ACOUSTIC PARAMETERS UNDERLYING THE RESPONSES OF SONG-SPECIFIC NEURONS IN THE WHITE-CROWNED SPARROW[1]

DANIEL MARGOLIASH

Beckman Laboratories of Behavioral Biology, Division of Biology 216-76, California Institute of Technology, Pasadena, California 91125

Abstract

Songbirds such as the white-crowned sparrow memorize the song of conspecific adults during a critical period early in life and later in life develop song by utilizing auditory feedback. Neurons in one of the telencephalic nuclei controlling song have recently been shown to respond to acoustic stimuli. I investigated the auditory response properties of units in this nucleus using a technique that permitted great flexibility in manipulating complex stimuli such as song. A few of the units exhibited considerable selectivity for the individual's own song. In wild-caught birds, song-specific units exhibited intradialect selectivity. In those birds that sang abnormal songs due to laboratory manipulation of song exposure during the critical period for song learning, units were selective for the abnormal songs. By systematic modification of a song, and by construction of complex synthetic sounds mimicking song, the acoustic parameters responsible for the response selectivity were identified. Song-specific units responded to sequences of two song parts but not to the parts in isolation. Modification of the frequencies of either part of the sequence, or increasing the interval between the parts, varied the strength of the response. Thus, temporal as well as spectral parameters were important for the response. When sequences of synthetic sounds mimicking song were effective in evoking an excitatory response, the response was sensitive to the aforementioned manipulations. With these techniques it was possible to elucidate the acoustic parameters required to excite song-specific units. All songs of the repertoire eliciting a strong excitatory response contained the appropriate parameters, which were missing from all weakly effective, ineffective, or inhibitory songs.

A fundamental problem in the study of integrative brain mechanisms is the question of the codes used by the nervous system. Behavioral phenomena such as learning, memory, and complex sensory perceptions may manifest themselves as particular states of activity in large aggregates of neurons, the activity of individual cells representing only limited information. Alternatively, these phenomena may be encoded in the brain by the activity of small, possibly redundant, populations of

highly specialized neurons. Furthermore, it is entirely conceivable that both types of organization coexist within the same neural system. There is a paucity of data directly supporting either hypothesis, although functional hierarchical organization is well established for many systems. Because the response properties of single neurons vary in complexity, and because the final output or product of the nervous system is behavior, this suggests that perhaps some very selective neurons exist that encode behaviorally relevant stimuli. Such neurons would represent a population near or at the top of a sensory hierarchy.

A wealth of behavioral observations on birdsong indicates that it is an attractive system for the study of neural encoding of complex stimuli. Song is essential for a songbird's reproductive success. In the majority of species, males sing both as an advertisement of territoriality and for mate attraction (Thorpe, 1961). For some species or subspecies of birds, songs may vary across but be similar within a locale; these song types are termed dialects (e.g., Marler and Tamura, 1962). Song is learned and is sensitive to the effects of sensory deprivation

during a critical period early in life (Thorpe, 1961). In some species, such as the white-crowned sparrow (*Zonotrichia leucophrys*), the critical period closes months before the bird enters a motor development phase that results in adult, "crystallized" song (Marler, 1970). Auditory feedback is necessary for normal motor development (Konishi, 1965). This observation led to the hypothesis that the song model learned during the critical period is stored as a neural memory trace or "template" that is recalled or accessed via auditory feedback during the song crystallization process (Konishi, 1965, 1978). A distinct group of nuclei involved in song production in adult birds has been described by Nottebohm et al. (1976). Two of these nuclei, HVc (hyperstriatum ventrale, pars caudale) and RA (nucleus robustus archistriatalis), are contiguous with areas that receive projections from field L (Kelly and Nottebohm, 1979), the primary auditory telencephalic nucleus in birds. The auditory recipient zone associated with HVc forms a layer along the ventral border of HVc and has been termed the "shelf." In an intracellular study of anesthetized birds, Katz and Gurney (1981) reported that many HVc neurons responded weakly to noise bursts. Furthermore, multiunit activity from HVc of unanesthetized and minimally restrained birds responded to recordings of conspecific songs, including the birds' own song, and gave weaker responses to songs played in reverse (McCasland and Konishi, 1981). The present study utilized an uncommon but behaviorally relevant stimulus search paradigm, with emphasis on the bird's own song, to investigate the response properties of single HVc and shelf units recorded extracellularly. This stimulus paradigm has proven to be a powerful tool.

Materials and Methods

White-crowned sparrows of the sedentary Nuttalli race were collected from the field either as nestlings, fledglings, or adults. Many adults and fledglings were collected at one time from one locale (Bodega Marine Laboratory, Bodega Bay, CA). Most but not all of these birds sang individual versions of the Bodega Bay dialect. Nestlings were housed individually in sound attenuation boxes and exposed to various song models ("tutored") during the critical period. Some nestlings copied the song(s) they were tutored in, with varying degrees of accuracy, whereas others sang crystallized but isolate-like songs wholly different from either their tutor song(s) or any wild-type song.

One to several days prior to an experiment, Equithesin-anesthetized birds (Jensen Salsbery: 2 ml/kg, i.m.) were implanted with a stainless steel post. On the days of experiment, birds were anesthetized with urethane (Sigma: 20%, 0.4 ml/kg, i.m.). Body temperature was monitored with a lubricated (KY Jelly, Johnson and Johnson) intercloacal probe and maintained with external heating. Glass-coated platinum/iridium electrodes were advanced with a hydraulic microdrive (Trent Wells) through a small opening in the skull centered at 2.0 mm lateral to the bifurcation of the midsagittal sinus. These procedures enabled recordings from individual units to be maintained for several hours. Electrolytic lesions were placed at the recording sites of most complex units, and

at other fiduciary locations. At the end of an experiment, the animal was administered a lethal dose of Equithesin and exsanguinated and fixed in 10% formaldehyde via intracardial perfusion. Standard frozen section histology (30-μm sections, cresyl violet staining) enabled the reliable recovery of electrolytic lesions.

Because the method of stimulus manipulation chosen for these experiments has not been commonly used, a somewhat detailed description follows, including basic technical results and discussion. The method stems from the observation of Greenewalt (1968) that many bird vocalizations, especially songs, fall into the class of signals that can be accurately modeled as a single sinewave ("voice") whose frequency and amplitude varies with time. For such signals, analysis is simplified in that frequency information can be extracted from the easily detected zero crossings of the signal (e.g., Straddon et al., 1978). Accurate resynthesis of the original signals is also simplified in that the frequency and amplitude of only a single oscillator need be controlled. Furthermore, signals thus represented can be systematically modified parameter by parameter, each new signal requiring only milliseconds of computation for a medium-sized laboratory computer.

Song analyses, on-line song resynthesis, and response rasters and histograms were implemented on a computer-based system (PDP 11/40). Birds were placed in small acoustic isolation chambers and were induced to sing by a subcutaneous testosterone implant. Songs were recorded with an analog tape recorder (TEAC). The recorded signals were then low-pass filtered at 9.5 kHz (8 pole filter) and digitized with 12-bit resolution at a 20-kHz sampling rate. Off-line programs extracted frequency and amplitude information for 1-msec intervals and established the starting and ending times for each part or phrase in the song. To resynthesize a song, successive values representing frequency and amplitude were converted to voltages (12-bit digital to analog converters) at 1-msec intervals. These voltages controlled a custom-designed amplifier. The analog hardware converted successive voltage steps to linear frequency and amplitude modulations. The amplitude was variable over 80 dB (adjusted to 0 dB = 20 μPa), and distortion products were 40 dB below signal intensity. Sound was delivered in freefield conditions by a small speaker (Lafayette: ± 2.5 dB, 1.5 to 11 kHz) mounted 1.7 m from the bird inside a 2 × 2 × 2 m sound attenuating chamber. The double walled, fiber glass-lined chamber attenuated room noises by at least 80 dB for frequencies above 300 Hz.

The accuracy of analysis/resynthesis is assessed by use of sound spectrographic analysis, which is not constrained to representing a single voice (Fig. 1). A white crown song typically consists of three to five parts or phrases. For example, the song of Figure 1A consists of three phrases refered to, in order, as whistle, buzz, trill. The trill comprises nine syllables. All of the syllables except for the initial one comprise several notes or elements, defined as the smallest individual unit of sound in a birdsong. Those phrases that best meet the single voice criteria, whistles and trills, show the least modification by this technique. Broader bandwidth signals such as buzzes may be less accurately resynthesized. Nuttalli
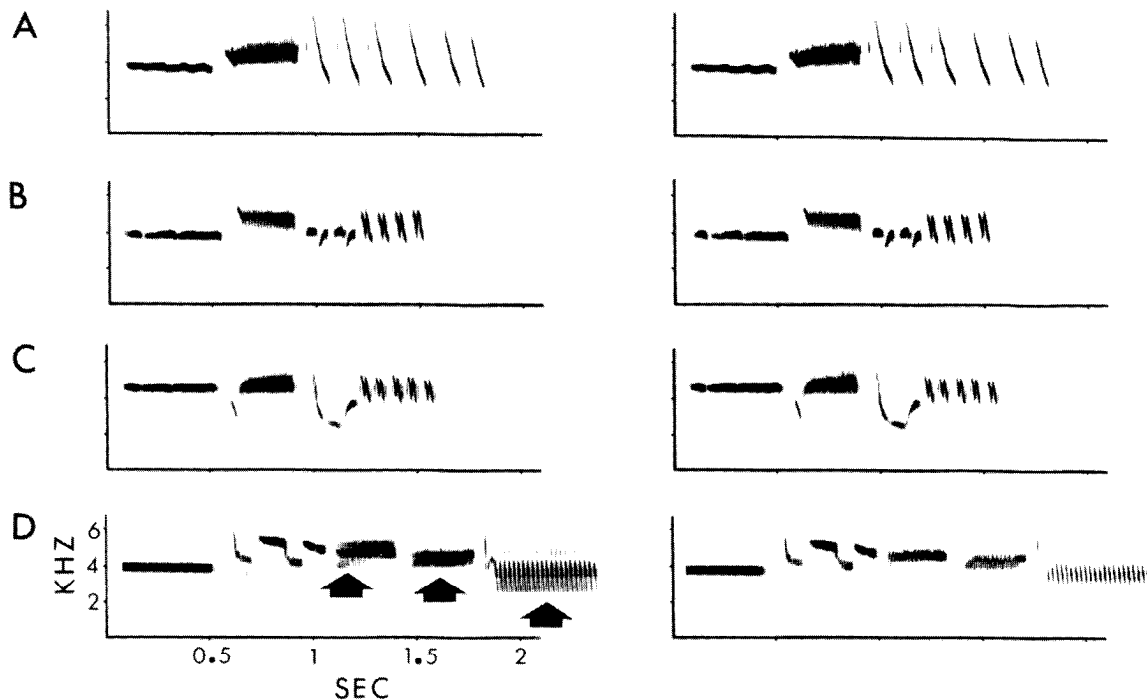
Figure 1. Resynthesis of white-crowned sparrow songs utilizing zero crossing analysis. Sonagrams—frequency versus time plots—in the *left column* are directly from tape recordings of singing birds. Sonagrams in the *right column* are computer-resynthesized versions using frequency and amplitude functions extracted from the original songs (see the text). *A*, Nuttalli (coastal) race. *B* and *C*, Oriantha (montane) race. *D*, Gambelii (Alaskan) race. In *A* and *B* reproduction is essentially perfect; in *C* small differences somewhat obscured by the photographic process appear in the buzz (second phrase); in *D* clear differences are present in many phrases (see *arrows*).

race white crowns, such as those found at Bodega Bay, sing songs that are relatively rich in whistles and trills. The computer, sound spectrogram (sonagram), and oscillogram representations of one such song are shown in Figure 2. Note the correspondance between the various representations: the computer-generated frequency function (Fig. 2B, *upper trace*) is similar to the sonagram (Fig. 2A), whereas the computer-generated amplitude function (Fig. 2B, *lower trace*) follows the outline of the oscillogram (Fig. 2C). Differences between the sonagram and frequency function are in part due to amplitude information in the sonagram, represented by the darkness of the line. In comparing the oscillogram and amplitude function, note that the ordinates for the two representations are linear and logarithmic, respectively. To demonstrate the flexibility of parameter modification available during the resynthesis, a highly modified version of the song is shown (Fig. 3). In contrasting Figure 3 with Figure 2, note that phrase order and interphrase intervals have been changed. One phrase has been eliminated. The whistle is increased in frequency and is played in reverse, the buzz has an artificial amplitude (but not frequency) envelope, and the frequencies in the trill have been low-pass filtered (smoothed). Expressing signals according to their frequency and amplitude versus time components enables independent modification of frequency and amplitude. For example, manipulation of the computer-generated frequency function (Fig. 3B, *upper trace*) is reflected in the sonagram (Fig. 3A), but with no change in the oscillogram (Fig. 3C); the converse

is apparently true for changes of the amplitude function. The expected results may be confounded by introducing or removing amplitude modulations that are sufficiently rapid as to induce significant frequency components. For the experiments reported in this paper, such modifications were avoided. It can, therefore, be seen that this approach combines a reasonable degree of accuracy in representation of the original stimulus with substantial on-line flexibility for parameter modification.

Sufficient analytic and synthetic power exists in this scheme to apply a systematic approach to the elucidation of the parameters underlying the responses of units to complex stimuli. In this study, units typically were tested with tone, narrow-band (600-Hz bandwidth), and wide-band (10-kHz bandwidth) noise bursts, the song of the experimental subject (bird's own song), and with the tutored song(s), available only for laboratory-reared birds. Numerous units were also tested with songs played in reverse. Spike activity during stimulus presentations were accumulated and displayed by the computer. Graphs were generated on-line to aid in deciding if a unit's responses to songs were predictable on the basis of the responses to simple stimuli. Those units exhibiting complex response properties, namely responses to song that were not predicted by the responses to simple stimuli, were investigated further by systematic parameter modification of the songs and by synthesis of complex artificial stimuli. The responses of numerous multiunit recording sites to various songs were recorded on magnetic tape for off-line analysis.
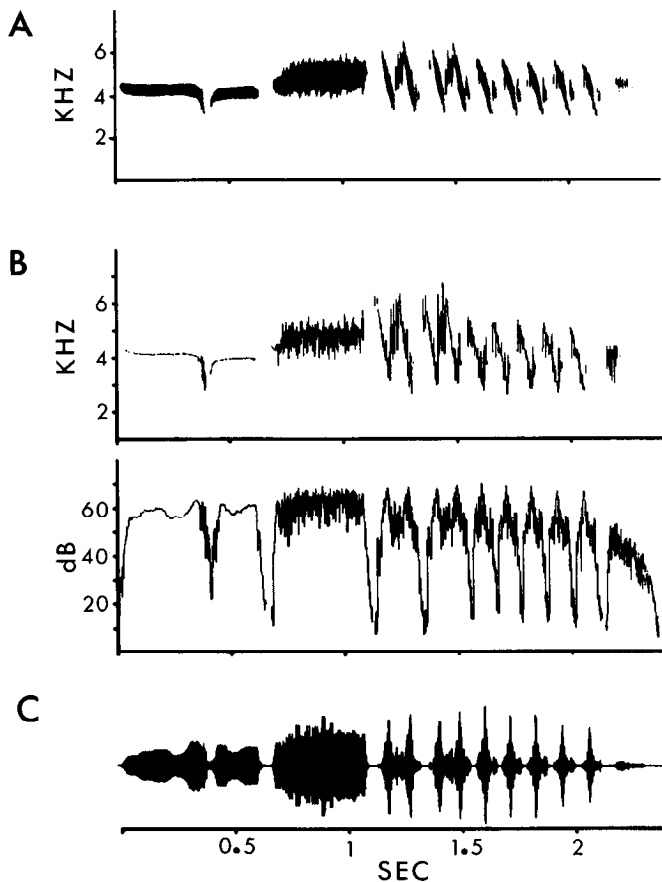
A

B

C

A

B

C

Figure 2. Relationship between computer, sonagram, and oscillogram representations. The sonagram in *A* and the oscillogram in *C* are analyses of a song produced by the computer-generated functions of *B*. The *top trace* in *B* is frequency versus time; the *bottom trace* is amplitude versus time. In this and all other examples in this paper of computer representations of frequency functions, the part of the otherwise continuous function corresponding to low amplitude values is blanked as a visual aid.

Figure 3. Modification of computer-generated functions. *B*, Computer-generated functions of the song of Figure 2 in highly modified form. Phrase order is rearranged from whistle-buzz-trill-terminal buzz to trill-whistle-buzz. Terminal buzz has been eliminated. Trill frequencies have been smoothed, and whistle frequencies have been increased 2 kHz. The whistle is played in reverse. The buzz has an artificial amplitude envelope. Note that frequency modifications change sonagram morphology (*A*) without modifying oscillogram morphology (*C*); the amplitude modification changes only the oscillogram.

Results

Overview. The response characteristics of 931 units, isolated in HVc and surrounding areas of 33 birds, were analyzed. Of these, 477 units were documented on the computer. Most units either did not respond or responded weakly to song and other sounds. A number of units responded vigorously to song in a manner predicted by their responses to tone bursts. These response classes are dealt with in a subsequent section (see "*Other response categories*"). The main emphasis of this report is on those few units that showed excitation to song and to complex artificial stimuli mimicking song, but not to simple stimuli such as tone bursts or narrow-band or wide-band noise bursts. The rationale for labeling these units as "song specific" is given in detail under "Discussion." For these units simple stimuli were ineffective or inhibitory; only a subset of songs from the repertoire elicited an excitatory response. Often, but not always, songs played in reverse proved to be ineffective. By elucidating the parameters common only to those natural and modified songs that were excitatory, it was possible on occasion to construct effective artificial stim-
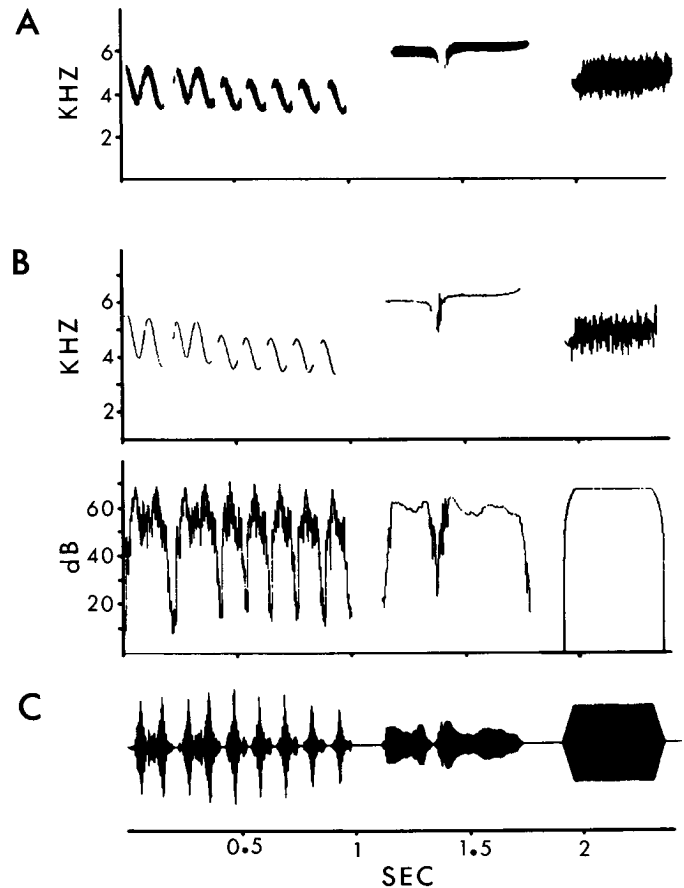
uli. Systematic modification of acoustic parameters within these artificial stimuli varied the units' responses in a manner consistent with the responses to various song parameters within the repertoire. Song-specific units were rare (27 well isolated units in 17 animals) and often habituated to stimulus repetition rates of greater than once every 5 to 7 sec. Thus, when necessary, randomly chosen repetition periods of 7 to 20 sec were used. These intersong intervals fall well within the range of behaviorally relevent intervals for singing white-crowned sparrows (personal observations). Song-specific units were typically encountered within an area that exhibited strong multiunit responses to song, and weaker responses to tone bursts of frequencies contained in the songs. Background responses were of a bursting nature. Such properties made the isolation of single units somewhat difficult and, combined with the need to maintain recordings for several hours, contributed to the paucity of song-specific units documented.

A striking and consistent observation was that two song phrases or two or more elements in the sequence
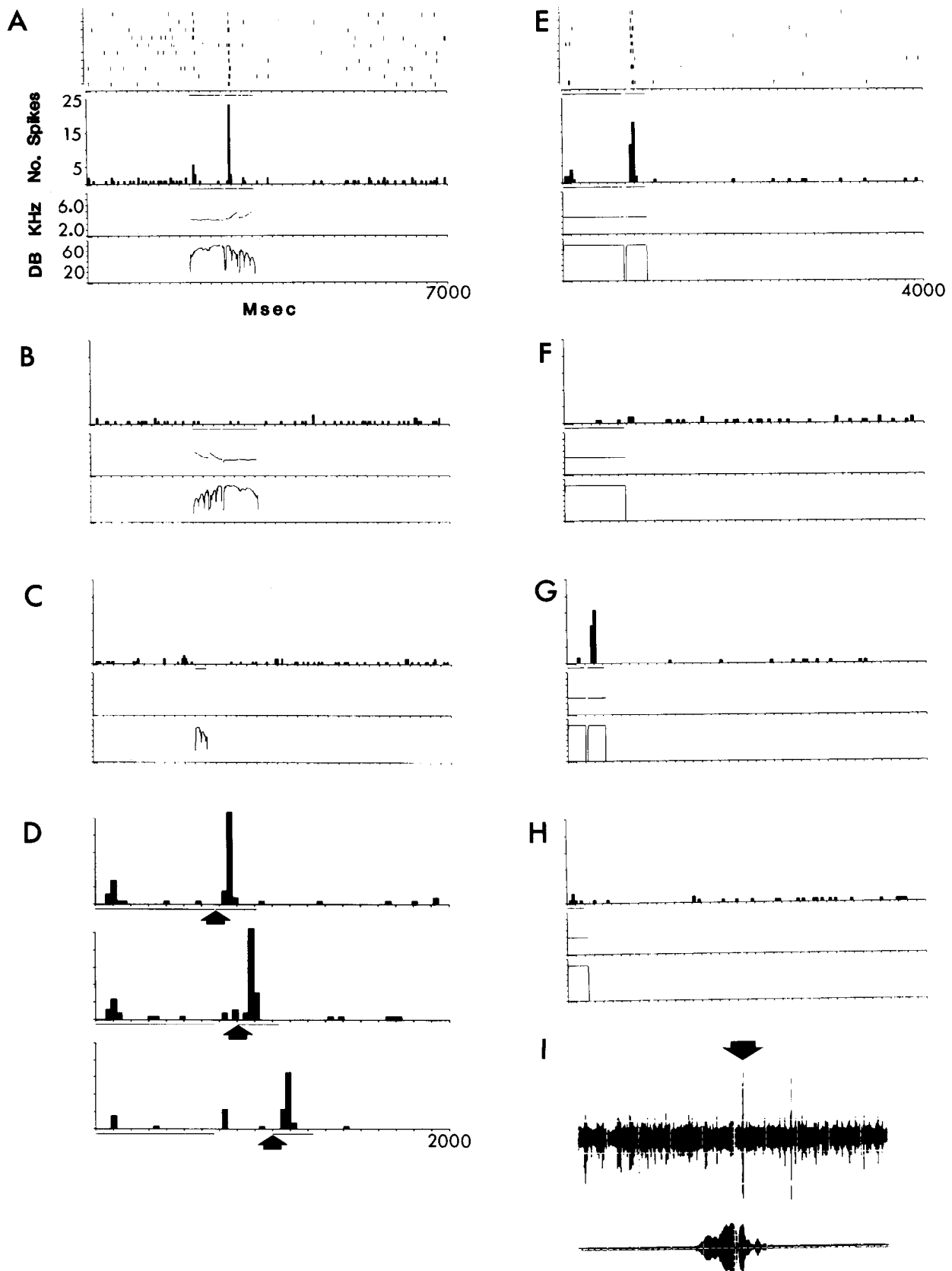
Figure 4. Response of song-specific whistle-whistle unit 560107. *A,* The *top section* is raster representation of unit response. Each vertical tick represents one spike; each row is marked on the ordinate by a horizontal tick and represents the response to one stimulus presentation. The *middle section* is a histogram of the raster; *ordinate:* number of spikes per bin. As for all response rasters and histograms in this paper, the *horizontal lines* directly underneath the plot are timing marks. The *bottom section* is a computer representation—frequency and amplitude versus time—of the stimulus, the bird's own song. The time axes of all plots are aligned. *B,* Histogram and reversed song stimulus only. *C,* Second phrase in isolation. *A* to *C,* Time axes 7000 msec. *D,* Histogram only for three stimuli: first and second phrase with 25-, 125-, 325-msec intervals. The *arrows* point to the onset of the second phrase. Note time axis is 2000 msec. *E* to *H,* Two-tone and single tone models; time axes 4000 msec. *I,* Actual recording, trace of unit response (*top trace, arrow*) to stimulus (*bottom trace,* bird's own song). The unit is well isolated. All histograms represent 10 repetitions, 30-msec binwidth.

from the song were required to elicit a response. For most song-specific units any individual phrase from the appropriate sequence elicited an inhibition, no excitation, and/ or weak excitation at the onset or offset. Sometime after the onset of the second phrase in the sequence, however, the unit responded vigorously. Modification of the frequency modulation or center frequency of the various phrases even without modification of the amplitude envelope could totally abolish the response. Temporal characteristics of the sequence strongly affected the response. A number of tests were designed to manipulate the temporal pattern within the song while leaving unaltered the total power spectrum. Changing the order of phrases within the sequence or reversing the direction of play of the otherwise normal sequence typically abolished the response. Increasing the interval between phrases over hundreds of milliseconds, however, diminished the response only gradually.

The song-specific units of a given bird were maximally stimulated by the song of that individual. Other songs were effective only if they contained similar phrases in the appropriate order. Thus, in birds that sang isolate-like songs—songs with abnormal morphologies never found in the wild—it was possible to find units that responded only to the bird's own song and not to any wild-type song in the repertoire. More remarkable, however, song-specific units in wild-caught Nuttalli race Bodega Bay adult birds only responded to a few songs within a large repertoire of quite similar songs all of the same dialect. Thus, such units exhibited specificity for song and intra- as well as interdialect selectivity.

Whistle-whistle units. Song-specific whistle-whistle units ($n = 7$) responded after the onset of the second of two whistles. Such whistle pairs occur frequently in the songs of Nuttalli race white crowns and are also common in isolate-like songs. The response of one such unit recorded from a bird that sang an isolate-like song is shown in Figure 4. The bird's own song consisted of an initial constant frequency phrase followed by two similar phrases of slowly increasing frequency, each of which started with a brief constant frequency segment (see Fig. 4A). The unit exhibited weak excitation at the onset of the first phrase followed by inhibition, but it responded vigorously after the onset of the second phrase (Fig. 4A). Inhibition followed that lasted throughout the final phrase and approximately 2 sec beyond. Reversed song was an inhibitory stimulus (Fig. 4B), indicating that the response was sensitive to the sequence of frequency or amplitude modulation within the song. Although the unit responded at the second phrase in the song, it did not respond to the second phrase presented alone (Fig. 4C). When the interval between the first and second phrases was increased even by as much as 300 msec, the response was not abolished but followed with constant latency the onset of the second phrase (Fig. 4D). Thus, the unit was not responding solely to the first phrase and instead required both phrases. The center frequency of the initial segment of the second phrase was essentially identical to the center frequency of the first phrase. With this in mind, the whistle pair was modeled as two successive tone bursts of identical frequencies. The duration of the tone bursts and the interval between them were taken from the song. These successive tone bursts of identical

frequencies elicited a weak excitation at the onset of the first tone followed by inhibition, and a strong excitation at the onset of the second tone followed by inhibition (Fig. 4E). Note the similarity with the response to the song (Fig. 4A). The first tone burst was by itself ineffective (Fig. 4F). Thus, the response properties of this unit were unequivocally not predictable on the basis of its responses to single tone bursts. The duration of the first of the tone bursts was not critical for the response. For instance, the two tone bursts could have identical durations (and all other parameters) yet the unit would respond only at the second tone burst of the pair (Fig. 4G). Note also that the ineffective single tone burst stimuli (Fig. 4, F and H) were intermediate in power to effective tone burst pairs (Fig. 4, E and G). In Figure 4I, the actual stimulus trace demonstrates the clear single unit nature of the recording.

These observations typifed some features of song-specific whistle-whistle unit response properties. The response could be elicited by consecutive tone bursts and thus did not require the fine frequency "jitter" from which a bird whistle derives its distinctive sound quality. Within the constraints of consecutive tone bursts, the timing information, namely duration of either of the
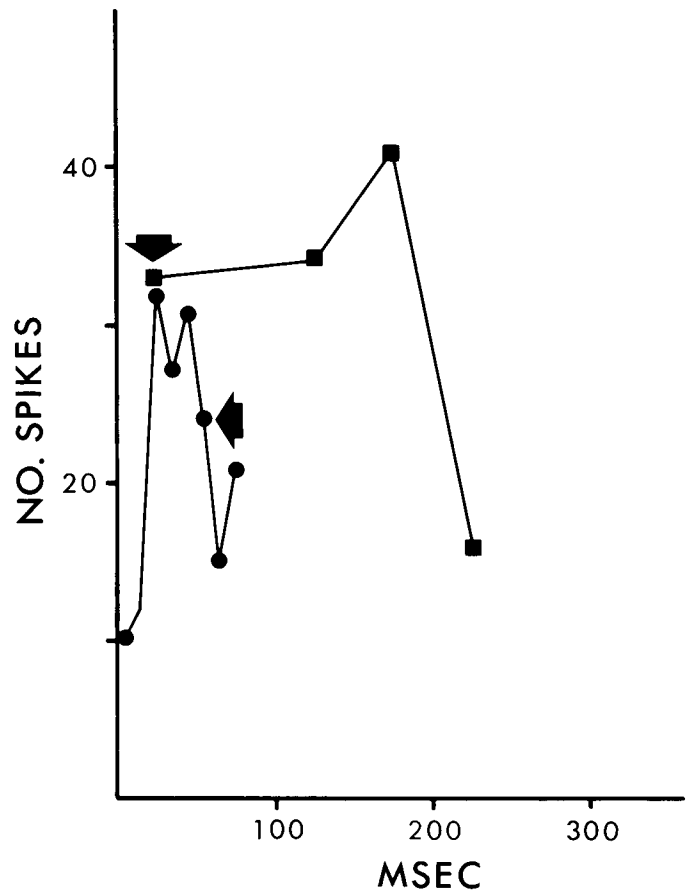

Figure 5. Intertone interval tuning. Each *point* represents the total number of spikes at the second of two tones as a function of the varying interval between the tones. Ten repetitions per *point*. *Arrows* point to the natural interval between corresponding whistles in the birds' own songs. The *leftmost curve* (*circles*) is the tightest intertone interval tuning seen to date. The *broader curve* (*squares*), for another unit in another bird, is a more typical example.

tones or the interval, was not critical to the unit's response. In fact, the timing information in the song was not necessarily optimal for evoking a response. Typical intertone interval tuning curves were broad and were not peaked around the interval in the bird's own song (Fig. 5). The maximum interval eliciting a response at the second tone could be several hundreds of milliseconds and was never less than approximately 100 msec (Fig. 5).
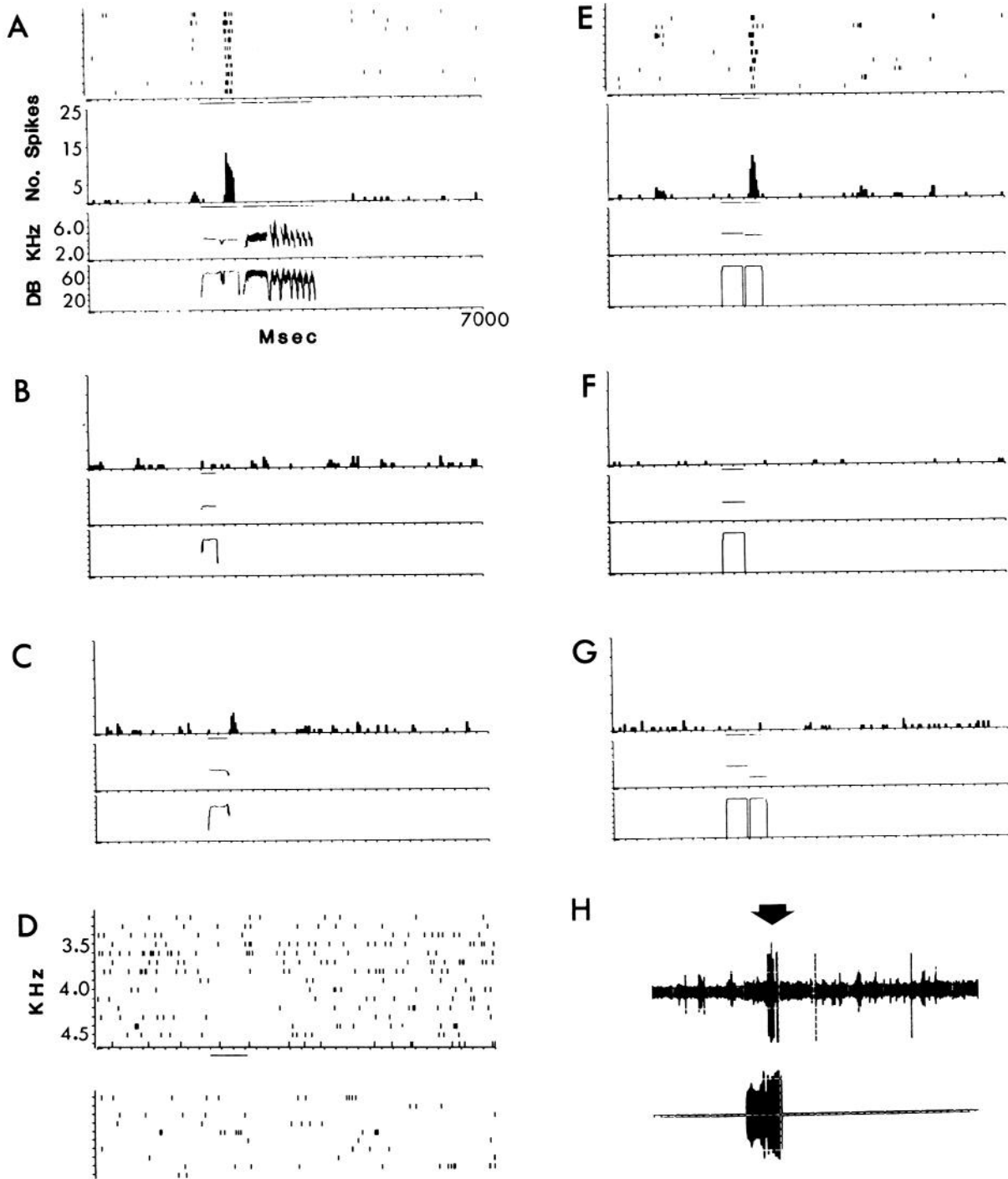

Figure 6. Response of whistle-whistle unit 810102. *A,* The stimulus is the bird's own song. Note the response at the second whistle of the first phrase, followed by approximately 2 sec of inhibition. *B,* Second whistle of first phrase. *C,* First whistle of first phrase. Note the weak response at offset. *D,* Response to single-tone bursts. The *top raster* is the response to individual tone bursts 3.2 kHz to 4.6 kHz in 100-Hz steps, covering a range of frequencies in the first phrase of the song. At 3.5 and 3.6 kHz, the unit seems to exhibit a slight offset excitation. The *bottom raster* is 10 repetitions of 3.6 kHz tone burst demonstrating no offset excitation. The frequency range of 500 Hz to 10 kHz in 500-Hz steps was also tested—no response (not shown). *E,* Consecutive tone model. Note the response at second tone. *F,* Second tone of model in isolation. The first tone was also tested—no response (not shown). *G,* Consecutive tone model as in *E* but second tone is decreased in frequency 1 kHz. *H,* Actual recording, trace of unit's response (*top trace, arrow*) to the isolated first phrase (*bottom trace*). Good single unit isolation. *A* to *G,* time axes 7000 msec; all histograms represent 10 repetitions, 30-msec binwidth.

Fine features of the amplitude modulation as found in the song, as well as the overall power, were not critical.

Song-specific units recorded in birds that sang abnormal, isolate-like songs exhibited considerable selectivity
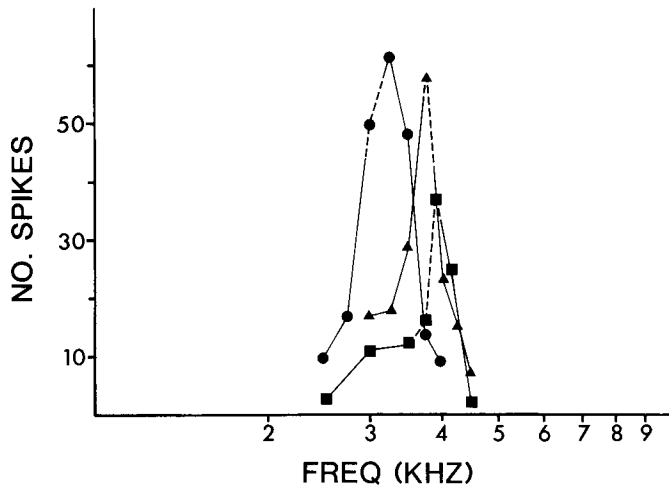

Figure 7. Consecutive tone tuning curves. Leading consecutive tone tuning for a song-specific whistle-whistle unit (*circles*). Following consecutive tone tuning for the same unit (*triangles*) and for another unit (*squares*)—see the text. The leading consecutive tone tuning curve represents a broadly tuned cell, the following consecutive tone tuning curves are examples of more sharply tuned cells. The *dashed portion* of each curve spans the range of frequencies present in the corresponding whistles from the original songs.

within a repertoire of naturally occurring songs. The most selective of the units responded only to the bird's own song and not to any of the six tutored songs ("choice" paradigm: five sympatric species and one white crown song) to which it was exposed to early in life, nor to any of nine other wild-type white crown songs tested. Song-specific whistle-whistle units were also recorded in birds that sang wild-type songs. The response of one such unit recorded in a bird singing a song of the Bodega Bay dialect is shown in Figure 6. The initial phrase, comprising two whistles, elicited a robust response at the second whistle of the pair, followed by approximately 2 sec of inhibition (Fig. 6*A*). In isolation, the second whistle of the phrase elicited no response (Fig. 6*B*), whereas the first whistle elicited only weak excitation at the offset (Fig. 6*C*). The unit was inhibited by single tone bursts (Fig. 6*D*) but was effectively stimulated by consecutive tone bursts of frequencies matching the center frequencies of the two whistles in the phrase (Fig. 6*E*). As predicted by the single tone burst tuning, individual elements of the model were ineffective (Fig. 6*F*).

The response decreased monotonically with variation in the frequency of either of the tone bursts of the model. For instance, the response was abolished by lowering the frequency of the second tone by 1 kHz while leaving all other parameters of the model unmodified (Fig. 6*G*). In this way, it was possible to quantify the response of whistle-whistle units as a function of the frequency of either tone burst (Fig. 7). This will be termed "leading," "following," or "dual consecutive tone tuning," depending
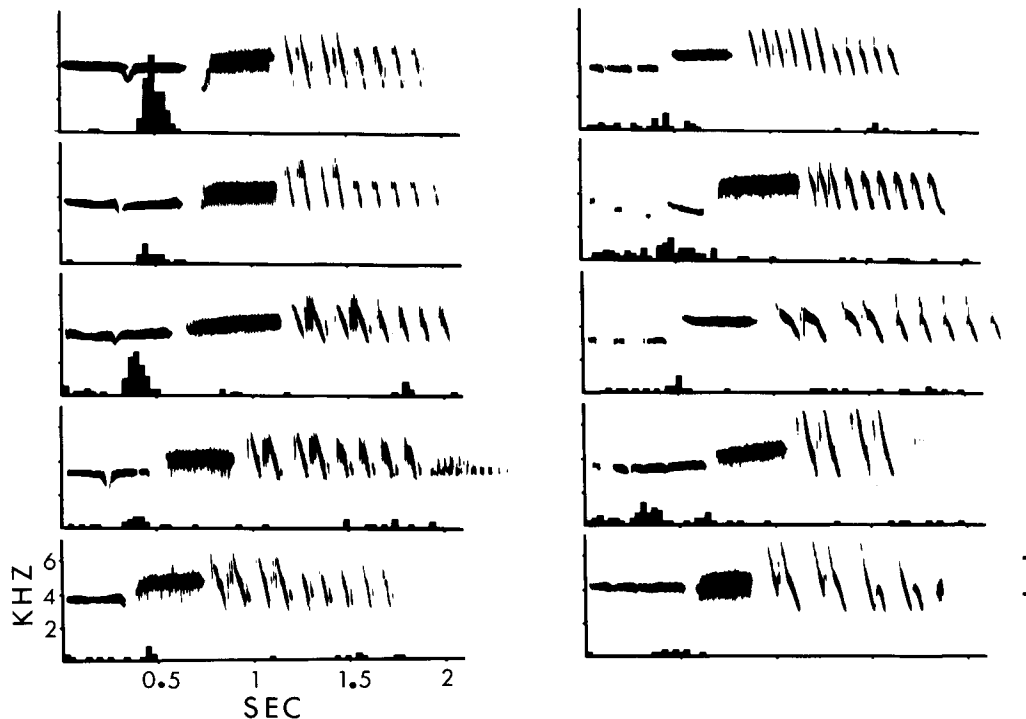

Figure 8. Song selectivity of unit 810102. Sonagrams of 10 white-crowned sparrow songs with response histograms superimposed. The five songs in the *left column* are all of the same Bodega Bay dialect; the *top song* is the bird's own song. Note that only one other song elicits strong excitation. The five songs in the *right column* are sample songs from five different dialects of the same race of white crowns (*top* to *bottom:* Berkeley, Tilden Park, Pacific Grove, Lompoc, Lake Merced, CA dialects). Note that no song elicits strong excitation. All histograms 20 repetitions, 30-msec binwidth. The *calibration mark* in *lower right* represents 10 spikes.

on which tone burst is being referred to. No systematic differences in tuning were observed for leading versus following consecutive tone tuning within the small sample size recorded ($n = 6$). Consecutive tone tuning for song-specific whistle-whistle units confirmed that the optimal frequencies for either of the tone bursts of the model typically fell within the range of frequencies present in the corresponding whistles of the natural (song) stimulus (Fig. 7).

The selectivity of the unit of Figure 6 within a group of wild-type songs is shown in Figure 8. Out of five Bodega Bay dialect songs, the bird's own song was most effective, and only one other song proved strongly excitatory. Eight other songs tested, including five samples of other Nuttalli dialects, elicited weak or essentially no excitation. These observations are consistent with the stimulus requirements as outlined above. That is, the two songs that evoked strong responses contained initial phrases comprising two whistles of appropriate frequencies (Fig. 9). The strength of response varied systematically with the frequencies of the two whistles, as predicted by the consecutive tone tuning for this unit. Differences in duration of the various whistles and intervals did not account for the response selectivity (Fig. 9). Five songs had whistles of approximately the same intensity as the whistle of the bird's own song; intensity did not account for the response selectivity. Those songs that contained single whistle phrases, even of the appropriate center frequency, did not elicit a response. An example is the song of Figure 8, *fifth row, first column,* center frequency = 3.8 kHz. For song-specific whistle-whistle units, all ineffective songs lacked one or more essential features, namely two whistles, or the appropriate frequencies. Significant variation in timing information did not affect the response.

Those whistle-whistle units that responded optimally to consecutive tones of different frequencies exhibited an interesting form of two-tone inhibition. For example, the

Figure 9. Parameters underlying song selectivity of unit 810102. Response histograms (20 repetitions, 30-msec binwidth) and corresponding frequency versus time plots for three Bodega Bay dialect phrases. The *leftmost* histogram is response to phrase in bird's own song. The dip in frequency at the middle of each phrase corresponds to the dip in amplitude. Note that the one other phrase that elicited a strong response has similar frequency, but not timing, parameters.

unit of Figure 10A responded at the second phrase of the isolate-like song of the bird but did not respond to either of the two whistle phrases in isolation (Fig. 10, *B* and *C*). While consecutive tones evoked a response at the second tone (Fig. 10D), either tone in isolation was ineffective (Fig. 10, *E* and *F*). When, in addition to the normal consecutive tone model, the second tone was played with the first tone (Fig. 11A) or the first tone was played with the second tone (Fig. 11C), the response was abolished. Because two tone bursts of equal intensity, played simultaneously produce a louder overall sound, control stimuli with the first (Fig. 11B) or second (Fig. 11D) tone burst 6 dB louder were tested. These control stimuli were effective, and thus the two-tone inhibition is not a result of an increase in stimulus intensity. The various observations outlined above are summarized under "Discussion," and a model is proposed.

Other song-specific units. Song-specific units were not restricted to a single response category. Combinations of phrases—whistle-buzz, buzz-trill, and buzz-buzz—as well as combinations of elements within single trill phrases, elicited responses from units exhibiting considerable selectivity and specificity. The complexity of stimulus requirements for these units complicated their detailed analysis. Within the small sample size (20 units for all classes), the response properties of few units were fully explored. However, to the extent that these units do not respond to simple consecutive tone models, they must represent one or more different classes of response.

An example of a song-specific buzz-trill unit is shown in Figure 12. The Bodega Bay dialect song of this subject evoked a strong response at the onset of the trill, the third phrase in the song, followed by approximately 2 sec of inhibition (Fig. 12A). Eliminating the whistle did not impair the response (Fig. 12B). Although the unit exhibited inhibition for the duration and/or weak excitation at the offset of tone bursts (Fig. 12C), it clearly responded to the isolated buzz at the offset (Fig. 12D). The isolated trill was only weakly excitatory (Fig. 12E). Attempts to stimulate this unit with consecutive tone models were unsuccessful. This particular unit showed the strongest response to an isolated phrase of any song-specific unit tested. Nevertheless, the response to the buzz in isolation showed some diminution in response strength and considerably greater temporal variation than the response to the buzz-trill combination. No song-specific unit responded as vigorously to an isolated phrase as it did to that phrase within its appropriate natural sequence.

Several song-specific units responded to the trill of the bird's own song ($n = 9$). The most extensively documented example is shown in Figure 13. The unit responded during the trill of the bird's own song (Fig. 13A) and equally well to the isolated trill (Fig. 13B), comprising three elements. Reversed trill (Fig. 13C) and tone or noise bursts (Fig. 13D) were ineffective stimuli. Slight filtering of the trill frequencies while leaving the amplitude function unaltered caused a slight decrease in the response (Fig. 13E), and extensive frequency filtering abolished the response (Fig. 13F). The stimulus requirements of this unit were quite closely matched to the trill parameters. Of six artificial models tested, only two (e.g., Fig. 13G) produced a moderately strong response. Four of the six models were of sufficiently similar acoustic
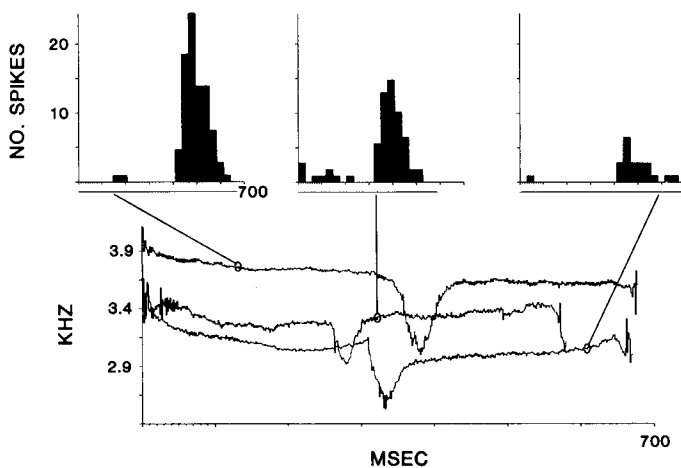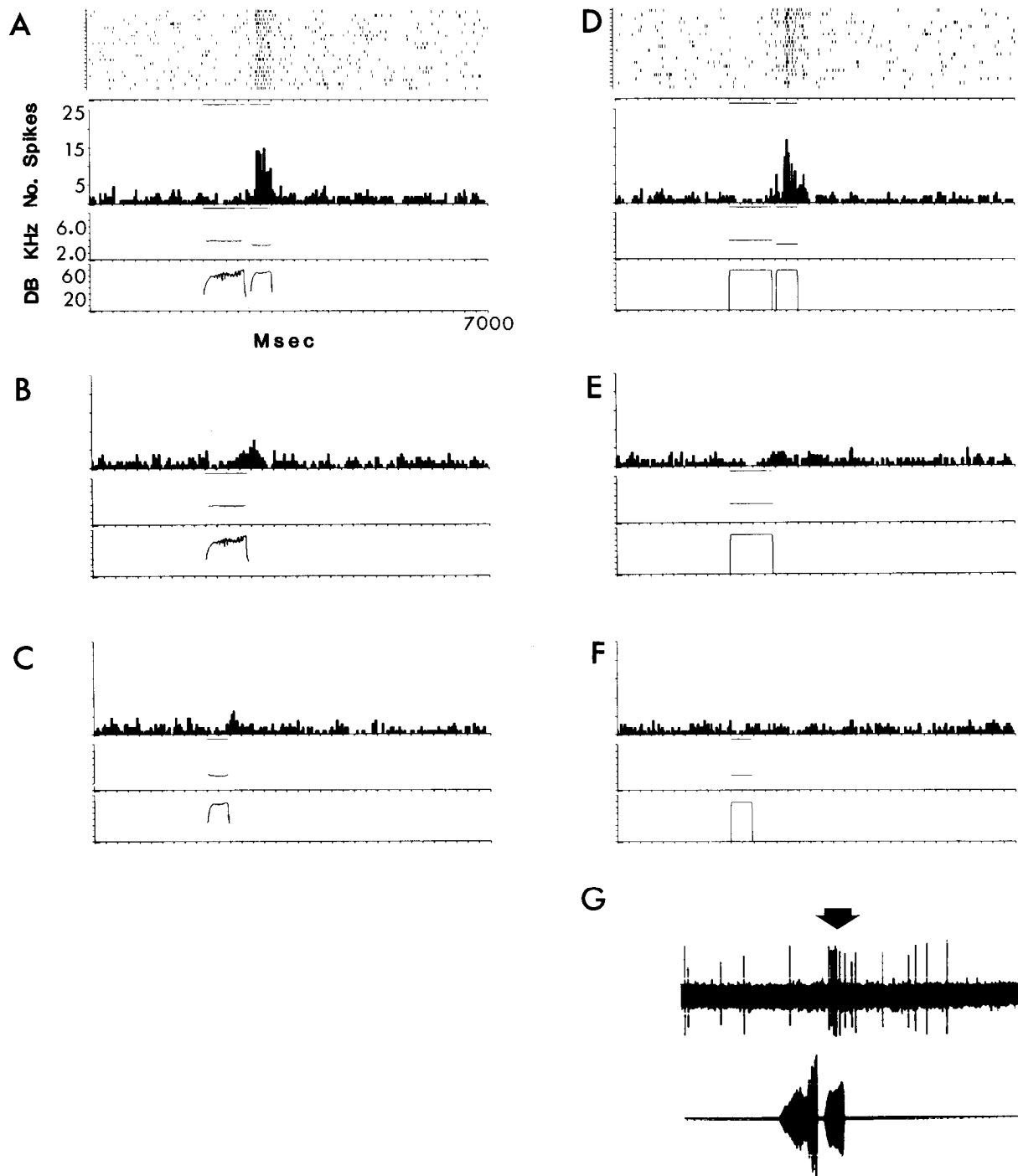
Figure 10. Response of whistle-whistle unit 840601. *A,* Strong response to bird's own song. *B,* Weak offset response or no response to first phrase in isolation. *C,* No response to second phrase in isolation. *D,* Strong response to consecutive tone model. Note similarity to *A. E,* Weak offset response or no response to first tone of model in isolation. Note similarity to *B. F,* No response to second tone of model in isolation. Note similarity to *C. A* to *F,* time axes 7000 msec; all histograms represent 20 repetitions, 30-msec binwidth. *G,* Actual recording trace of unit's response (*top trace, arrow*) to bird's own song (*bottom trace*).

parameters as to make the sounds hard to distinguish by a trained listener. Thus, this unit exhibited exquisite response specificity. Elimination of the initial element of the model totally abolished the response (Fig. 13*H*). Although numerous tests suggested themselves, the unit was lost shortly after the control series of Figure 13*A*, and thus the essential acoustic parameters underlying the response were not fully determined. This and other

song-specific trill units exhibited significant selectivity for song. In a remarkable example, a bird exposed to the "choice" paradigm during the critical period eventually sang a song very similar to the white crown song of the tutor repertoire. The bird's own song differed from the original tutored song only in slight detail in the trill section. Nevertheless, a song-specific trill unit isolated in that bird gave robust responses to the bird's own song,

while remaining virtually unresponsive to playback of the original tutor song.

The song-specific buzz-buzz unit of Figure 14 responded at the terminal buzz of the bird's own song, a Gambelii or Alaskan dialect song (Fig. 14A). Reversing the song abolished the response (Fig. 14B). Because individual buzzes are represented as rather symmetrical functions in time, this suggested that the unit was responding to a sequence of buzzes, which is not symmet-
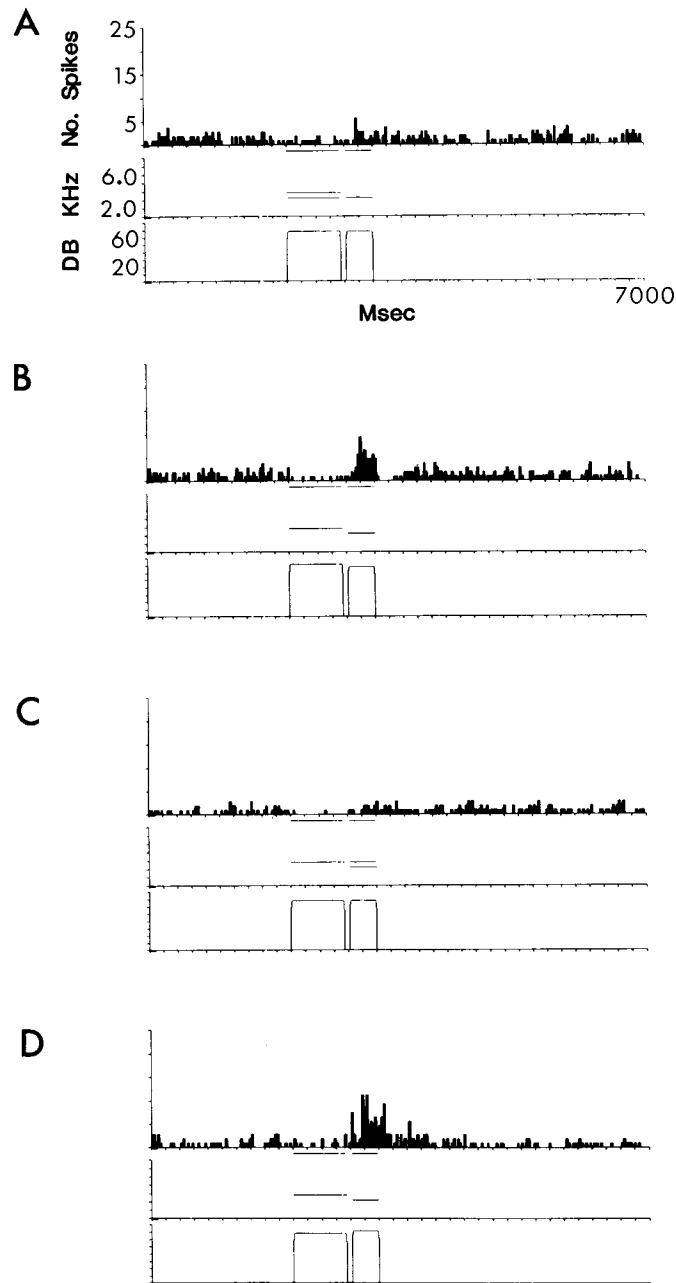

Figure 11. Two-tone inhibition of unit 840601. *A,* Response to consecutive tone model of Figure 10D is abolished by addition of another tone to the first part of the model. Additional tone is of same intensity and of frequency equal to tone of second part of model. *B,* Control stimulus; model of Figure 10D with first tone 6 dB louder does not inhibit response. *C,* Addition of tone to second part of model abolishes the response. *D,* Model with second tone 6 dB louder does not abolish the response. All parameters are as in Figure 10.

rical. Indeed, although tone bursts and narrow-band and wide-band noise bursts did not drive the unit (Fig. 14C), the final two buzzes in isolation elicited a strong response (Fig. 14D). As could be predicted, neither of those two buzzes in isolation elicited a response (Fig. 14, *E* and *F*). Because realistic but artificial buzzes were not available at the time of this experiment, attention was focused on this unit's response selectivity. Out of seven other songs, three Gambelii samples and four Oriantha or montane race dialects, only one elicited a response; this also was a Gambelii song. Taken as a whole, these observations suggest the existence of several classes of song-specific responses, but a firm conclusion awaits the complete elucidation of the response properties of such units.

Multiunit responses. Electrode penetrations through the HVc complex (see Fig. 15) often encountered areas with strong background multiunit activity, often with physiologically distinct borders. At these recording sites the considerable spontaneous activity was quite variable and of a bursting nature. Within these areas acoustic stimuli often evoked strong responses. For some recording sites, tone bursts were effective stimuli; in others song was far more effective. Multiunits responding to song often exhibited surprising selectivity. Reversed song was less effective and on occasion appeared to inhibit the spontaneous activity. Of various songs drawn from the repertoire, only those similar to the bird's own song elicited a strong response. Thus song selectivity was evident even in multiunit recordings. Unfortunately, the inherent inability to clearly quantify such recordings precludes a detailed analysis of the response properties (see "Discussion").

Anatomical distribution of functional properties. Of 27 song-specific units, 22 were isolated in areas that yielded responses clearly selective for the bird's own song within the song repertoire. In an attempt to map the associated lesion data, it was observed that HVc is shaped as an ellipsoid, with the major axis parallel to the midsagittal plane (Fig. 15A). Sagittal sections at any mediolateral position generate an elliptical shape for HVc. Lesion locations thus can be mapped onto an ellipse while preserving relative position. Although this mapping technique could in principle distort the distribution of locations near the lateral borders of HVc, where the shape and extent of the nucleus are not clear, this is unlikely to have affected the present data because virtually all of the lesions were confined to areas near the central sagittal plane of the nucleus. When lesions of units were so treated, an obvious clustering became apparent (Fig. 15B). To emphasize this clustering, lesions associated with tone burst unit recording sites were also mapped. These two groups of lesions segregated themselves within the HVc complex (Fig. 15B). Of 21 lesions of song-specific units, 19 occur at or above the major axis; of 17 tone burst unit lesions, only three occur at or above the major axis. This difference indicates a spatial organization of units with differing functional properties within the HVc complex. Note that although the sites of most song-specific units were lesioned at the end of a recording, only a small percentage of the more numerous tone burst responding units were so marked. That is, only in rare cases was the electrode advanced after lesioning the site of recording for a song-specific unit. As
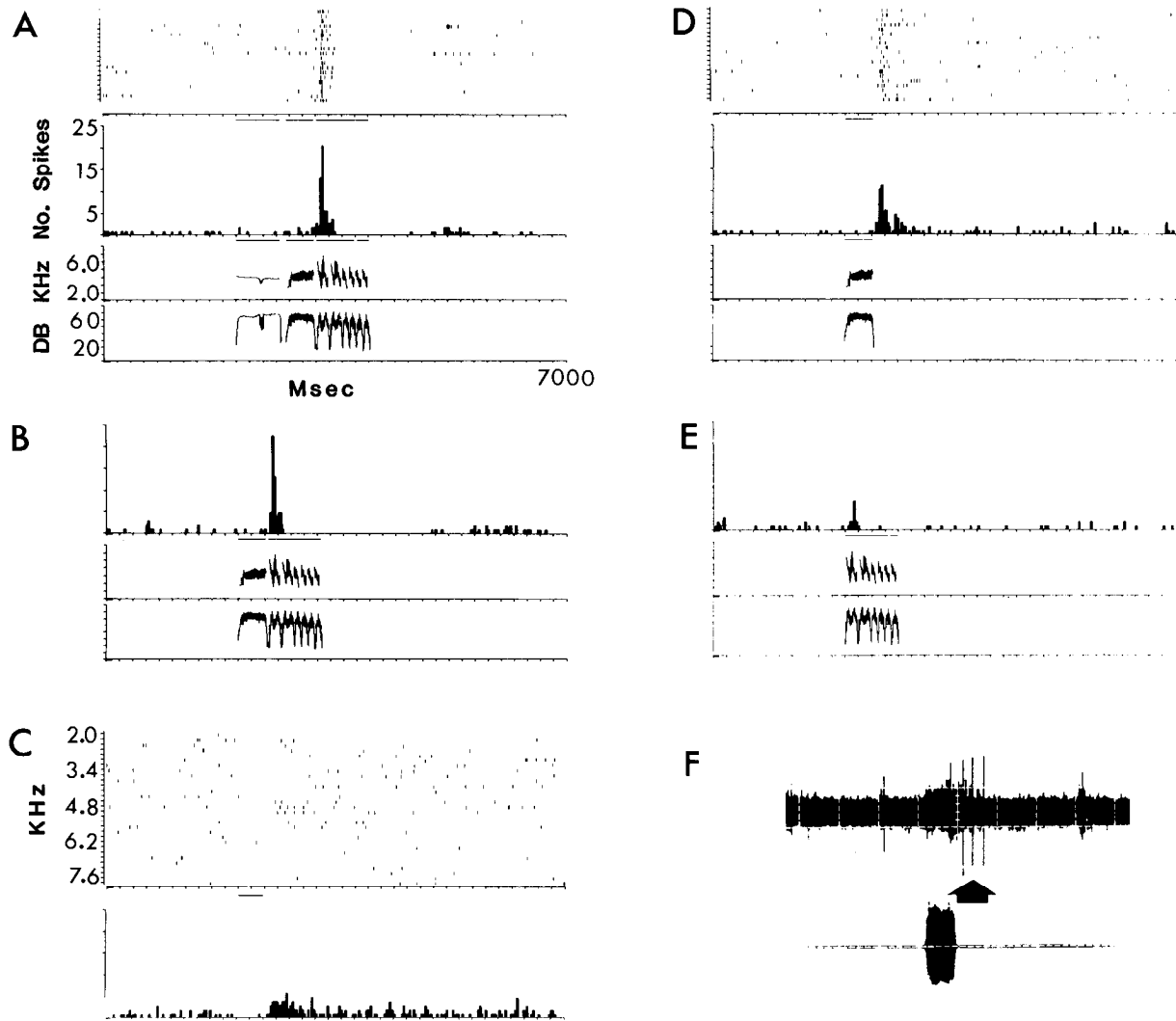
Figure 12. Response of buzz-trill unit 810402. *A,* Response to bird's own song. *B,* Response to buzz-trill combination. *C,* Response to single-tone bursts. *Top raster,* 2.0 kHz to 7.8 kHz in 200-Hz steps, covering the range of frequencies of buzz and trill. Note weak response at offset of 4.8- and 5.0- kHz tone bursts. Frequency range of 500 Hz to 10 kHz in 500-Hz steps was also tested; no other response (not shown). *Bottom histogram,* 20 repetitions of a 4.8-kHz tone burst. Note weak response at offset. *D,* Response to isolated buzz. Note response at offset is weaker and exhibits greater temporal variation than buzz-trill combination (*A* or *B*). *E,* Very weak response to trill. *A* to *E,* time axes 7000 msec; all histograms 20 repetitions, 30-msec binwidth. *F,* Actual recording showing isolation; response (*upper trace*) is to buzz (*lower trace*).

only 27 song-specific units were isolated out of 239 penetrations, this procedure is unlikely to have introduced a significant bias in the number of lesions located along the dorsal perimeter of the HVc complex. On the other hand, other lesions were often placed at the end of a penetration, a procedure clearly biasing locations toward the ventral border of the HVc complex. Thus, the distribution of the tone burst unit lesions, together with location data for the rest of the tone burst responding units, does not at this time demonstrate any clear organization within the HVc complex. The observed clustering of the song-specific units in the other two dimensions also cannot be assumed significant, because the distribution of penetrations for these experiments is skewed toward the medioposterior aspect of the HVc complex. Thus, the relationship between the locations of song-specific units and the anteroposterior organization of HVc recently

elucidated on the basis of anatomical techniques (Nottebohm et. al., 1982) cannot presently be assessed.

The presumed source of auditory input to HVc, the "shelf" area, exists in canaries as a parvocellular strip contiguous with the ventral border of HVc (Kelly and Nottebohm, 1979). A similar morphology occurs in white crowns (personal observations; Fig. 15*A*). Song-specific units concentrate in the layer between HVc and the ventricle, and/or the dorsal perimeter of HVc, and not in the shelf area. This dorsal area has not been implicated as a recipient zone for field L (auditory telencephalon) projections. Thus, it is reasonable to propose that the input to song-specific units may come via the shelf area and, therefore, that circuits local to the HVc complex are responsible for the integrative properties of song-specific units.

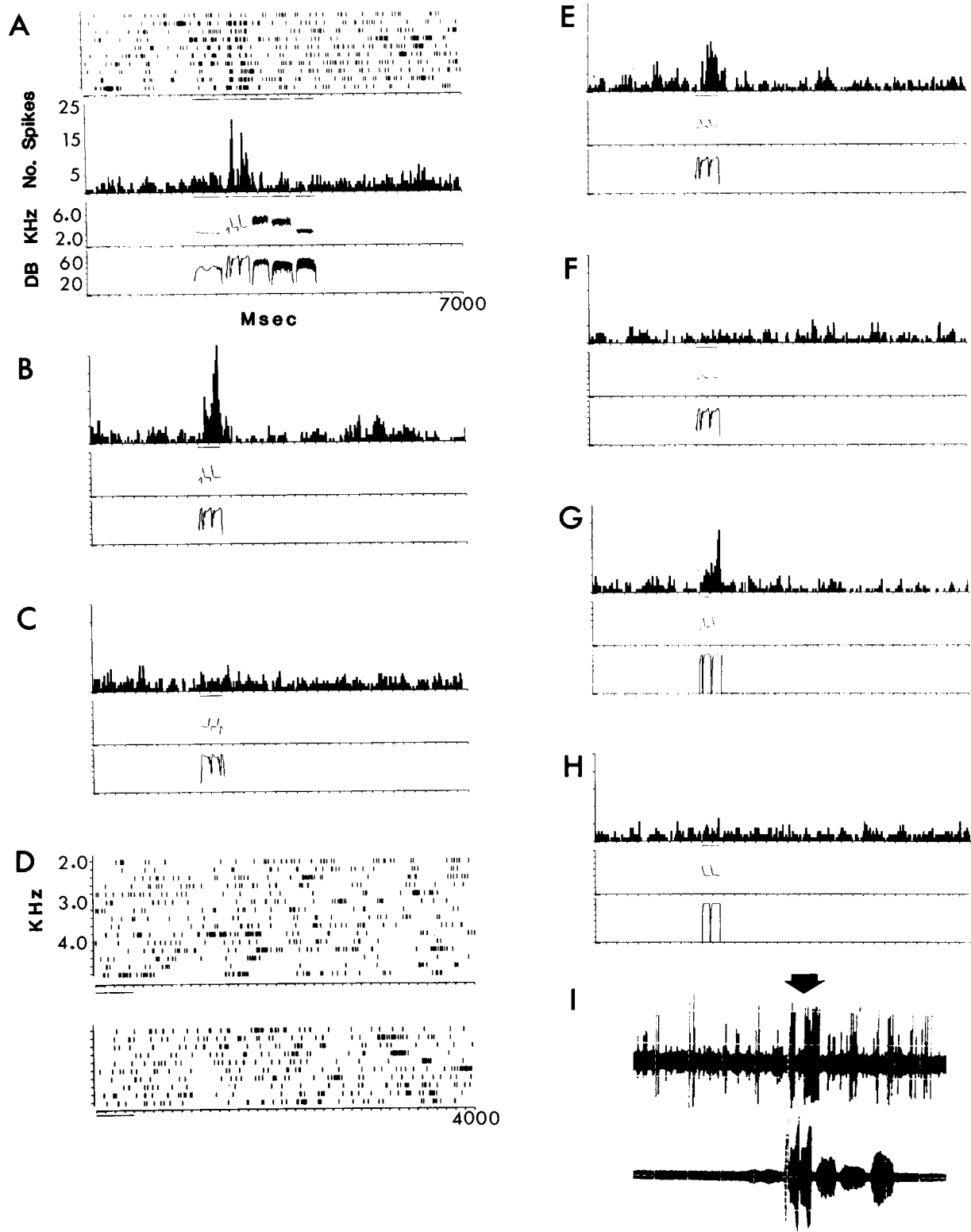*Other response categories.* The majority of single units

Figure 13. Response of trill unit 570401. *A,* Response to bird's own song. *B,* Response to isolated trill. *C,* Response is abolished by reversing trill. *D, Top raster:* individual tone bursts 2.0 kHz to 4.8 kHz in 200-Hz steps. Frequency range 500 Hz to 10 kHz in 500-Hz steps was also tested; no response (not shown). *Bottom raster:* 10 repetitions of wide-band noise burst. *E,* Trill frequencies slightly smoothed (low-pass filtered); response slightly decreased. *F,* Trill frequencies substantially smoother; response abolished. Note in *E* and *F* amplitude waveform is unaltered. *G,* Only effective model devised consists of three complex frequency and amplitude elements. Note response is not quite as vigorous as to natural trill. *H,* Response abolished by eliminating first element of model. *A* to *C* and *E* to *H,* time axes 7000 msec; all histograms 10 repetitions, 20-msec binwidth. *D* time axis is 4000 msec. *I,* Response of unit (*top trace, arrow*) to bird's own song (*bottom trace*). A second unit is present in the recording but is of sufficiently smaller amplitude as to not contaminate the recording.
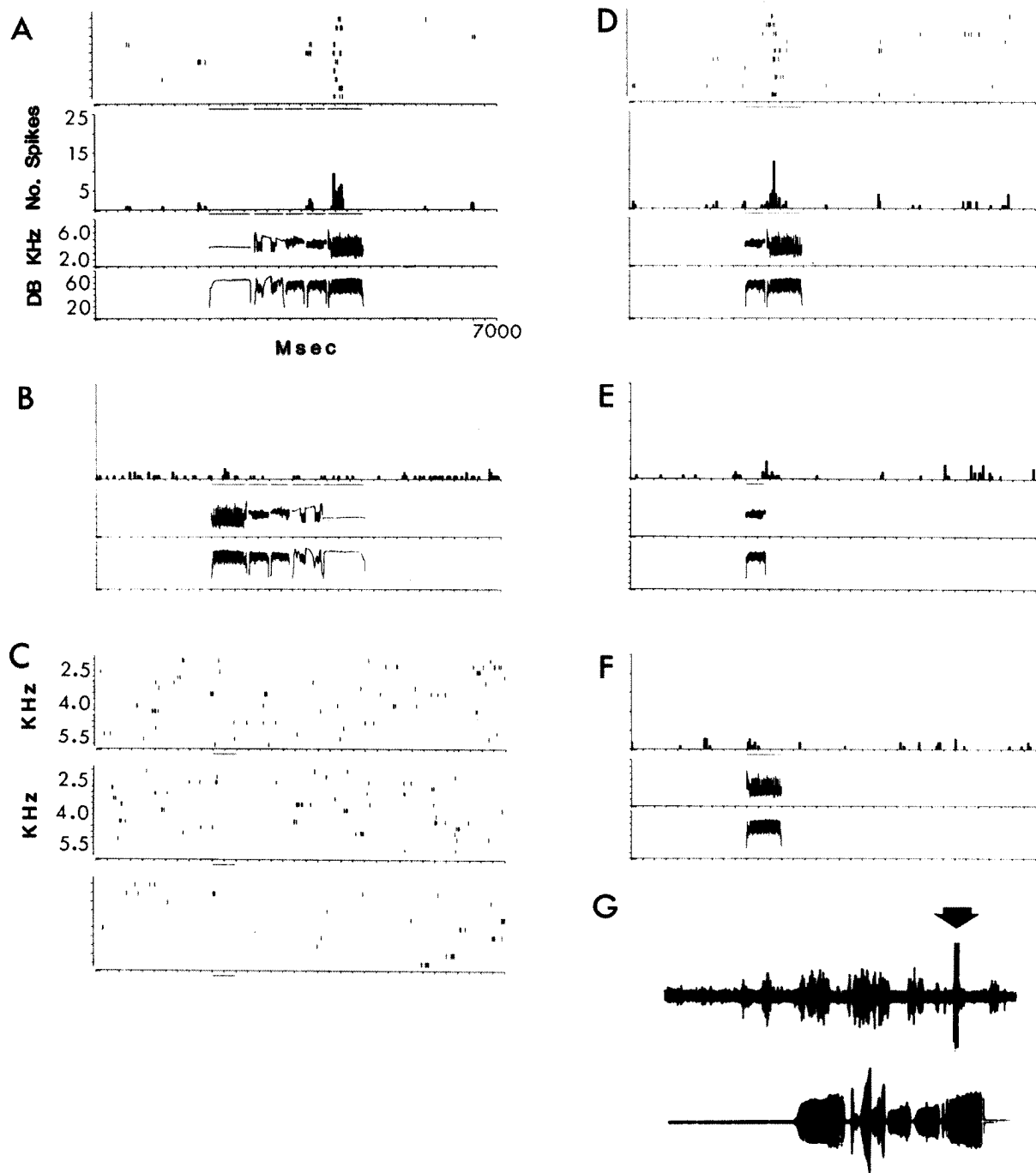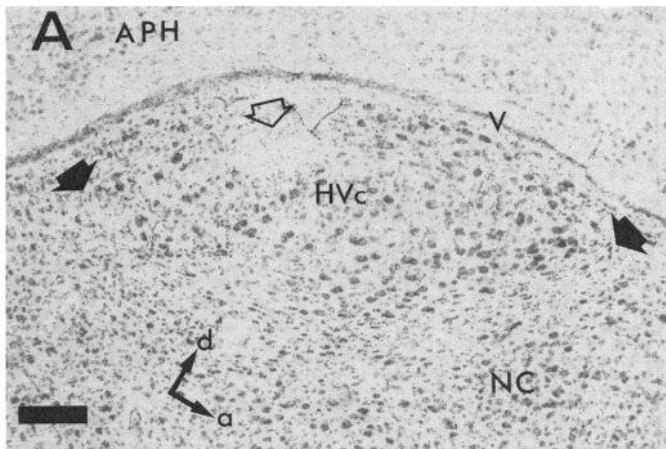
Figure 14. Response of buzz-buzz unit 600301. *A,* Response to bird's own song. *B,* Reversed song abolished the response. *C,* *Top, middle, bottom rasters:* tone bursts, narrow-band (600 Hz) noise bursts, wide-band noise bursts, respectively—no response; 2.0 kHz to 5.75 kHz in 250-Hz steps, covering frequency range of last two buzzes of the song. Frequency range of 500 Hz to 10 kHz in 500-Hz steps for tone bursts was also tested—no response (not shown). *D,* Last two phrases in isolation. *E,* Fourth phrase in isolation. *F,* Fifth phrase in isolation. *A* to *F,* time axes 7000 msec; all histograms 10 repetitions, 30-msec binwidth. *G,* Good single unit response *(arrow)* to bird's own song. This trace was chosen to emphasize the strong multiunit activity that song elicits.

isolated in HVc showed responses similar to those reported in an earlier intracellular study. These units either did not respond whatsoever to any stimulus tested or exhibited sufficiently weak onset, sustained, offset, or onset/offset responses to tone and noise bursts that many repetitions were required to clearly demonstrate the response (Katz and Gurney, 1981). In the present study a number of units responded with short latency and brisk onset responses to wide-band noise bursts or to tone bursts of a limited frequency range. The best frequencies of these units varied from 1.1 to 9.5 kHz. The distribution of latencies to response for some of these units is shown in Figure 16, and a particularly clear example of the response is shown in Figure 17*A*. No latencies shorter than 18 msec were encountered. The shortest latencies observed in field L were 8 msec (Leppelsack, 1974), and in field GA of the frontal telencephalon a recent study demonstrated units with as short a latency as 14.5 msec
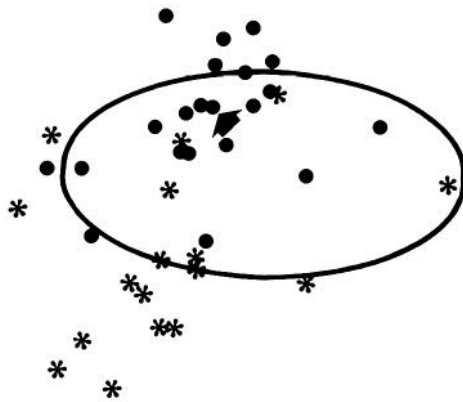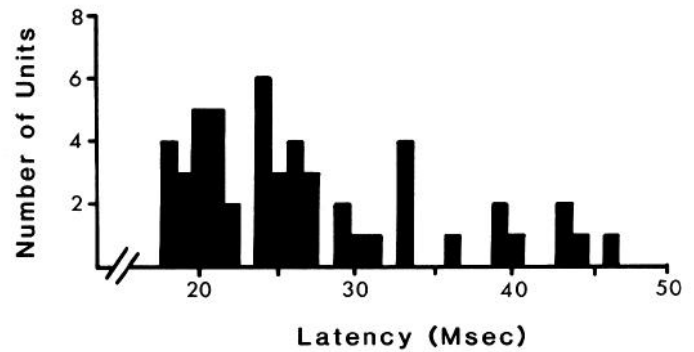
Figure 16. Latency of HVc units. The minimal time to response for short latency HVc units ($n = 51$). Stimulus was optimal one, either wide-band noise burst or tone burst, at 65 dB. Shortest latency observed was 18 msec.


Figure 15. Distribution of song-specific and tone burst units within the HVc complex. A, A cresyl violet-stained, 30-μm thick section showing the HVc complex. HVc is the magnocellular, elliptically shaped nucleus in the center. *Solid arrows* delimit the anteroposterior extent of HVc and point to the "shelf," a thin parvocellular strip contiguous with the ventral border of HVc. A relatively cell-free layer interposes between the dorsal surface of HVc and the lateral ventricle. The *open arrow* points to an electrolytic lesion of the buzz-buzz unit 600301 of Figure 14. *a,* anterior; *APH,* area parahippocampus; *d,* dorsal; *HVc,* hyperstriatum ventrale, pars caudale; *NC,* neostriatum caudale; *V,* lateral ventricle. *Calibration bar* = 100 μm. *B,* Ellipse onto which lesion data have been mapped (see the text). Orientation as in *A. Circles* are sites of song-specific units; *stars* are sites of tone burst responding units. No lesion site was dorsal to the lateral ventricle. *Arrow* indicates location on map of lesion shown in *A.*

(Kirsch et al., 1980). These data suggest the existence of multiple parallel auditory pathways in the avian telencephalon. Some units responded with brisk sustained activity throughout a noise or tone burst (e.g., Fig. 17*B*). A few units exhibited rather complex responses to tone bursts. For example, the unit of Figure 17*C* responded to tone bursts only of long duration, greater than 500 msec. The response latency was similar at threshold (30 dB) and at 35 dB above threshold. The unit was tuned rather sharply to 3.8 kHz, the center frequency of the 700-msec duration whistle in the bird's own song. On the basis of these response characteristics, this unit responded vigorously to the bird's own song and responded poorly or
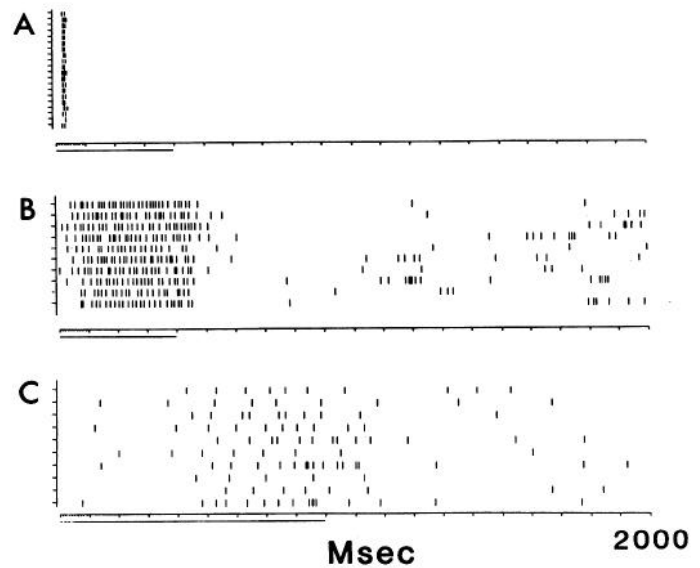
Figure 17. Tone burst units in HVc. *A,* Particularly good example of short latency onset responding unit. Stimulus is wide-band noise, 65 dB. Unit also responded well to tone bursts of 7 to 9 kHz. *B,* Particularly good example of HVc unit sustained response. Stimulus as in *A. C,* Responds to 65-dB, 900-msec tone bursts of 3.8 kHz, the center frequency of the whistle in the bird's own song.

not at all to 10 other white crown songs, half of which were of the same dialect. The significance of the aforementioned response properties, seemingly intermediate in complexity between simple tone burst responding units and song-specific units, remains unclear.

Discussion

Criteria for specificity. A number of studies employing complex natural sounds as stimuli have relied on a variety of criteria in examining whether unit responses exhibited "specific," "selective," "sensitive," "feature detector," or "focal" properties. The potential significance of such a discovery is in its relationship to the presumed existence of functional sensory hierarchies that analyze behaviorally relevant stimuli, and the assumption that activation of elements at the top of such a hierarchy is an integral part of the process of perception for that specific stimulus. Thus, the response properties of a single unit

must meet strict criteria for functional integration and stimulus selectivity if it is to be considered stimulus specific. In those systems such as echolocation in bats, where natural vocalizations are sufficiently simple that the attendant response requirements for acoustic specificity are clear, convincing examples of function-specific complex units have been reported (O'Neill and Suga, 1982). In other systems, the results to date have not been entirely satisfactory.

In the current study, the following criteria were used as the basis for deciding that a unit's response was song specific.

1. Simple stimuli such as single tone bursts, and narrow-band and wide-band noise bursts, did not elicit responses that could be used to explain the responses to complex stimuli. This included inhibitory as well as excitatory effects.

2. The unit only responded to a subset within a repertoire of stimuli chosen on the basis of observations of naturally occurring behaviors. In every case, this subset of songs included the bird's own song. It was observed that this limited set of stimuli shared parameters that were missing in the ineffective stimuli of the repertoire.

3. Artificial stimuli synthesized on the basis of these shared parameters elicited the response. The systematic modification of the parameters of the artificial stimuli delineated the response properties and verified the acoustic parameters responsible for the efficacy of the various natural stimuli.

The first criterion attempts to distinguish between response properties that reflect peripheral frequency analysis and those that are dependent on central integrative processes. Even simple peripheral response properties can generate seemingly complex responses to complex stimuli. With the addition of inhibitory side-bands, complex responses of primary-like units can easily be misinterpreted as reflecting integrative mechanisms (e.g., see discussion of forwards/backwards paradigm). Furthermore, perhaps the only units with responses to complex sounds that can be fully predicted on the basis of simple tone burst tuning are those with primary-like response properties. Two-tone excitatory/inhibitory interactions, and excitation to narrow-band or broad-band noise, are relatively simple mechanisms that nevertheless can generate significant response complexity not predictable simply on the basis of single tone burst tuning. Thus, in the absence of compelling observations to the contrary, it is essential to carefully search with tone bursts, including two-tone stimulation, as well as with narrow-band and wide-band noise bursts, for simpler explanations for seemingly complex responses.

The second criterion is commonly accepted and applied by most researchers. Knowledge of the behavioral context of various vocalizations affords powerful insights into the potential significance of a unit's stimulus selectivity. For example, in this study the abundance of passerine neighbor-stranger discrimination studies (Verner and Milligan, 1971; Emlen, 1972; Brooks and Falls, 1975; Baker et al., 1981), and the well defined existence, if not role, of song dialects, has focused attention on the intra-dialect selectivity exhibited by some units. However, since the observed selectivity of a neuron reflects in part the repertoire chosen, this criterion applied in isolation

results in an inadequate description of neuronal response properties.

Few studies have employed realistic but artificial sounds that mimic complex natural stimuli (criterion 3). This may, in part, reflect a technical limitation; nevertheless, the consequences are serious. For instance, it has been reported that in the starling field L units exhibited different levels of selectivity within a large repertoire of 80 naturally occurring call and song elements (Leppelsack and Vogt, 1976). Some units responded to many vocalizations; one extreme example responded to only one sound out of the entire repertoire and did not respond whatsoever to simple stimuli—tone and wide-band noise bursts. A very large repertoire of sounds with overlapping acoustic parameters limits the possibility that any observed response selectivity is capriciously generated by relatively simple parameters. Nevertheless, the use of synthetic yet effective sounds, modified parameter by parameter, could establish this conclusion convincingly by delineating precisely the parameters underlying the response selectivity.

In the present study, the aforementioned response property criteria were met by song-specific units. When less rigorous requirements are applied, the results can be ambiguous. For example, Scheich et al. (1979) described 58 units that selected within a group and between groups of species-specific bird calls. Although approximately half of those units responded to nine or more out of 16 guinea fowl Iambus calls, few exhibited any response to the other guinea fowl calls comprising the repertoire. Forty-eight of these units responded to tone bursts, a high proportion responding to 1 kHz and to 1.8 to 1.9 kHz, which are prominent bands in the Iambi spectra. Response versus tone burst frequency histograms typically contained one or two major peaks, several or many minor peaks, and intervening inhibitory regions. Thus, a simple relationship between the spectral energy distribution of Iambi and the neuronal frequency tuning may account for many of the observed responses. Furthermore, it was concluded that for many units the observed selectivity was based on spectral sensitivity to "frequency complexes" that define "focal properties" of Iambi. Whether this sensitivity was actually based on the entirety of the spectra or on simple combinations of spectral components cannot be ascertained because two-tone stimulation studies were not undertaken.

The great majority of units in the HVc complex responded to complex stimuli such as songs in a manner predictable on the basis of the excitatory and inhibitory responses to tones and narrow-band and wide-band noise bursts. Similar results have been reported in most studies of telencephalic and midbrain auditory nuclei in birds and mammals (Suga, 1965; Whitfield and Evans, 1965; Abeles and Goldstein, 1972; Watanabe and Katsuki, 1974; Leppelsack and Vogt, 1976; Bonke et al., 1979; Scheich et al., 1979; Symmes et al., 1980; Langner et al., 1981).

In contrast to the conclusions of these authors stand the conclusions of Newman and Wollberg (1973), who investigated the functional properties of the auditory cortex of the squirrel monkey. They reported that most of the units responded to tone bursts and 89% responded to over half of 12 different vocalizations that constituted the chosen repertoire. They concluded that neither the

tone burst tuning nor the parameters in the various call groups could explain the majority of unit responses. One explanation for this discrepancy is that the functional properties of the squirrel monkey auditory cortex are indeed different from what has been delineated for all other animals tested to date. However, the lack of artificial and modifiable versions of each call and the incomplete testing with tone and narrow-band and wide-band noise bursts of various amplitudes, as well as the lack of two-tone stimulation studies, severely limit the possible interpretation of those results.

For many multiunit and single unit recordings and for most song-specific units, forward but not reversed song was an effective stimulus. The limitations in the interpretation of this manipulation must be stressed. Such discriminations can potentially be accounted for by the almost ubiquitous occurrence in sensory systems of neurons with antagonistic center-surround receptive fields. A complex stimulus moving in time from an excitatory center into an inhibitory surround of a unit could elicit an entirely different response if reversed (Suga, 1965). Furthermore, for multiunit recordings (McCasland and Konishi, 1981), it is difficult to distinguish between increased activation of the same population of units and a change in the composition of the population of units responding to different stimuli. Nevertheless, these and other data (McCasland and Konishi, 1981) suggest that in the HVc complex, the forward/reversed song paradigm is suggestive, but not demonstrative, of complex stimulus processing.

Possible physiological mechanisms of the responses. The present data afford some insight into the neural circuit generating the selectivity of song-specific units. This analysis concentrates on the simplest of these units, the whistle-whistle units. It is assumed that all song-specific whistle-whistle units derive their response properties from the same underlying neural circuit. The observations pertaining to the responses of song-specific whistle-whistle units, that is, units that respond to a tone burst of frequency F2 if and only if it is preceded by another tone burst of frequency F1, can be summarized as follows:

1. Maximal response is elicited for frequencies F1 and F2. The consecutive tone tuning for these frequencies is reasonably sharp and peaked, typically with inhibitory side bands.

2. F2 can equal F1.

3. Simultaneous two-tone stimulation inhibits the response to consecutive tones, even if the additional tone burst is of frequency F1 or F2.

4. The timing information—the length of the tone bursts, or the interval duration—can vary over hundreds of milliseconds with only slight variation in response strength.

5. The first tone often is inhibitory.

Numerous two- or multi-cell models can account for the responses of song-specific whistle-whistle units. Single cell models fail either in that they cannot account for the timing or, if constrained to incorporate the timing requirements, do not have the necessary frequency requirements. A simple circuit that meets all the criteria is the two cell model of Figure 18, the second cell representing a song-specific whistle-whistle unit. The first cell
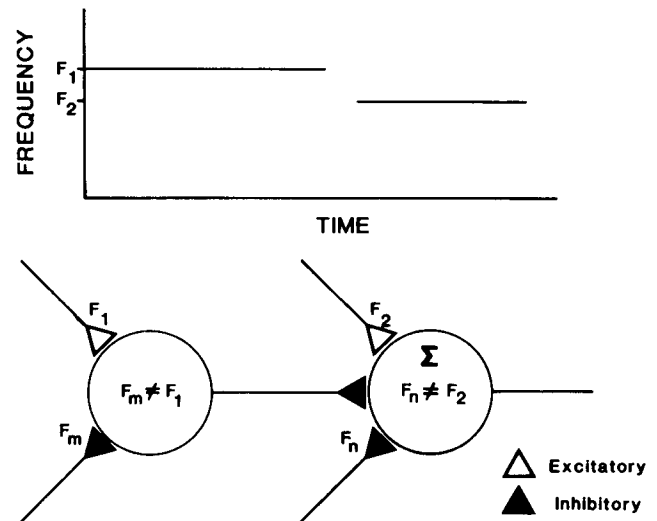

Figure 18. Simple neural circuit to account for the response properties of song-specific whistle-whistle units. The stimulus model is represented by *upper* graph. *Lower* schematic is of the excitatory and inhibitory input to two cells (*circles*). The second cell responds if rebound excitation from release from tonic inhibition by the first cell is synchronized with activation of excitatory input to the second cell (see the text).

is tuned to F1 and is inhibited by tone bursts of frequencies other than F1. The output of the first cell inhibits the second cell for the duration of the first tone. At the offset of the first tone, a rebound from release from inhibition in the second cell produces a perithreshold depolarization. The second tone elicits a subthreshold depolarization in the second cell. The summed currents elicit a suprathreshold response.

All of the response properties of song-specific whistle-whistle units can be accounted for by the proposed model. The leading consecutive tone tuning reflects the frequency tuning of the first cell in the model, and the following consecutive tone tuning reflects the tuning of the second cell. Even when F1 = F2, a single tone burst elicits only weak excitation at the onset (see Fig. 4) because excitation at the second cell is rapidly followed by inhibition from the first cell. When F1 ≠ F2, two tone stimulation, even if of frequencies F1 and F2, elicits inhibition in either the first cell or the second cell and thus abolishes the response. The first tone in isolation activates an inhibitory input to the second cell. Perhaps the weak offset excitation to initial phrases or tone bursts seen in several recordings reflects the second cell's rebound from inhibition at the offset of the sound. If each cell in the circuit has elaborate frequency requirements, the more complicated responses seen in units such as the whistle-buzz and buzz-buzz units, may possibly be modeled in the same general way.

Clear and convincing examples of temporal facilitation have already been reported at the level of single units—delay-sensitive neurons of the bat auditory cortex (O'Neill and Suga, 1979, 1982; Sullivan, 1982). Those units also respond strongly to specific temporal combinations of sound elements that in isolation elicit weak or no response. Other aspects of their response, however, such as sharp temporal tuning, lack of two tone inhibition, and response to high repetition rates, are different

from song-specific units, potentially reflecting the differing behavioral requirements of bats and songbirds. It will be interesting to see if the current model has general applicability. The proposed model has several other implications. Intracellular recordings from the second cell proposed by the model would establish whether the functional integration seen in these recordings resides at the level of the HVc complex or indeed reflects complex processing of sounds that occurs in field L (Leppelsack and Vogt, 1976; Scheich et al., 1979; Langner et al., 1981). The model makes specific predictions about the synaptic events expected in an intracellular recording of the second cell—inhibition to frequencies other than F2, e.g., F1; excitation to F2; and excitation at the offset of the first tone. Finally, if these cells are indeed modified by the song learning and memory experience, then song learning should involve not only changes at synapses (Hebb, 1949), but also the establishment of multineuronal circuit properties.

Behavioral significance of the responses. It is tempting to conclude that the song-specific responses reported here reflect some aspect of song learning. However, in the current study a significant sampling bias was introduced by searching with one particular song, the bird's own song. In the extreme case, it is possible that the HVc complex contains a galaxy of highly complex and specific unit responses and that the search paradigm merely selected for those few units with parameter requirements fortuitously satisfied by a specific song. A number of observations argue against this interpretation and indicate that at least song-specific HVc auditory responses are modified during song ontogeny.

Because auditory feedback is known to be neccessary for normal motor (song) development (Konishi, 1965), it is significant that song-specific auditory units are found in an area with motor activity specifically related to song production (Nottebohm, 1976; McCasland and Konishi, 1981). In this study, song-specific units were recorded from white-crowned sparrows of the following groups: wild-caught Gambelii and Nuttalli, laboratory-reared birds that successfully copied computer tutor songs, and birds that sang isolate-like songs. The degree of specificity exhibited by song-specific units would require a large parameter space to be embodied at the single unit level by the HVc complex of each bird if the system were innately specified. The two response classes accounting for the majority of the recordings in the HVc complex responded weakly to most sounds or did not respond at all. Song-specific units, however, were typically inhibited by stimuli other than song and thus represented a minority population. Furthermore, some laboratory-reared birds were successfully tutored with songs containing parameters never found in the wild (e.g., reversed trill—see Konishi, 1978). Isolate-like songs also did not exhibit wild-type morphologies. Nevertheless, song-specific units in these birds reflected the abnormal songs (Figs. 4, 10, and 11). When multiunit recordings responded vigorously to song, the bird's own song and allied songs always were the effective stimuli. Distantly related songs were never as effective. Finally, although many units were tested with reversed song, a complex stimulus, only a few units responded preferentially to reversed as opposed to forward song. If the response properties of complex units in

HVc were specified independently of song, one would predict reversed song to be as effective a stimulus as forward song.

These observations cannot be construed as definitive proof, but they are highly suggestive of a role for song ontogeny in modifying the responses of units of the HVc complex. If this contention is accepted, it is interesting that for those birds that failed to copy their tutor song(s) and sang either variations of the tutor song(s) or isolate-like songs, no units specific for tutor song were found. For these experiments tutor song(s) were also used as a search stimulus, although the relative number of presentations of tutor song(s) versus the bird's own song was not controlled. If this initial observation is further validated, it suggests experiments that could delineate the ontogenetic processes responsible for the song of an isolate bird, a behavioral phenomena for which no current explanation exists. Either the failure to accept a tutor model during the critical period produces isolate song-specific units that then direct the song production process, or song-specific units undergo a second round of plasticity at or after the time of song crystallization.

Concluding remarks. The existence of neurons with highly selective response properties that clearly reflect a significant degree of central integration strongly supports suggestions (e.g., Barlow, 1972; Bullock, 1983) that at least some sensory pathways are organized in a clear functional hierarchy, with elements at the top of the hierarchy specialized for specific tasks. However, even the most complex responses observed in this study required appropriate parameters of only some parts of the song—two phrases, or several elements within a single phrase; and the songs of various birds that contained the appropriate phrase(s) elicited a response. No unit in the HVc complex responded at the end of the song and only if all the song phrases were the appropriate ones. Thus, whichever behavioral task these units are involved in, the present data require the sequential and possibly parallel activation of different populations of neurons to encode the entire song. It is possible that the apex of sensory processing for birdsong at the level of single neurons would be only one stage of integration beyond that described in this study. There is no logic available to predict whether further levels of integration do exist, but the well characterized projections of HVc neurons suggests area X as a candidate site.

Many investigations of complex stimulus discrimination in sensory systems have been greatly aided by the insight provided from the neuroethological approach. In this study, the use of song as a search stimulus has uncovered responses that otherwise might have escaped notice. The significance of these responses has been assessed in relation to the natural history, critical period, crystallization, and other behavioral observations pertaining to bird song. One of the serious limitations of "pure" neurophysiology is in the interpretation of results in terms of their function in the context of animal behavior. However, the suggestion that the song-specific units described here reflect some aspect of the song-learning experience can be directly tested by recording from young birds that have yet to sing but are past the critical period for acquisition of an appropriate model. It may be possible to directly assess the relationship between song-

specific units and the hypothesized neural template. The neural correlates of acoustic imprinting seen in birds may reside at the level of single cells. Perhaps some of the neural mechanisms underlying human vocal imprinting are similar.

References

Abeles, M., and M. H. Goldstein (1972) Responses of single units in the primary auditory cortex of the cat to tones and tone pairs. Brain Res. *42:* 337-352.

Baker, M. C., D. B. Thompson, and G. L. Sherman (1981) Neighbor/stranger song discrimination in white-crowned sparrows. Condor *83:* 265-267.

Barlow, H. B. (1972) Single units and sensation: A neuron doctrine for perceptual psychology? Perception *1:* 371-394.

Bonke, D., H. Scheich, and G. Langner (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. I. Tonotopy and functional zones of field L. J. Comp. Physiol. *132:* 243-255.

Brooks, R. J., and J. B. Falls (1975) Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition. Can. J. Zool. *53:* 1749-1761.

Bullock, T. H. (1983) Perspectives on the neuroethology of sensory convergences. In *Multimodal Convergences in Sensory Systems,* E. Horn, ed., Academy of Science and Literature of Mainz, Mainz, Germany, in press.

Emlen, S. T. (1972) An experimental analysis of the parameters of bird song eliciting species recognition. Behavior *41:* 130-171.

Greenewalt, C. H. (1968) *Bird song: Acoustics and physiology,* Smithsonian Institution Press, Washington, D.C.

Hebb, D. O. (1949) *The Organization of Behavior, A Neuropsychological Theory,* John Wiley & Sons, New York.

Katz, L. C., and M. E. Gurney (1981) Auditory responses in the zebra finch's motor system for song. Brain Res. *211:* 192-197.

Kelly, D. B., and F. Nottebohm (1979) Projections of a telencephalic auditory nucleus—field L—in the canary. J. Comp. Neurol. *183:* 455-470.

Kirsch, M., R. B. Coles, and H. -J. Leppelsack (1980) Unit recordings from a new auditory area in the frontal neostriatum of the awake starling (*Sturnus vulgaris*). Exp. Brain Res. *38:* 375-380.

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

Konishi, M. (1978) Auditory environment and vocal development in birds. In *Perception and Experience,* R. D. Walk and H. L. Pick, Jr., eds., pp. 105-118, Plenum Publishing Corp., New York.

Langner, G., D. Bonke, and H. Scheich (1981) Neuronal discrimination of natural and synthetic vowels in field L of trained mynah birds. Exp. Brain Res. *43:* 11-24.

Leppelsack, H. J. (1974) Funktionelle Eigenshaften der Hörbahn im Feld L des Neostriatum Caudale des Staren (*Sturnus vulgaris* L., Aves). J. Comp. Physiol. *88:* 271-320.

Leppelsack, H. -J. (1978) Unit responses to species-specific sounds in the auditory forebrain center of birds. Fed. Proc. *37:* 2236-2241.

Leppelsack, H. -J., and M. Vogt (1976) Responses of auditory neurons in the forebrain of a songbird during stimulation with species-specific sounds. J. Comp. Physiol. *107:* 263-274.

Marler, P. (1970) A comparative approach to vocal learning: Song development in white-crowned sparrows. J. Comp. Physiol. Psychol. *71 (2), Pt. 2:* 1-25.

Marler, P., and M. Tamura (1962) Song dialects in three populations of white-crowned sparrows. Condor *64:* 368-377.

McCasland, J., and M. Konishi (1981) Interaction between auditory and motor activities in an avian song control nucleus. Proc. Natl. Acad. Sci. U. S. A. *78:* 7815-7819.

Newman, J. D., and Z. Wollberg (1973) Multiple coding of species-specific vocalizations in the auditory cortex of squirrel monkeys. Brain Res. *54:* 287-304.

Nottebohm, F., T. M. Stokes, and C. M. Leonard (1976) Central control of song in the canary, *Serinus canarius.* J. Comp. Neurol. *165:* 457-486.

Nottebohm, F. E., D. B. Kelly, and J. A. Paton (1982) Connections of vocal control nuclei in the canary telencephalon. J. Comp. Neurol. *207:* 344-357.

O'Neill, W. E., and N. Suga (1979) Target range-sensitive neurons in the auditory cortex of the mustache bat. Science *203:* 69-73.

O'Neill, W. E. and N. Suga (1982) Encoding of target range and its representation in the auditory cortex of the mustached bat. J. Neurosci. *2:* 17-31.

Scheich, H., G. Langner, and R. Koch (1977) Coding of narrowband and wide-band vocalizations in the auditory midbrain nucleus (MLD) of the guinea fowl (*Numida meleagris*). J. Comp. Physiol. *117:* 245-265.

Scheich, H., G. Langner, and D. Bonke (1979) Responsiveness of units in the auditory neostriatum of the guinea fowl (*Numida meleagris*) to species-specific calls and synthetic stimuli. II. Discrimination of Iambus-like calls. J. Comp. Physiol. *132:* 257-276.

Staddon, J. E. R., L. W. McGeorge, R. A. Bruce, and F. F. Klein (1978) A simple method for rapid analysis of animal sounds. Z. Tierpsychol. *48:* 306-330.

Suga, N. (1965) Analysis of frequency-modulated sounds by auditory neurons of echo-locating bats. J. Physiol. (Lond.) *179:* 26-53.

Sullivan, W. E., III (1982) Neural representation of target distance in auditory cortex of the echolocating bat *Myotis lucifugus.* J. Neurophysiol. *48:* 1011-1032.

Symmes, D., G. E. Alexander, and J. D. Newman (1980) Neural processing of vocalizations and artificial stimuli in the medial geniculate body of squirrel monkey. Hear. Res. *3:* 133-146.

Thorpe, W. H. (1961) *Bird song: The biology of vocal communication and expression in birds,* Cambridge University Press, Cambridge, England.

Verner, J., and M. M. Milligan (1971) Inter-population song dialect discrimination in the white-crowned sparrow. Condor *73:* 208-213.

Watanabe, T., and Y. Katsuki (1974) Response patterns of single auditory neurons of the cat to species-specific vocalization. Jpn. J. Physiol. *24:* 135-155.

Whitfield, I. C., and F. E. Evans (1965) Responses of auditory cortical neurons to stimuli of changing frequencies. J. Neurophysiol. *28:* 655-672.