by birds in recognizing their own species' song (DABELSTEEN & PEDERSEN, 1985).

The second hypothesis predicts that features that are distinctive relative to other species in the local environment will be preferred recognition cues (MARLER, 1960; FALLS, 1963; EMLEN, 1972; DABELSTEEN & PEDERSEN, 1985). This 'sound-environment' hypothesis

---

[1]) This research was supported by a fellowship from the U.S. National Institutes of Health (No. 5 F32 NS07629) and Biomedical Research Support Grant BRSG S07 RR07065 to Rockefeller University. I thank Vassar College for access to computing facilities, and the Institute of Ecosystem Studies at the New York Botanical Garden Mary Flagler Cary Arboretum for permission to work on their property. I thank T. DABELSTEEN, C. EVANS, P. MARLER, and an anonymous reviewer for helpful comments on the manuscript, and Jörg BÖHNER for preparing the German summary.

assumes that variability is only one component of species distinctiveness; the other component is the relationship between different species in an 'acoustic space' defined by features that represent song structure (BREMOND, 1976; MILLER, 1982; SORJONEN, 1986). Thus, relatively variable features could be useful song recognition cues if a species' song differed sufficiently from other species.

The goal of this research was to compare the relative importance of different sound features in species song recognition by the field sparrow (*Spizella pusilla*). Experiments by GOLDMAN (1973) and MORRISON-PARKER (1977) demonstrated that field sparrows discriminate their songs from the songs of two species that frequently occur in the same habitat: the chipping sparrow (*Spizella passerina*) and prairie warbler (*Dendroica discolor*). Field sparrow songs usually consist of about 16 notes (range = 7 to 28) grouped into two or more 'phrases', with the durations of notes and intervals between notes decreasing in successive phrases. A note is defined as a continuous tracing on a spectrogram, uninterrupted by more than 3 ms of silence. Most songs end with a steady trill of about 12 notes. Spectrograms illustrating variation in field sparrow song are shown in Fig. 1.

A discriminant analysis of 14 features of field sparrow song and the songs of 12 other species that sing in or near old fields in the eastern United States (including chipping sparrows and prairie warblers) revealed that invariant frequency features were the most effective in discriminating field sparrow songs from the songs of other local species. Other features such as the number of phrases within a song, inter-note interval, and note duration were equally effective in statistically distinguishing songs (NELSON, unpubl. data). Therefore, the sound-environment hypothesis predicts that field sparrows should be more sensitive to experimental manipulations of frequency than to changes in any other feature. The invariant features hypothesis makes a similar prediction, but differs in predicting that the number of phrases in a song, a relatively invariant feature, should be a more important feature than the significantly more variable features note duration and inter-note interval.

The general experimental strategy was to produce four series of songs in each of which one acoustic feature (maximum frequency, number of phrases, trill note duration, and inter-note interval was systematically varied relative to a fixed control while other features were kept constant within the normal range of variation in field sparrow song. On the basis of earlier studies (*e.g.*, SCHUBERT, 1971) I expected that small changes in

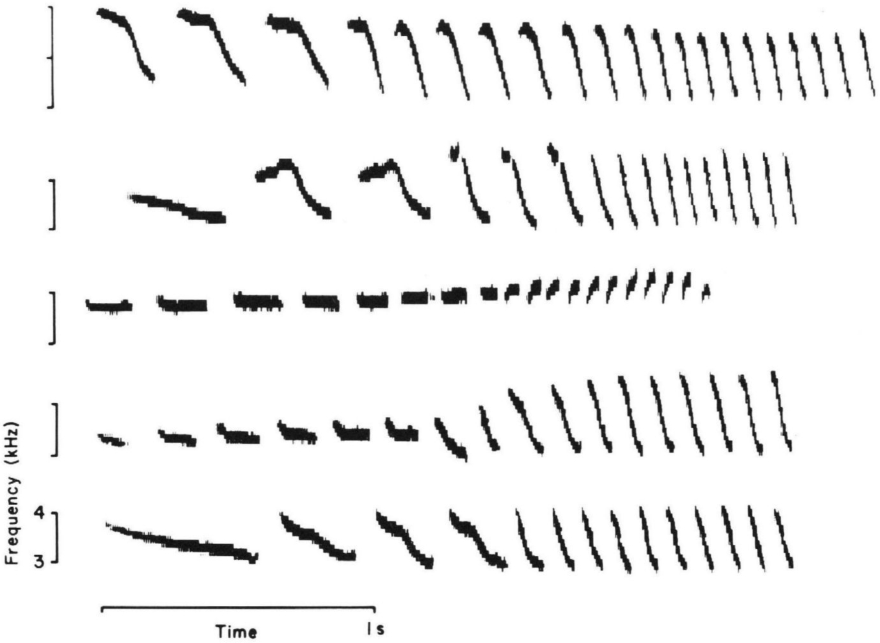

Fig. 1. Spectrograms representing variation in natural field sparrow song. A 256 point DFT (temporal resolution = 12 ms, frequency resolution = 81 Hz) was used to produce the plots on a computer graphics printer.

the value of a feature relative to the control would have little or no effect on the salience of the experimental song, but that at some point, further alterations to a feature would produce a significant decrease in responsiveness. Initially, large step sizes were used to identify the approximate limits of variation the birds would respond to. I then used smaller steps within each series, limited in three series by the temporal or frequency resolution of the computer synthesis techniques. These experimental series served to identify the minimal acoustic difference along each feature continuum that resulted in a significant response decrement relative to the control, the 'just-meaningful difference'. The 'just-meaningful difference' describes how much a feature must be changed to significantly reduce the birds' natural territorial response to song playback. It does not bear any necessary relation to the just-noticeable difference (JND) measured in laboratory psychophysical tasks. An additional experimental series tested the importance of various aspects of a complex feature, the 'shape' of a note, but did not identify the just-meaningful difference in note 'shape'.

After having scaled the four acoustic features in this manner, I then did six experiments (all possible pair-wise comparisons among four stimuli) to determine the relative importance of each feature in song recognition. Birds were presented with a choice of two songs that differed in two features (A & B). Each feature had two states (A & a, B & b), representing the smallest acoustic difference that resulted in a significant response decrement. One song had the 'species typical' (control) state for A and the 'just-meaningfully different' state for b, while the alternative had the opposite arrangement (aB). The song that elicited the strongest response was inferred to contain the most important species-typical feature-state.

Methods

Stimuli.

Stimuli were synthesized on a computer-based system (ZOLOTH *et al.*, 1980; DOOLING *et al.*, 1982). Stimuli were digitized by a 12 bit analog-to-digital converter at a sampling rate of 20.8 kHz, and were converted to a spectral representation by a 128 point digital Fourier transform (DFT). This yielded frequency resolution of 163 Hz and temporal resolution of 6 ms in manipulating signals using an editing program (EDFFT). A brickwall low-pass filter set at 40% of the sampling frequency was used to prevent aliasing during all signal input and output operations. To produce two stimuli (in Experiments 7 & 18), a sampling rate of 41.6 kHz was used to achieve better temporal (3 ms) resolution during the editing process. The experimental stimuli used in the frequency series were produced from a 256 point DFT (resolution = 81 Hz). In 1985 stimulus parameters were analyzed on a Kay digital sonagraph; in 1986 and 1987, a digital analysis program (Engineering Design, 1987) was used. Temporal resolution was 3 and 6 ms, and frequency resolution 40 and 81 Hz, respectively, using the two methods.

Two experimental series involved manipulations of a single note 'prototype', selected because it was a high quality recording of a constant FM downsweep. The 'shape' of this note (defined by its duration, bandwidth, and frequency modulation rate) was altered by sending it through the computer's digital-to-analog converter to a Nagra III tape recorder at a sampling rate different from the 20.8 kHz originally used to enter the note into the computer. Thus, a faster output rate produced a shorter note with wider bandwidth, higher frequency, and faster FM rate; while a slower output sampling rate 'stretched' the note. The appropriately 'stretched' or 'shortened' note was re-entered into the computer at the usual sampling rate of 20.8 kHz. The modified note was then edited to the desired duration, bandwidth and frequency using the filter and frequency shift functions in EDFFT. Once a test note with the desired parameters was produced, it was copied in EDFFT to produce a trill. Fig. 2 depicts the prototype note and the four notes derived from it that formed the note duration series. Notice that the 84 ms-long note that served as the control trill note in most experiments went through the same sequence of editing steps used to construct experimental notes. Design specifications of stimuli are presented in Table 1.

Other manipulations of test stimuli were straightforward. Inter-note interval was controlled by adding or deleting 6-ms-long time blocks (3 ms in Experiment 18) with zero amplitude between notes. Frequency was altered by addition of integral multiples of 81 Hz to the control song. Phrase structure was varied by 'splicing' in EDFFT an introductory phrase edited from a natural song to a terminal trill of notes, thereby producing a

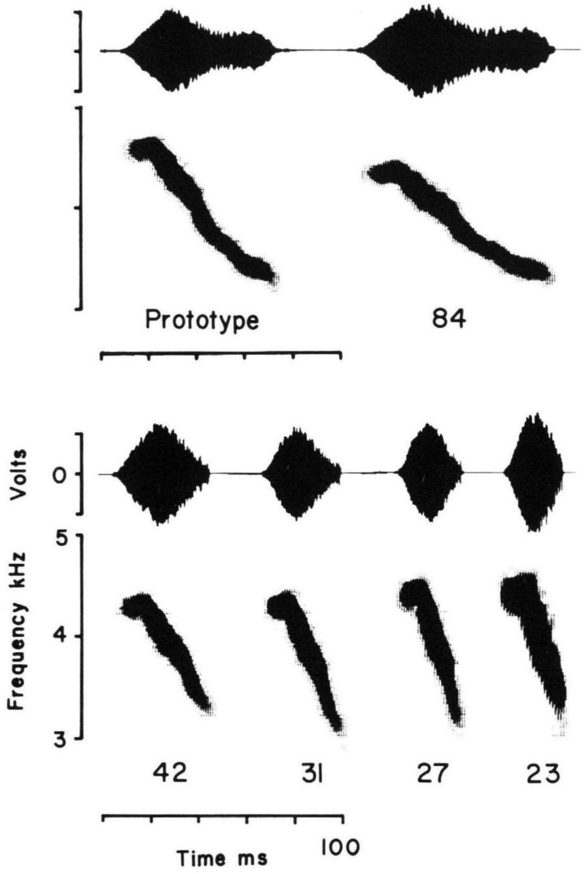

Fig. 2. Digitally-generated amplitude-time (top) and frequency-time (bottom) plots of the prototype note and four notes derived from it used in the note duration continuum (Experiments 5-8). The spectrograms were derived from a 256 point DFT.

two-phrased song of approximately 2 s duration, or by making a 2-s-long trill without an introduction. Two different introductory phrases were used to prevent exposing birds to an identical stimulus when tested two or more times in one year. The peak amplitudes of the last note in both introductory phrases were identical, and matched that of the trill notes, as occurs in natural songs.

Experimental procedure.

Twenty-five experiments were performed from 1985 to 1987 using a two-speaker choice paradigm (Lanyon, 1978; Peters *et al.*, 1980). In Experiments 1 to 19, one song was a control, and the alternative stimulus song differed in one feature. Experiments 20 to 25 were the feature-weighting tests. Two NAGRA DH loudspeakers were set on the ground facing up 16 m apart near the middle of a subject's territory. Insofar as possible, speakers

TABLE 1. Design parameters of experimental notes

Experiment	Stimulus	Maximum frequency (Hz)	Note duration (ms)	Bandwidth (Hz)	FMR[1] (Hz/ms)	Inter-note interval (ms)
1-8 (Control)	84 ms	4400	84	1150	-14	47
5	23 ms	4400	23	1000	-43	47
6	27 ms	4400	27	1100	-41	47
7	31 ms	4350	31	1150	-34	46
8, 10	42 ms	4350	42	1180	-28	50
9	3 kHz	4300	87	3070	-35	50
10	2.2 kHz	4300	43	2200	-51	55
11	downsweep	4400[2]	108	1000	-10	50
11	"tone"	3700	108	80	-.8	50
12	3 elements	4400	107	1500	-22, +150, -9	50
12	1 element	4460	107	960	9	50

[1] FMR = frequency modulation rate: + if increasing frequency, - if decreasing.

[2] The mid-frequency of the downsweep (3900 Hz) was 200 Hz higher than the "tone" it was paired with in Experiment 11.

were positioned such that trees and bushes were uniformly distributed near them. To facilitate measuring approach distances, poles were set a 4 m intervals between speakers. In 1986 and 1987 each speaker was covered with a piece of brown, acoustically-transparent cloth. I kneeled or stood 15-20 m away and recorded observations on a cassette recorder. Observations were usually transcribed later the same day.

The two stimuli being presented were recorded at equal amplitudes on separate tracks of a SONY TC-D5M stereo tape recorder, so that they alternated, with 10 s (in 1985) or 13 s (in 1986 and 1987) inter-song intervals. A trial in 1985 consisted of two 90 s playback blocks separated by 120 s of silence. In 1986 and 1987 each playback block was 120 s long. Thus, in 1985 each block contained 10 tokens of each song, while eight tokens of each were used in 1986 and 1987. Before each trial, two coin flips determined which speaker played which song, and which song was played first in the first block. In the second block, the speaker cables and order of song presentation were reversed. The distances between the bird and each speaker were noted every five s during the playback blocks using six distance categories: 0-2 m (mean = 1 m), 2-4 (3), 4-8 (6), 8-14 (12), 14-18 (16), and 18+ (24). Mean distances from each stimulus were calculated over the total time of playback. The difference between the mean distances for each song were compared using the Wilcoxon paired-signed-ranks test. One-tailed probabilities of 0.05 were used in Experiments 1 to 19 in which one acoustic parameter was varied relative to a control; two-tailed probabilities ($P<0.05$) were applied in Experiments 20 to 25. After testing eight individuals, I analyzed the data and if the null hypothesis was rejected, stopped the experiment. Otherwise, four more males were tested, yielding an upper limit of 12 birds per experiment. The Wilcoxon test is conservative with a sample size of eight. I adopted this procedure to conserve time and the number of birds needed. Sample sizes in some experiments depart from 8 or 12 because of tied observations.

Approach distance was the only suitable response measure in the two-speaker design. Most males did not sing during song playback, but even if they did, it was not possible to decide which speaker the male was responding to. Call notes (chips, trills) were infrequently given; and field sparrows did not exhibit any obvious visual displays or postures (other than flights and hops) during playback.

Male field sparrows were tested from 14 May to 16 July 1985, 21 April to 1 August 1986, and 23 April to 28 June 1987. Thirty-four males were tested in 1985 (25 birds tested once, 9 twice), 80 males in 1986 (60 once, 20 twice) and 104 males in 1987 (67 once, 37 twice). I avoided experimenting within a two-territory radius of a completed trial area for two days, and did not retest individuals within three weeks. No male was used more than once in the same experiment. Trials were run from dawn until 10:30 on days of little wind and no rain.

Variation in natural song.

I recorded songs from individuals on 98 different territories in 1986 using a Nagra III tape recorder (19 cm/s) and Sennheiser MKH 816 shotgun microphone. Songs were entered into a PDP 11/23 computer at a sampling rate of 20.8 kHz for analysis by a software program (Signal, Engineering Design, 1987). Songs were displayed on a computer terminal, and measurements were recorded using a graphics cursor. Temporal measurements were made on spectrographic plots produced with a 128 point DFT (resolution = 6 ms); while frequency measures were derived from a 512 point DFT (resolution = 41 Hz). Measurements made were: number of phrases in a song (only complete songs were used), maximum frequency in the song, duration of notes in the terminal trill phrase, and duration of inter-note intervals in introductory and terminal trills. All variables except the number of phrases had log-normal distributions. In reporting results below, variable means and values one, two and three standard deviations above or below the mean were backtransformed from logs into the original linear scale. Standard deviations for these variables are not given in the text, but are presented in

Figures 4, 6, and 7 for comparison to the birds' responsiveness to experimental manipulation of each feature. The coefficient of variation (c.v.) was calculated as (standard deviation/mean) × 100.

Results

Experiments on phrase structure.

Natural field sparrow songs have a mean of 2.8 phrases (range = 2-5, c.v. = 19%, N = 98). Compared to notes in the terminal trill, notes in the introductory phrase(s) are longer and are separated by longer inter-note intervals (NELSON, unpubl. data). Birds occasionally sang incomplete songs that, with one exception, always consisted of the long introductory notes without the terminal trill.

Experiment 1 compared a two-phrased song consisting of an introduction taken from a natural song spliced before a trill of eight 84-ms-long notes, to a song containing a single trill of 16 84-ms-long notes. Males approached significantly closer to the two-phrased song that more closely resembled normal field sparrow song (Fig. 3).

Experiment 2 employed a song with a synthetic introduction made of two 336-ms-long notes spliced to a trill of eight 84-ms-long notes. The two introductory notes were made by 'stretching' and duplicating an 84-ms-long note. Inter-note intervals were 47 ms in both introduction and trill. Males approached significantly closer to the two-parted song than to the single trill (Fig. 3).

Experiment 3 tested whether the normal temporal arrangement of phrases within a song was important in recognition. Males approached significantly closer to the normal syntax song than to a song with the phrase order reserved (Fig. 3).

The results of Experiments 1 and 2 could have resulted from the males' preference for two-parted songs, a preference for a shorter trill, or from a preference for the longer notes that occur in the introduction. To exclude the latter possibility, Experiment 4 compared an introduction alone to a trill. Both of the introductions used in different experiments were tested, matched by length to the trill each was paired with. Of 21 males tested, only nine approached within 24 m of either speaker. The nine birds did not approach significantly closer to either introduction or trill (Fig. 3). Of the 12 non-respondents, six approached playback of a normal song immediately after the experimental trial, and six did not. The low response rates were probably the result of two factors: the stimuli were brief (about 1 s), and the experiment was performed late in the breeding season (22-30 July 1986). To summarize this series of experiments, males significantly preferred control songs with a normal

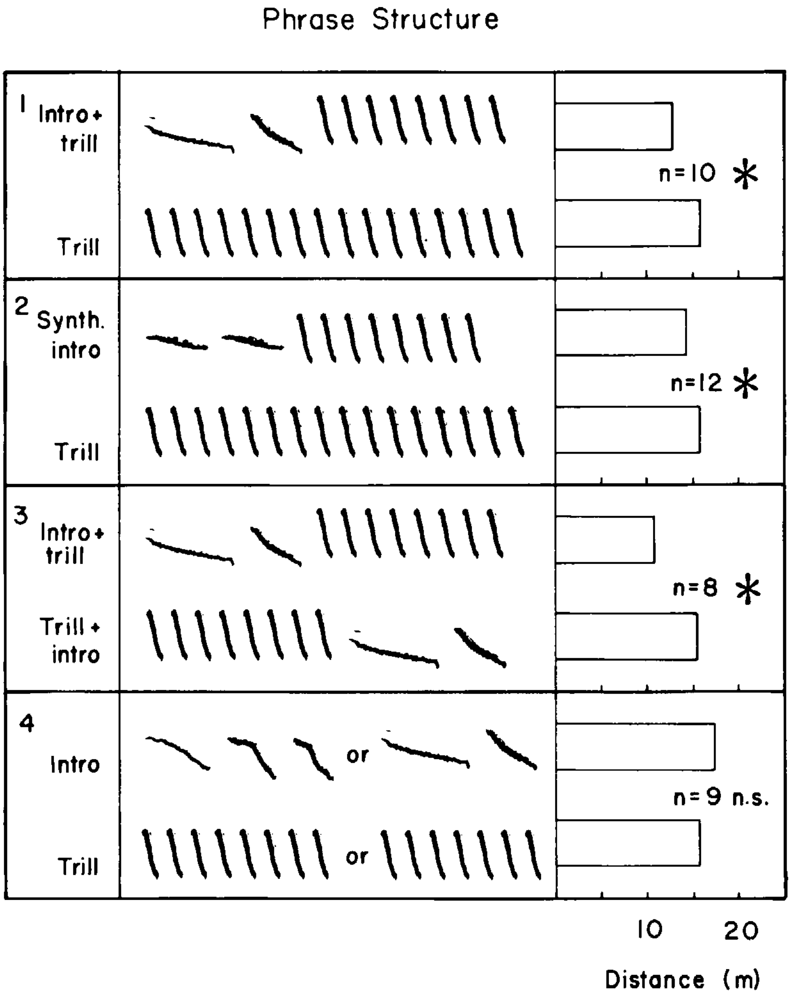

Fig. 3. Stimuli used in Experiments 1-4 to test the importance of phrase structure. Each stimulus except those in Experiment 4, is approximately 2 s long. Spectrograms were made using a 40 Hz analyzing bandwidth on a Kay Digital Sonagraph. Frequency is scaled on the ordinate in 1 kHz intervals; the longer tick is 4 kHz. Histograms on the right depict the mean approach distance made to each song. Note that a smaller distance indicates a stronger response. Standard errors of the mean were less than 2 m in all experiments, and are omitted. Asterisks denote significantly different distances (* = one-tailed $P<0.05$, ** = $P<0.025$, *** = $P<0.01$) using a Wilcoxon-matched-pairs signed-ranks test.

number of phrases (two) over a song with one phrase, and a song with phrases in the reversed order.

Experiments on trill-note duration.

Notes in the terminal trill phrase of natural field sparrow songs have a mean duration of 63.2 ms (c.v. = 38%, Fig. 4, upper panel). The synthetic note used as the control note in most experiments was longer (84 ms) than the species mean but had an FM rate and bandwidth nearly equal to the species mean. The difference in duration arose because when I began the experiments, I had a smaller sample of recordings of natural song (N = 22) with a mean of 97 ms. The notes tested against the control were 42, 31, 27, and 23 ms long. Since the inter-note intervals in the trill were held constant at 47 ± 2 ms, the series represented decreasing note duration and a faster temporal pattern.

Males approached significantly closer to a song composed of a single trill of 16 84-ms-long notes than to a single trill of 28 23-ms-long notes (Fig. 4: Experiment 5). Different numbers of notes were used in the two songs to control for total song duration. Experiments 6 to 8 below all employed two-parted songs in which the duration of notes comprising the terminal trill was varied relative to the 84 ms control. The number of trilled notes in the experimental songs varied from 8 to 16, because I decided to control for total song duration. This variation in number of

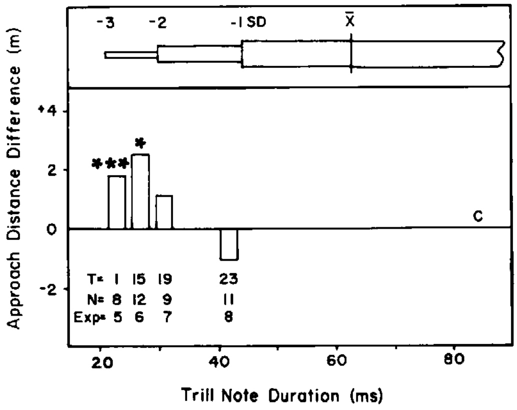

Fig. 4. Results of Experiments 5-8 in the trill note duration series. The top panel shows distributional statistics based on a sample of 98 natural songs. The bottom panel shows experimental results. Histograms represent the difference in approach distances between each experimental song and the control (C). Positive differences resulted if subjects were closer to the control song, and negative differences if closer to the experimental song. T. is the Wilcoxon statistic, N the sample size, and asterisks denote statistical significance as in Fig. 3.

trill notes was within the normal range of variation (x = 10.3 ± 4.4, N = 77).

A song with a terminal trill comprised of 27-ms-long notes elicited a significantly weaker response than did the control, while 31- and 42-ms-long notes elicited the same response as the control (Fig. 4: Experiments 6-8). Thus, terminal-trill notes with durations within approximately two standard deviations of the mean (31, 42, 84 ms) elicited similar responses, while note durations below this range (23, 27 ms) produced significantly weaker responses.

Experiments on trill-note shape.

In simple frequency sweeps, bandwidth and FM rate covary with note duration in determining a note's 'shape'. Bandwidth is the difference between maximum and minimum frequencies within a note, and FM rate is the bandwidth divided by duration. In Experiments 5-8 bandwidth was held approximately constant at 1100-1200 Hz while duration and FM rate covaried (Table 1). Thus, the response differences observed in Experiments 5 and 6 could have resulted from differences in note duration, FM rate, or both. The next series of experiments (Numbers 9-11) tested the relative importance of note duration, bandwidth, and FM rate in species song recognition.

In Experiment 9 the control song was the trill of 84-ms-long notes used in Experiment 5 (bandwidth = 1150 Hz, FM rate = 14 Hz/ms, close to species' mean in duration, bandwidth [1200 Hz], and FM rate [18 Hz/ms]). The second song was made of notes with bandwidth (3080 Hz) outside the normal range of variation, and duration (87 ms) and FM rate (35 Hz/ms) within the normal range. In natural trill notes a bandwidth of 1940 Hz and FM rate of 36 Hz/ms were two standard deviations greater than their corresponding means, and define the upper limits to the natural ranges of variation. Both experimental trills had the same maximum frequency. Males made weak, statistically indistinguishable responses to these songs (Fig. 5: Experiment 9).

Experiment 10 compared terminal-trill notes of similar duration (42 ms and 43 ms), but differing in bandwidth and FM rate. One note, identical to that used in Experiment 8, had a bandwidth of 1180 Hz and a FM rate of 28 Hz/ms. The other note had a bandwidth of 2200 Hz and FM rate of 51 Hz/ms. Despite the latter note's bandwidth and FM rate being outside the normal range of variation in terminal trill notes in field sparrow song, males did not 'prefer' one note over the other (Fig. 5:

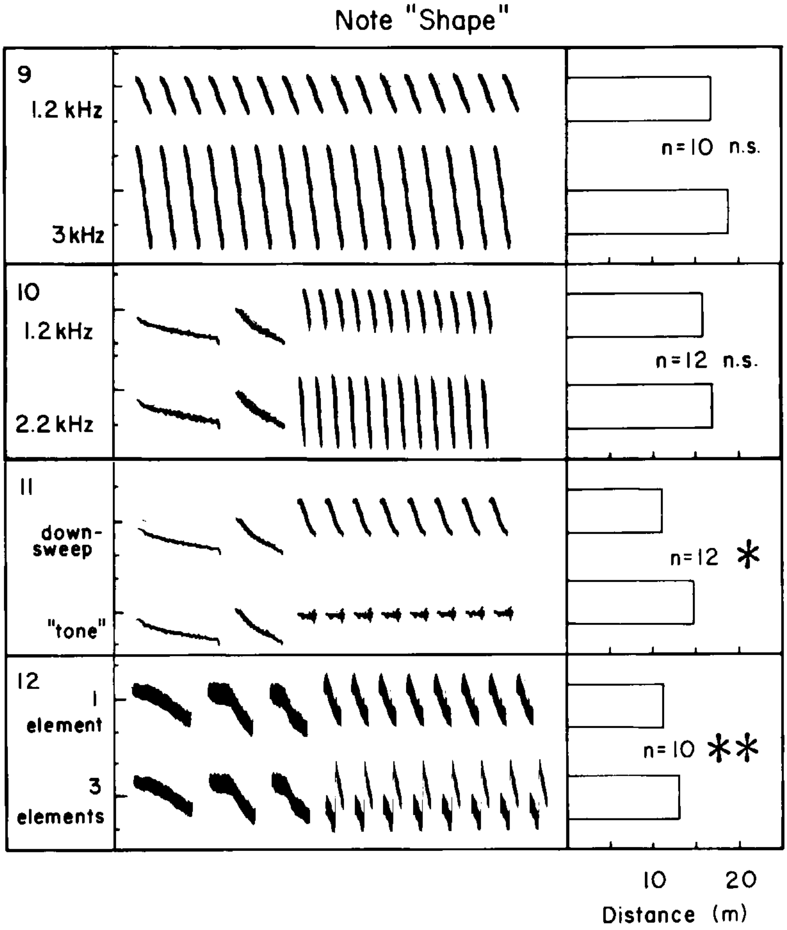

Fig. 5. Stimuli used in Experiments 9-12 to test the importance of note 'shape'. Results shown as in Fig. 3. Audiospectrograms of stimuli used in Experiment 11 were prepared with a 300 Hz analyzing bandwidth, in order to show the brief middle element in the 3 element note.

Experiment 10). The results of Experiments 9 and 10 suggest that large bandwidths and high FM rates do not decrease responsiveness, and further suggest that differences in note duration, or associated changes in temporal pattern, were controlling responses in Experiments 5 and 6.

An element is a note sub-unit with nearly constant rate of FM, separated from other such units by silence or a 'marked' change in FM. In normal song, most elements (92% of 125) in terminal trill notes

(N = 98, 27 of these had two elements) are frequency downsweeps. In New York field sparrow songs, elements with upsweeps or nearly-constant tones were rare in terminal trills but common in introductory or middle phrases. To test if note bandwidth was used in song recognition, I synthesized a 107-ms-long note by removing 32 ms from the last third of a natural tonal note sampled from the middle phrase of a song. The electronic 'splice' left a 1-ms-long notch of 1.1 dB (relative to peak amplitude) in the waveform, resulting in a brief period of spectral spreading (Fig. 5: Experiment 11). A trill of eight of these notes preceded by an introduction was tested against a control song composed of the introduction followed by eight 107-ms-long frequency-downswept notes. Twelve males approached significantly closer to the control song (Fig. 5: Experiment 11), suggesting that males 'prefer' the normal FM structure of terminal trills. A possible confound in this experiment, however, was the greater loudness relative to the introduction of the tonal-note trill (apparent to me) as compared to the downswept-note trill.

All notes in terminal trills of natural songs are composed of one (72%) or two (28%) elements. Attempts to synthesize a note with three elements with constant durations and FM rates from a natural note were unsuccessful, so I selected a note composed of three elements from the introduction of a natural field sparrow song. The first and third elements were within the normal range of element duration (43, 54 ms) and FM rate (22, 9 Hz/ms), but the middle element was short (10 ms) with a fast FM rate (150 Hz/ms). This was paired against a single-element control note with the same total duration (107/ms) as the three-element note (Fig. 5: Experiment 12). Both notes had the same maximum frequency, but the three-element note had a wider bandwidth that was well within the normal range of variation (Table 1). Males approached significantly closer to the song with one-element notes. Further experiments would be needed to eliminate the possibility that spectral differences or the extremely short middle element, and not the number of elements per se, was responsible for the response decrement (cf. Experiments 5 & 6).

Experiments on song frequency.

Experiments 13 to 15 tested the effect of increasing the frequency of the entire song relative to a control song with maximum frequency of 4400 Hz. The mean maximum frequency of 98 natural songs is 4365 Hz (c.v. = 6%, Fig. 6, upper panel). In Experiment 13 the entire experimental song was shifted 652 Hz higher to 5052 Hz. Males approached the same distance to both songs (Fig. 6: Experiment 13). In Experiment 14 the frequency difference between control and experimental songs was increased to 815 Hz. Males approached significantly closer to the control (Fig. 6, Experiment 14). Males also approached closer to the control song when it was paired against a song shifted 978 Hz higher to 5378 Hz (Fig. 6: Experiment 15). I attempted to produce a song with a maximum frequency between 5052 and 5215 Hz. However, the song shifted to a maximum frequency of 5133 Hz had deep amplitude modulation and an audible buzzy quality in the song that differed from normal field sparrow song. I chose not to test a song with both AM and frequency altered. Thus, frequency increases greater than two standard deviations relative to the mean resulted in significantly weaker approach responses.

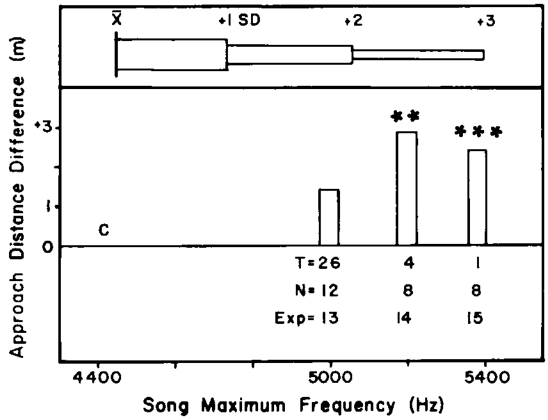

Fig. 6. Results of Experiments 13-15 in the song frequency series, presented as in Fig. 4.

Experiments on inter-note interval.

The mean inter-note interval in natural song is 79.3 ms in the introductory phrases and 37.1 ms in the terminal trill (c.v. = 26%, Fig. 7: upper panel). I used 87 ms and 47 ms (range = 46-48 ms), respectively, as the inter-note intervals in control songs. When inter-note intervals were increased in three experimental steps of 108, 120, and 132 ms in the introduction, and 65, 72, and 76 ms in the terminal trill, approach distances to the experimental songs did not differ significantly from the control (Fig. 7; Experiments 16-18). When inter-note intervals were increased to 144 ms in the introduction and 84 ms in the trill, the dif-

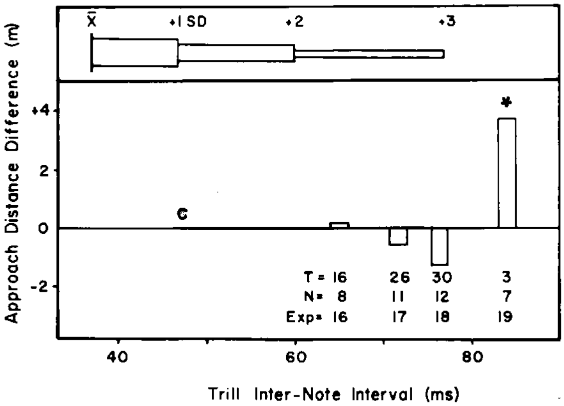

Fig. 7. Results of Experiments 16-19 in the inter-note interval series, presented as in Fig. 4. Inter-note intervals of the introductory phrase were also varied relative to the control (87 ms). In Experiment 16: 108 ms; 17: 120 ms; 18: 132 ms; and 19: 144 ms. The mean introductory inter-note interval is 79.3 ms (+1 SD = 106 ms, +2 SD = 140 ms).

ference in approach distances increased significantly to the control song (Fig. 7; Experiment 19). The number of trill notes was held constant in this series. The songs varied in duration from 2.2 to 2.6 s, well within the normal range of variation in natural songs (x = 2.5 ± 0.35 s, N = 98). To summarize, responses decreased significantly when inter-note intervals were increased by two to three standard deviations relative to the species' mean, resulting in a slower temporal pattern.

The results thus far have shown that manipulations to each of five acoustic features, phrase structure, note duration, note shape, song frequency, and inter-note interval, lead to significant decrements in response. By themselves, these data do not demonstrate that male field sparrows differentially 'weight' certain features over others in recognizing songs, since all features were shown to affect responses. One exception to this conclusion is that extreme note bandwidths or FM rates did not decrease responses as long as the note durations were within the normal range of variation (Fig. 5).

Relative importance of features.

The final six Experiments consisted of all possible pair-wise comparisons among four features: number of phrases, trill-note duration, song frequency, and inter-note interval. In each Experiment one song had the control or 'right' value for one feature (A), and the just-meaningfully different value, as determined in Experiments 1-19, for the second feature (b). The second song had the feature-states reversed (aB).

Experiment 20 compared the relative importance of trill-note duration and song frequency in recognition. One song had 27-ms-long trill notes at a peak frequency of 4400 Hz ('wrong' duration, 'right' frequency), while the other had 84-ms-long notes at 5215 Hz ('right' duration, 'wrong' frequency). Males approached significantly closer to the song with the 'wrong' trill-note duration but 'right' frequency, suggesting that song frequency is a more important cue than trill-note duration (Fig. 8: Experiment 20).

Experiment 21 compared phrase structure to song frequency. Males approached closer to a one-part song at 4400 Hz ('wrong' phrases, 'right' frequency) than to a two-parted song with peak frequency of 5215 Hz ('right' phrases, 'wrong' frequency) (Fig. 8: Experiment 21). This suggests that song frequency is more important than the number of phrases in the song in recognition.

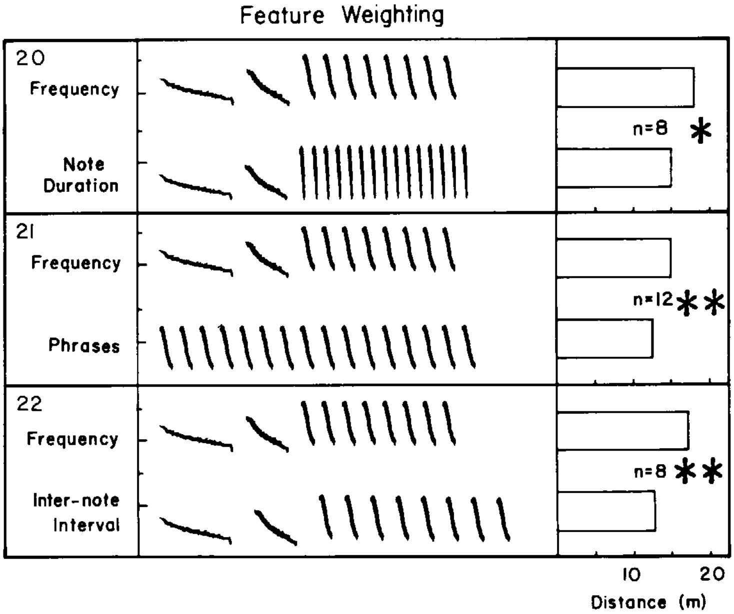

Fig. 8. Results of Experiments 20-22 comparing the importance of frequency to three other features in species song recognition. Results are displayed as in Fig. 3. The stimulus descriptions to the left describe which feature of the two compared in each Experiment was changed to the state representing the 'just meaningful difference' determined from Experiments 1-19. See text for further explanation. Songs labelled 'Frequency' at left are 2 s long. Asterisks denote two-tailed probabilities: * = $P<0.05$, ** = $P<0.025$, *** = $P<0.01$.

The importance of song frequency and inter-note interval were compared in Experiment 22. Males approached significantly closer to the song with the 'wrong' inter-note interval and 'right' song frequency than to the song with 'right' inter-note interval and 'wrong' song frequency (Fig. 8; Experiment 22), again suggesting that song frequency is the more important cue.

Experiment 23 compared phrase structure to trill-note duration. Four males approached closer to the song with the 'wrong' number of phrases and 'right' trill-note duration, and four approached closer to the alternative song; suggesting that phrase structure and trill-note duration are of equal importance in species song recognition (Fig. 9; Experiment 23). (The Wilcoxon statistic obtained with eight males exceeded the critical value for N = 12, so four additional males were not tested.)

Experiment 24 compared phrase structure to inter-note interval. The song with the 'wrong' phrase structure and 'right' inter-note interval was a single trill with 47-ms-long inter-note intervals. The other song had two phrases but with inter-note intervals increased in the introductory phrase to 123 ms and 84 ms in the terminal trill. Males approached significantly closer to the two-parted song, leading to the conclusion that extreme

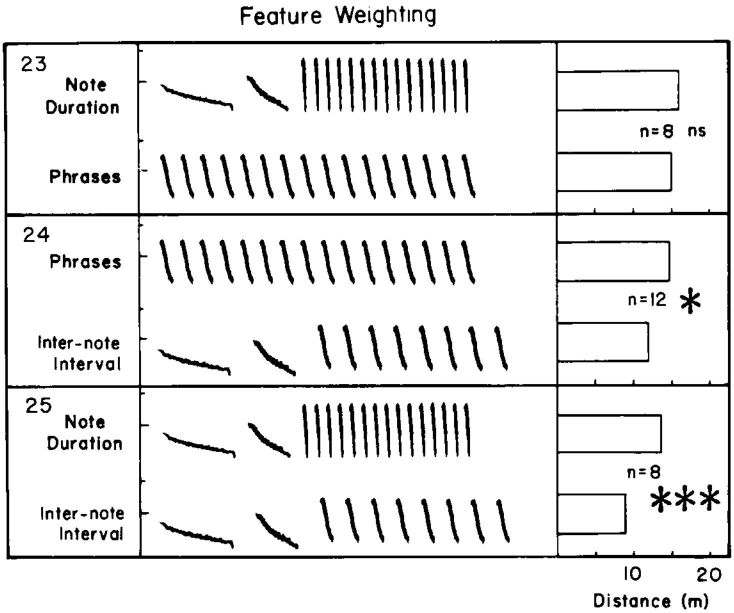

Fig. 9. Results of Experiments 23-25 comparing the relative importance of song features in species song recognition. Results presented as in Fig. 8.

inter-note intervals disrupt song recognition less than does extreme phrase structure (Fig. 9: Experiment 24).

The trill-note duration and inter-note interval features were compared in Experiment 25. Seven of eight males approached closest to the song with the 'wrong' inter-note interval and 'right' trill-note duration (Fig. 9; Experiment 25), suggesting that alterations to inter-note interval are tolerated more than alterations to trill-note duration.

Discussion

The major findings of this research are: 1) Field sparrows are sensitive to changes in each of the five features tested, 2) Approach responses decreased significantly when four features (number of phrases, trill-note duration, frequency, and inter-note interval) were varied by approximately two to three standard deviations relative to the control, and 3) Multiple features are integrated in species song recognition, and features are differentially weighted. An invariant feature, song frequency, is weighted more heavily than another invariant feature, number of phrases, and more variable features (trill-note duration, inter-note interval).

It is clear from these results that field sparrows derive information about species identity from multiple features. The observation that responses decreased significantly when a feature was changed by more than two standard deviations relative to the mean (control) value, appears to be similar to the situation described in the European blackbird, where responses declined when two of three features were changed to just outside the limits of natural variation in song (DABELSTEEN, 1984; DABELSTEEN & PEDERSEN, 1985). DABELSTEEN & PEDERSEN (1985) used a measure to represent the limits of natural variation in blackbird song different from, and not obviously related to, the measure used here. Thus, it is difficult to compare these results to DABELSTEEN & PEDERSEN's. Standard deviations calculated from a large sample of natural conspecific songs were used here to represent the variability in song because in a normally-distributed sample, 96% and 99% of the observations fall within two and three standard deviations, respectively, of the mean. Thus, field sparrows and blackbirds appear to be 'tuned' to the normal range of variation they encounter in conspecific song. Parametric manipulations of song features in other studies have not been compared to the range of variation present in song (BREMOND, 1968; M. SCHUBERT, 1971; HELB, 1973; BECKER *et al.*, 1980).

In field sparrows, manipulations of certain features by more than three standard deviations did not influence responses. Increasing FM rate and bandwidth by 3.5 and 7 standard deviations to outside the normal range of variation did not decrease responses (Experiments 9 & 10). However, decreasing the bandwidth to yield a tone (Fig. 5: Experiment 11) did result in a significant response decrement. Thus the effects of some manipulations may depend upon the direction in which the feature is changed relative to the control.

There are at least three possible explanations for asymmetric effects of experimental manipulations. Response sensitivity may be heightened by changes in one direction because such changes bring the song into 'acoustic overlap' with (1) other species (BREMOND, 1978), or (2) other conspecific vocalizations (GAIONI & EVANS, 1986). The first possibility seems unlikely in the present case because an acoustic analysis of variation in field sparrow song and songs of 12 other species did not reveal any obvious asymmetries in the distribution of alien songs relative to field sparrow song (NELSON, unpubl. data). The second possibility also seems unlikely because field sparrow call notes are so much simpler than song. A third possible explanation is that constant ratios between the two stimuli were not preserved when FM rates were increased and decreased. A larger relative difference between stimuli should be more noticeable to the birds. The ratio between FM rates of the ''tone'' note and frequency-downsweep in Experiment 11 was much greater (10/0.8 = 12.5) than the ratios between the FM rates contrasted in Experiments 9 (2.5) and 10 (1.8). BECKER (1976) and DABELSTEEN & PEDERSEN (1985) both found that frequency decreases had greater effects on response levels that did physically equivalent frequency increases, presumably because decreases represented larger relative changes than increases.

These results are consistent with the 'feature variation tolerance' model described by SHIOVITZ & LEMON (1980) and similar models developed to describe call perception in crickets (DOHERTY & HOY, 1985), songbirds (DABELSTEEN & PEDERSEN, 1985), and ducklings (GAIONI & EVANS, 1986). That is, field sparrows appear to compare song stimuli to an internal representation of song, perhaps in the form of an auditory template (KONISHI, 1965; MARLER, 1976, 1977; MORTON, 1982; MCGREGOR *et al.*, 1983; MARGOLIASH, 1983, 1986). Response strength is then dependent on the disparity between the stimulus and representation along five or more features. Response strength as assayed here appears to be approximately constant provided the stimulus disparity is within two or three standard deviations of the mean, but responses then decrease as the acoustic difference increases. This is not to say that field sparrows cannot discriminate song variants within the normal species-typical range of variation. Other experimental paradigms have shown that males discriminate the songs of familiar and unfamiliar field sparrows (GOLDMAN, 1973; MORRISON-PARKER, 1977; NELSON, unpubl. data).

EMLEN's (1972) and BECKER's (1982, p. 223) generalization that only a subset of acoustic features, invariant ones, are used in song recognition does not accurately describe these results. Field sparrows use both invariant and variable features, a conclusion also reached by DABELSTEEN & PEDERSEN (1985) for the European blackbird. The reason that previous workers concluded that only invariant features are used as song recognition cues may be as follows. If only a limited number of manipulations are performed without reference to the natural range of variation of features, then changes made to a variable feature are more likely to fall within the 'tolerated' range than are random changes made to a relatively invariant feature. Thus, limited changes made to a variable feature would not influence the birds' responses, while changes to an invariant feature would.

Although field sparrows use all tested features in species song recognition, they *weight* information provided by some features more than they do information from variable features. Disparities along certain features, namely phrase structure and frequency, are more likely to interfere with recognition of a stimulus than are equally-noticeable disparities in other features, *e.g.*, inter-note interval and note duration (Experiments 20-25). By scaling each feature separately I was able to ensure that the acoustic differences between the two states of each feature compared in Experiments 20 to 25 had approximately equivalent perceptual effects when tested singly. Otherwise, any attentional preferences that the birds might have would be masked or confounded with differences in the degree to which acoustic differences between feature states were noticeable.

Song frequency is the most important feature in species-song recognition. This may be inferred from the ranking of features according to their importance as species-song recognition cues derived from the results of Experiments 20-25. Variation in song frequency interfered with recognition more than any of the other three features (Experiments 20-22). Number of phrases and trill-note duration are tied for intermediate importance, and inter-note interval is the least important feature, as variations in it were tolerated more than any other feature.

The results do not cleanly distinguish between the sound-environment and invariant feature hypotheses discussed in the introduction. The prediction shared by both hypotheses, that frequency is the most important species song recognition cue, is borne out by these results. The ranking of feature weightings apparently used by the birds agrees in part with a ranking of features based on their relative variability in natural song. Frequency is both the least variable feature (coefficient of variation = 6%) and the most important recognition cue. However, if relative variability is the only determinants of feature's importance as a species song recognition cue, then the birds should have weighted trill-note duration, the most variable feature (c.v. = 38%), the least important. Instead, trill-note duration appears to be as important as the number of phrases (c.v. = 19%), and inter-note interval (26%) is accorded the least importance. The equivalent perceptual ranking of number of phrases and trill-note duration is consistent with the sound-environment hypothesis, but inter-note interval was also predicted to be as important as these two features. EMLEN (1972) pointed out that not all invariant features are apparently used in species song recognition. These results may provide a explanation. Although number of phrases varies significantly less than trill-note duration, it is no more effective in statistically distinguishing among species' songs (NELSON, unpubl. data). Field sparrows may ignore the variability difference and weight these features equally because they are equally effective discrimination cues.

Field experiments on other species have shown that frequency information is used in song recognition, but its importance relative to other cues is not known in such cases (FALLS, 1963; BREMOND, 1968, 1986; M. SCHUBERT, 1971; HELB, 1973; BECKER *et al.*, 1980; DABELSTEEN & PEDERSEN, 1985). Laboratory studies using operant techniques on artificial or natural stimuli indicate that absolute frequency is an important constraint on birds' abilities to discriminate stimuli (HULSE & CYNX, 1985; PARK & DOOLING, 1985; DOOLING *et al.*, 1987). Thus, there is converging evidence from different experimental paradigms indicating that frequency is an important feature influencing the perception of avian vocal signals. Also noteworthy is MARGOLIASH's (1986) demonstration by multi-unit recordings in a song system nucleus (HVc) in the brain of the white-crowned sparrow (*Zonotrichia leucophrys*). Responses at locations reactive to 'autogenous song' decreased monotonically with increasing frequency-shift relative to the bird's own song. Frequency shifts of ± 500 Hz, produced test songs within the normal range of intra-dialect variation for 'several' songs (MARGOLIASH, 1986) and resulted in normalized response decrements of only about 25%; whereas frequency shifts of ± 1000 Hz produced decrements on the order of 75%. There is a highly-suggestive correspondence between the neurophysiological response to frequency-shift and the territorial response assayed in this paper.

In conclusion, these results indicate that field sparrows integrate information from at least five different song features in the process of species song recognition. The use of multiple features is in accord with a statistical analysis of song variation in a community of 13 species which revealed that single features were not sufficient to accurately recognize song, but that seven features used simultaneously provided the best recognition performance (NELSON, unpubl. data). Although there is a high degree of redundancy among features, each feature has small, additive effects in improving recognition accuracy. The partial redundancy among different song features in recognition has been stressed by previous workers (BRÉMOND, 1968; BECKER, 1982) and possibly is an adaptation facilitating communication in an acoustic channel subject to perturbations and masking by different types of noise. The present results are novel in demonstrating that field sparrows differentially weight the information provided by different features. The perceptual mechanisms underlying these attentional preferences, apparently designed to optimize song recognition accuracy, remain to be determined.

Summary

Computer-modified songs were presented to wild territorial male field sparows to assess the relative importance of different acoustic features in species-song recognition. The features tested exhibited differing degrees of relative variability within field sparrow song. The experiments demonstrate that field sparrows are sensitive to changes in song frequency, number and order of phrases, trill-note duration, note 'shape', and internote interval. Field sparrows appear to be 'tuned' to the normal range of varition in song features. Responses decreased significantly when features in the experimental songs were changed by two to three standard deviations relative to a control song. These results contrast with previous conclusions that birds only use 'invariant' features in species-song recognition. However, manipulations of invariant features, frequency and number of phrases, interfered with recognition more than did an equally-noticeable change to a variable feature, inter-note interval. The number of phrases in a song and the duration of trill notes appear to be equally important recognition cues. Thus, field sparrows integrate information from at least five different features in species song recognition, but give more weight to information from an invariant feature, song frequency, than they do to information from variable features.

References

BECKER, P. H. (1976). Artkennzeichnende Gesangsmerkmale bei Winter- und Sommergoldhänchen (*Regulus regulus*, *R. ignicapillus*). — Z. Tierpsychol. 42, p. 411-437.

—— (1982). The coding of species-specific characteristics in bird sounds. — In: Acoustic communication in birds, Vol. 1 (Kroodsma, D. E. & Miller, E. H., eds). Academic Press, NY, p. 214-252.

——, Thielcke, G. & Wustenberg, K. (1980). Der Tonhöhenverlauf ist entscheidend für das Gesangserkennen beim mitteleuropäischen Zilpzalp (*Phylloscopus collybita*). — J. Orn. 121, p. 229-244.

Brémond, J.-C. (1968). Recherches sur la sémantique et les élémentes vecteurs d'information dans les signaux acoustiques du rouge-gorge (*Erithacus rubecula*). — La Terre et Vie 2, p. 109-220.

—— (1976). Specific recognition in the song of Bonelli's warbler (*Phylloscopus bonelli*). — Behaviour 58, p. 99-116.

—— (1978). Acoustic competition between the song of the wren (*Troglodytes troglodytes*) and songs of other species. — Behaviour 65, p. 89-97.

Dabelsteen, T. (1984). An analysis of the full song of the blackbird *Turdus merula* with respect to message coding and adaptations for acoustic communication. — Ornis. Scand. 15, p. 227-239.

—— & Pedersen, S. B. (1985). Correspondence between messages in the full song of the blackbird *Turdus merula* and meanings to territorial males, as inferred from responses to computerized modifications of natural song. — Z. Tierpsychol. 69, p. 149-165.

Doherty, J. & Hoy, R. (1985). Communication in insects. III. The auditory behavior of crickets: some views of genetic coupling, song recognition, and predator detection. — Quart. Rev. Biol. 60, p. 457-472.

Dooling, R. J., Brown, S. D., Park, T. J., Okanoya, K. & Soli, S. (1987). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): I. Pure tones. — J. Comp. Psychol. 101, 139-149.

——, Clark, C., Miller, R. & Bunnell, T. (1982). Program package for the analysis and synthesis of animal vocalizations. — Behav. Res. Methods & Instr. 14, p. 487.

Emlen, S. T. (1972). An experimental analysis of the parameters of bird song eliciting. — Behaviour 41, p. 130-171.

Engineering Design (1987). Signal software manual. — Engineering design, 43 Newton St., Belmont, MA, p. 1-158.

Falls, J. B. (1963). Properties of bird song eliciting responses from territorial males. — Proc. 13th Internl. Orn. Congr. p. 259-271.

Gaioni, S. J. & Evans, C. S. (1986). Perception of distress calls in mallard ducklings (*Anas platyrhynchos*). — Behaviour 99, p. 250-274.

Goldman, P. (1973). Song recognition by field sparrows. — Auk 90, p. 106-113.

Helb, H. W. (1973). Analyse der artisolierenden Parameter im Gesang des Fitis (*Phylloscopus trochilus*) mit Untersuchungen zur Objektivierung der analytischen Methode. — J. Orn. 114, p. 145-206.

Hulse, S. H. & Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (*Mimus*, *Molothus*, and *Sturnus*). — J. Comp. Psychol. 99, p. 176-196.

Konishi, M. (1965). The role of auditory feedback in the control of vocalization in the white-crowned sparrow. — Z. Tierpsychol. 22, p. 770-783.

Lanyon, W. E. (1978). Revision of *Myiarchus* flycatchers of South America. — Bull. Amer. Mus. Nat. Hist. 161, p. 429-627.

McGregor, P. K., Krebs, J. R. & Ratcliffe, L. M. (1983). The reaction of great tits (Parus major) to playback of degraded and undegraded songs: the effect of familiarity with the stimulus song type. — Auk 100, p. 898-906.

Marler, P. (1960). Bird songs and mate selection. — In: Animal sounds and communication (Lanyon, W. E. & W. N. Tavolga, eds). Am. Inst. Biol. Sci. Publ. No. 7. Washington, D.C., p. 348-367.

—— (1976). Sensory templates in species-specific behavior. — In: Simpler networks and behavior (FENTRESS, J. C., ed.). Sinauer, Sunderland, MA, p. 314-329.

—— (1977). Development and learning of recognition systems. — In: Recognition of complex acoustic signals (BULLOCK, T. H., ed.). Dahlem Konferenzen, Berlin, p. 77-96.

MARGOLIASH, D. (1983). Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. — J. Neurosci. 3, p. 1039-1057.

—— (1986). Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. — J. Neurosci. 6, p. 1643-1661.

MILLER, E. H. (1982). Character and variance shift in acoustic signals of birds. — In: Acoustic communication in birds, Vol. 1 (KROODSMA, D. E. & MILLER, E. H., eds). Academic Press, NY, p. 253-295.

MORRISON-PARKER, C. (1977). Experimental investigations of the vocal behavior of the field sparrow (*Spizella pusilla*). — Unpubl. Ph.D. thesis, Indiana University.

MORTON, E. S. (1982). Grading, discreteness, redundancy, and motivation-structural rules. — In: Acoustic communication in birds, Vol. 1 (KROODSMA, D. E. & MILLER, E. H., eds). Academic Press, NY, p. 183-212.

PARK, T. J. & DOOLING, R. J. (1985). Perception of species-specific contact calls by the budgerigar (*Melopsittacus undulatus*). — J. Comp. Psychol. 99, p. 391-402.

PETERS, S. S., SEARCY, W. A. & MARLER, P. (1980). Species song discrimination in choice experiments with territorial male swamp and song sparrows. — Anim. Behav. 28, p. 393-404.

SCHUBERT, M. (1971). Untersuchungen über die reaktionauslösenden Signalstrukturen des Fitis-Gesanges, *Phylloscopus trochilus* (L.), und das Verhalten gegenüber arteigenen Rufen. — Behaviour 38, p. 250-288.

SHIOVITZ, K. A. & LEMON, R. E. (1980). Species identification of song by indigo buntings as determined by responses to computer generated sounds. — Behaviour 74, p. 167-199.

SORJONEN, J. (1986). Factors affecting the structure of song and the singing behaviour of some northern European passerine birds. — Behaviour 98, p. 286-304.

ZOLOTH, S. R., DOOLING, R. J., MILLER, R. & PETERS, S. S. (1980). A minicomputer system for the synthesis of animal vocalizations. — Z. Tierpsychol. 54, p. 151-162.

Zusammenfassung

Männchen einer Freilandpopulation der Rotschnabelklapperammer *Spizella pusilla* wurden künstlich veränderte Gesänge vorgespielt, um die relative Wichtigkeit verschiedener akustischer Eigenschaften bei der Erkennung des arteigenen Gesanges zu erfassen. Die getesteten Parameter weisen im typischen Gesang der Art ein unterschiedliches Ausmaß an relativer Variabilität auf. Die Ergebnisse zeigen, daß *S. pusilla* auf Veränderungen 1. in der Tonhöhe, 2. in der Anordnung und der Anzahl der Gesangsphrasen ("phrases"), 3. in der Dauer der Elemente des Endtrillers, 4. in der "Gestalt" der Elemente und 5. in dem Abstand zwischen einzelnen Elemente reagieren. *S. pusilla* ist offensichtlich "abgestimmt" auf die im Artgesang auftretende Variabilität bestimmter Parameter; Reaktionen wurden signifikant schwächer, wenn diese Parameter sich von einem Kontrollgesang um zwei bis drei Standardabweichungen unterschieden. Die Ergebnisse sprechen gegen frühere Auffassungen, daß Vögel nur solche Gesangscharakteristika bei der Erkennung des arteigenen Gesanges benutzen, die wenig variabel ("invariant") sind. Gleichwohl wirken sich Veränderungen in den sehr konstanten Parametern Tonhöhe und Anzahl der Gesangsphrasen stärker aus als entsprechende Änderungen bei dem stärker variablen Parameter Elementabstand. Die Anzahl Gesangsphrasen und die Dauer der