Temporal Hierarchical Control of Singing in Birds

Albert C. Yu and Daniel Margoliash*

Songs of birds comprise hierarchical sets of vocal gestures. In zebra finches, songs include notes and syllables (groups of notes) delivered in fixed sequences. During singing, premotor neurons in the forebrain nucleus HVc exhibited reliable changes in activity rates whose patterns were uniquely associated with syllable identity. Neurons in the forebrain nucleus robustus archistriatalis, which receives input from the HVc, exhibited precisely timed and structured bursts of activity that were uniquely associated with note identity. Hence, units of vocal behavior are represented hierarchically in the avian forebrain. The representation of temporal sequences at each level of the hierarchy may be established by means of a decoding process involving interactions of higher level input with intrinsic local circuitry. Behavior is apparently represented by precise temporal patterning of spike trains at lower levels of the hierarchy.

The neural codes that define discrete units of episodic behavior and organize these units into temporal sequences are not well established. Vocalizations constitute a group of behaviors for which correct temporal sequencing of discrete, often stereotyped events is fundamental to proper execution (*1*). Participation of midbrain structures in the generation of simple calls is well known in both mammals and birds (*2*). Less is known about the contribution of forebrain structures, particularly in the production of more complex vocalizations such as human speech and bird songs. Here, we characterize singing-related neuronal activity in the nuclei HVc and robustus archistriatalis (RA) of the zebra finch (*Taeniopygia guttata*). We present evidence for the hierarchical organization of neural codes that corresponds to the hierarchical organization of the singing behavior.

Zebra finch songs are hierarchically organized vocalizations formed by discrete acoustic elements (syllables) separated by intervals of silence (*3*). Song syllables can be classified into distinct classes (types) on the basis of acoustic features. Each syllable, in turn, can be further divided into acoustically distinct notes. The typical zebra finch song begins with a variable number of identical, simple introductory syllables comprising one or two notes, followed by a fixed sequence (motif) of multinote syllables. The motifs are repeated in longer versions of songs and are often separated by introductory syllables or other simple "connecting" syllables.

We developed techniques to record single-unit and multiple-unit neuronal activity in the HVc and RA of singing adult male zebra finches (*4*). Multiple sites were recorded in each nucleus in several birds who were good singers, resulting in a large database of vocalizations and associated neuronal activities [94 ± 92 (mean ± SD) songs per bird, $n = 13$ birds]. The onset and offset time and the identity of each syllable and note were established manually or by an automatic technique (*5*) whose output was verified manually. This procedure was essential for veridical analysis because the exact timing of the sequence of song elements varied from song to song. Additional long records (300 s) of ongoing activity

A. C. Yu, Committee on Neurobiology, University of Chicago, Chicago, IL 60637, USA.
D. Margoliash, Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA.

*To whom correspondence should be addressed.

during behavioral quiescence were collected to quantify baseline activity.

Each HVc unit was strongly recruited starting before the first introductory syllable, with overall excitation throughout the entire duration of the song and with activity terminating before the end of the song (Fig. 1, A and B). We determined the exact pattern of activation by calculating motor activity histograms (MAHs), representing each unit's activity relative to the onset of a syllable or note type (6). For each HVc neuron, activity levels close to the maximum firing rate were found during the production of almost all syllable types. Each syllable type was associated with a stable and unique pattern of neuronal activity (Fig. 1B). The activity pattern for the same syllable type varied across HVc neurons, and for each HVc neuron, the activity pattern varied with syllable type. Temporal features of neuronal activity associated with each syllable type were sufficiently distinctive that correct inference of vocal output could be easily made in many instances by inspection of the associated MAH. Aligning neuronal activity with the onset (or offset) of the associated syllable was essential for detecting these features. The distinct features of the MAHs were lost when the onset or offset of one syllable was used

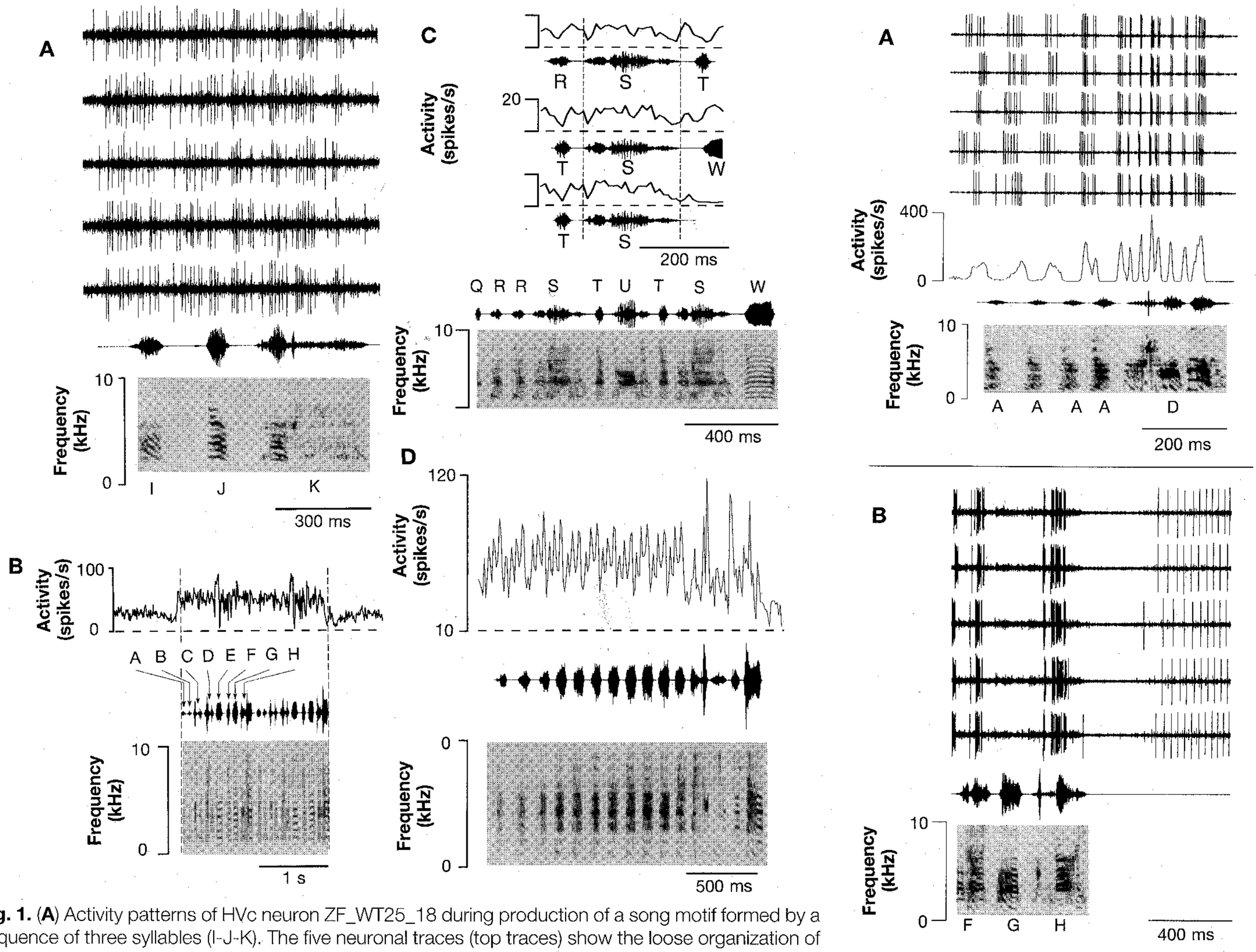

Fig. 1. (**A**) Activity patterns of HVc neuron ZF_WT25_18 during production of a song motif formed by a sequence of three syllables (I-J-K). The five neuronal traces (top traces) show the loose organization of bursts of activity preceding each syllable. Neuronal traces were aligned at the onset of syllable K (the exact timing of syllable onsets and offsets was slightly different for each neuronal trace). The oscillograph (amplitude envelope) and spectrograph (frequency representation) of an exemplar of the motif are time-aligned below the neuronal traces, and the syllable type designation is indicated below the spectrograph. (**B**) HVc activity during a song with two motifs. An extended MAH (eMAH; top trace) (*20*) of neuron ZF_YL49_4 shows the similarity of HVc activity across two song motifs (note similar peaks of activity around syllable E). The oscillograph and spectrograph of the canonical song of the bird are shown in the middle and bottom traces, respectively, and the syllable type designation is indicated above the oscillograph. Vertical dashed lines indicate the relative positions of song onset and offset. (**C**) Activity of HVc neuron ZF_GR46_2 for syllable S in different contexts (syllable sequences). The top three traces show eMAHs corresponding to three different sequences in which syllable S occurs (R-S-T, n = 137 entries; T-S-W, n = 133 entries; T-S-end of song, n = 67 entries). Vertical dashed lines indicate onset and offset of syllable S. The oscillograph and spectrograph of the canonical song are shown in the middle and bottom traces, respectively. (**D**) Similarity of HVc activity for the repeated syllables in introductory sequences of songs. An eMAH of neuron ZF_GR43_1 (top trace) is time aligned with the oscillograph and spectrograph (bottom traces) of the canonical song. The typical song of ZF_GR43 was unusual, with many introductory syllables but only one motif.

Fig. 2. (**A**) Activity patterns of RA neuron ZF_RA01_7 during a sequence of introductory syllables followed by the first syllable of the motif. Neuronal traces were aligned at the onset of syllable D. Individual MAHs that contributed to the eMAH were calculated starting 40 ms before their corresponding syllable. (**B**) Activity patterns of RA neuron ZF_RA02_5 during a motif that ends a song. Neuronal traces were aligned at the onset of syllable H. There was strong inhibition, lasting 400 to 800 ms, after song offset in all RA neurons, after which the neurons returned to their nonsinging ongoing oscillatory activity. In contrast, excitation is seen in the background during suppression of activity associated with syllable G. Other single units recorded in bird ZF_RA02 exhibited excitation during syllable G.

to calculate the MAH for even immediately preceding or following syllables. This effect resulted primarily from variation across songs in the intervals between syllables (there was much lower variation in the duration of syllables), demonstrating that the strong modulation of neuronal activity in HVc during singing was related to the timing of syllables.

In all birds, MAHs for syllables of the same type occurring in different motifs of a song were strikingly similar (Fig. 1B). To quantify this, we calculated the linear correlation coefficient r for pairs of MAHs corresponding to syllables of the same type and for pairs of MAHs corresponding to syllables of different types (7). Correlations were high comparing pairs of MAHs from the same neuron for syllables of the same type drawn from different motifs (r = 0.918 ± 0.05, n = 644 MAH pairs). By comparison, correlations were very low between pairs of MAHs from the same neuron for different syllable types drawn from the first motifs (r = 0.042 ± 0.263, n = 497 MAH pairs). The two distributions of correlation coefficients were nonoverlapping (Mann-Whitney U test, $Z = -28.997$, P = 0.0001). Additionally, several birds produced songs in which syllables of a given type occurred in two or more distinct sequences (different preceding or following syllables, or where the focal syllable ended the song) (Fig. 1C). Without fail, syllables of the same type occurring in different syllable sequences also had similar MAHs (r = 0.919 ± 0.065, n = 14 MAH pairs, four birds). The repeated introductory syllables at the beginning of a song also presented the same MAHs whether in the middle of a sequence of such syllables or as the last syllable before the first motif (r = 0.895 ± 0.112, n = 78 MAH pairs, nine birds) (Fig. 1D). These observations demonstrate that motor activity in the zebra finch HVc is centered on the syllable, is based on syllable type, and is independent of syllable context.

In contrast to the relatively tonic discharge patterns of HVc neurons, neuronal activity in the RA during singing was characterized by trains of short bursts of spikes separated by periods of profound inhibition (Fig. 2, A and B). The spike bursts associated with all introductory syllables up to the last one had imprecise and variable timing (compare Figs. 1D and 2A). Otherwise, each spike burst was characterized by a stereotyped and unique pattern of intraburst timing. The reliability of activity patterns was sufficient to allow correct inference of vocal output from individual spike trains. The RA neurons could be recruited after the onset of some syllable types, remain active after the offset of some syllable types, or exhibit complete suppression of activity even for complex syllables that formed part of a motif (Fig. 2B), phenomena never observed for HVc neurons.

During singing, each RA spike burst pattern was associated with a unique subsyllabic acoustic event. For example, two of the four birds sang a pair of syllable types that had different initial note types but thereafter shared the same sequence of note types (Fig. 3). For these four syllable types, all of the RA neurons (n = 8) exhibited similar activity patterns corresponding to the shared note types (r = 0.883 ± 0.081,

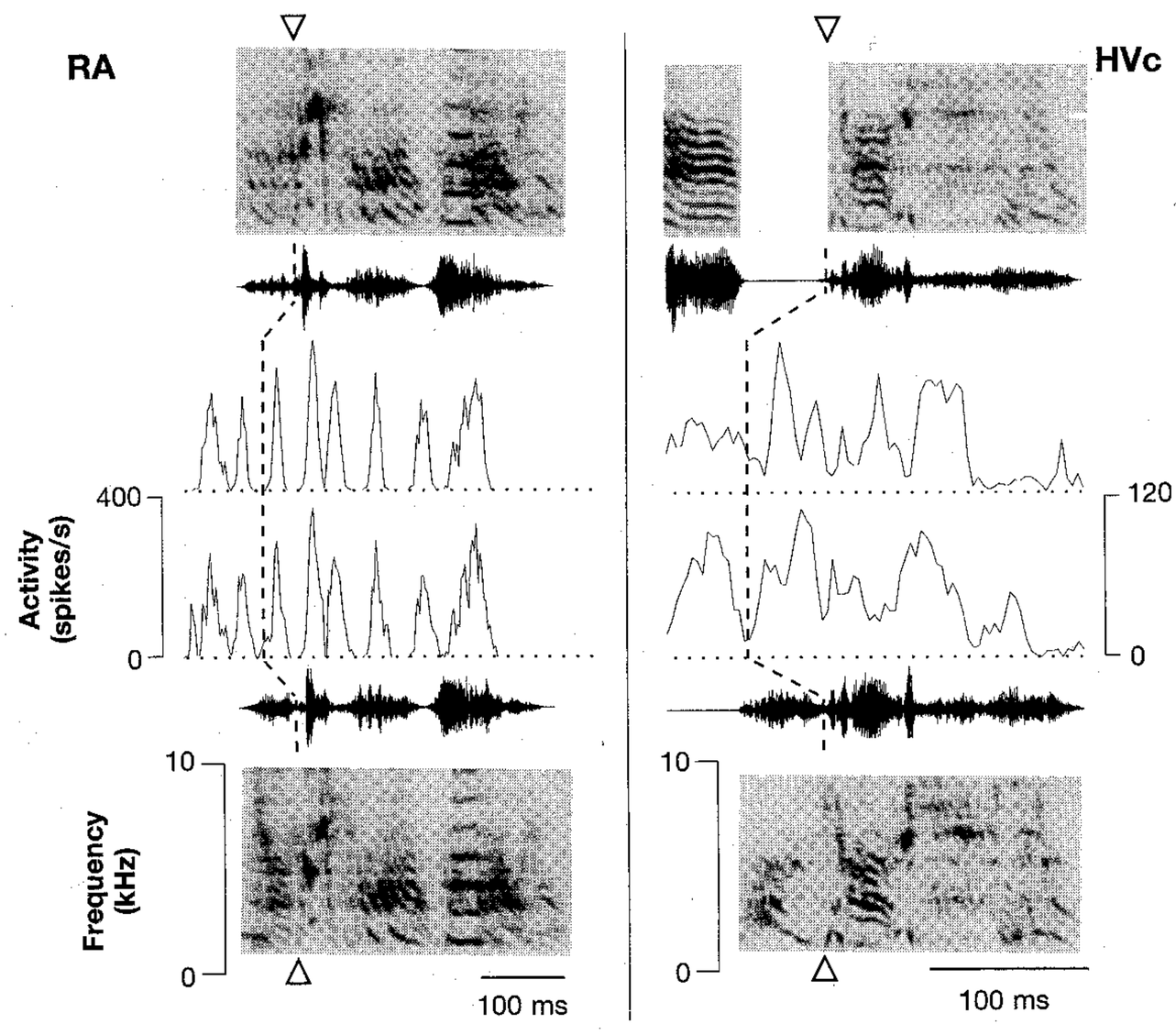

Fig. 3. MAHs for syllable pairs that start with different note types but otherwise comprise the same sequence of notes. (**Left**) Two syllables of bird ZF_RA01 share the same sequence of notes, except for different introductory notes (to the left of the open triangles). MAHs are shown for RA neuron ZF_RA01_7 for the two syllable types (top MAH, n = 91 entries; bottom MAH, n = 85). The dashed line through the MAHs marks 40 ms before the start of the first shared note. (**Right**) An equivalent analysis for HVc neuron ZF_WT25_18 for two similar syllable types. One syllable type (bottom panel, associated MAH has 70 entries) has an introductory note that is missing from the other syllable type (top panel, associated MAH has 93 entries).

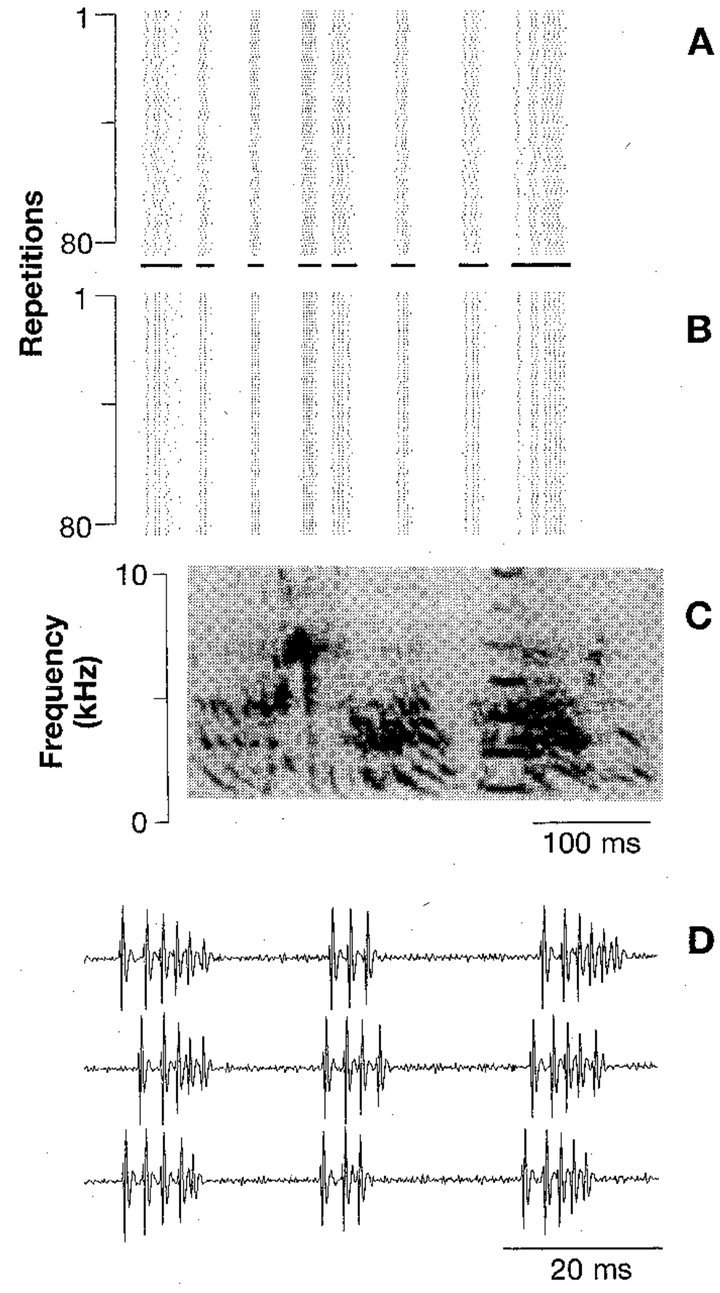

Fig. 4. (**A** and **B**) The discharge pattern of an RA neuron during 86 occurrences of the same syllable type. In (A), the time of each spike train is adjusted relative to the acoustics of the corresponding syllables by a cross-correlation technique (*9*). In (B), the same spike data are aligned on a per-burst basis, minimizing time differences (*10*). The horizontal bars below (A) indicate the time windows used to define each burst. (**C**) The spectrograph of the syllable, time-aligned with (A) and (B). (**D**) The burst pattern of another RA neuron, demonstrating reduction in spike amplitude for very fast bursts. As in this case, we commonly observed that the first interspike intervals of very fast bursts were of longer duration than the subsequent interspike intervals, which were relatively constant in duration.

$n = 8$ comparisons of MAHs corresponding to the sequence of notes of the same type) and dissimilar activity patterns corresponding to the dissimilar note types drawn from the same syllables ($r = 0.190 \pm 0.304$, $n = 8$ comparisons of MAHs) (*8*). Thus, the pattern of activity of RA neurons depends on note type. There were also two examples of birds with HVc recordings who sang a pair of syllable types that differed in the sequence of initial note types but shared the same sequence of subsequent note types. In contrast to the RA neurons, however, all HVc neurons ($n = 12$) displayed dissimilar MAHs for all segments of these similar syllable types (Fig. 3). For HVc neurons, the correlation coefficients for MAHs corresponding to the sequence of notes of the same type were low ($r = 0.401 \pm 0.158$, $n = 12$ comparisons of MAHs), much lower than those for the corresponding analysis for RA neurons but higher than those for comparisons of HVc activity patterns associated with completely different syllables. Thus, the pattern of activity of HVc neurons associated with each note depends in part on the identity (type) of the syllable in which the note is embedded.

For the fixed sequences of syllables that form a motif, the discharge patterns of RA but not HVc neurons exhibited highly reliable and precise timing at the level of individual spikes. For all RA neurons, application of an analysis procedure to improve the temporal registry of syllables (9) resulted in a visually striking alignment of bursts, exposing a consistent temporal structure within each burst, albeit with some remaining temporal jitter (Fig. 4A). Adjustment of the temporal registry of individual bursts within a syllable (*10*) revealed a remarkable precision of temporal patterning in individual bursts, often associated with a reliability of spike occurrence approaching 100% (Fig. 4B). In many cases, the temporal jitter at a given spike position within a burst was of the order of the rate at which the original analog wave forms had been sampled (50 μs per sample). Additionally, during singing, RA neurons operated through most of the dynamic range of activity available to nervous systems. Against a background of complete suppression of activity, some neurons exhibited zero or one spike for a given syllable, whereas the spike rate of the fastest burst for each neuron was 383 ± 119 spikes/s ($n = 15$ neurons), in many cases with consistent intraburst instantaneous spike rates > 700 Hz. For many neurons, we observed bursts that consistently drove the neuron into its relative refractory period, causing dramatic reduction in spike amplitude and potentially approaching spike failure (Fig. 4D). Thus, the stereotype of bird song is achieved by means of a neuronal code within the forebrain (RA) operating over a wide dynamic range that exhibits precise and reliable temporal patterning of spikes.

Singing is a motor program, representing the coordinated spatiotemporal activation of many syringeal and vocal tract muscles in conjunction with the respiratory and postural systems (*11*). Our data indicate that in the zebra finch, successively smaller units of vocalization—syllables and presumptively notes—are reliably coded in the activity patterns of single HVc and RA neurons, respectively [compare with (*12*)]. The physiological signature of a note depends on note sequence in HVc and is independent of note sequence in RA. Thus, on the basis of these data, the precise sequencing of notes apparently emerges from the interaction of HVc input with RA local circuits. Similarly, the temporal sequence of syllables may result from an interaction of afferent input to HVc (*13*) with local HVc circuitry. These data imply a hierarchical organization for the forebrain control of bird song production (*14*). A hierarchical organization has long been anticipated in the neural control of behavior (*15*).

Our results indicate that the neural code for syllables (movements) is transformed in the projection of HVc onto RA (motor control). Neurons of the RA have simple, oscillatory, ongoing discharge patterns that probably result from combinations of intrinsic properties and local circuits (*16*). The driving and coupling of simple oscillators could produce during singing the complex burst patterns that we observed. Such processes must be dynamically regulated at the rate of note production, resulting in rapid resynchronization of the neuronal population, as has been seen in other systems (*17*). Dynamic modulation of RA burst patterns by HVc input may result from regulation of the phase relations of stable groups of oscillators organized into functional units, as well as from dynamic coupling of different groups of RA neurons into functional units. A temporal structure similar to that of the fastest RA bursts has been seen in the burst patterns of inferior olivary neurons, where the first spike results from a somatic sodium spike and the timing of subsequent spikes is modulated by dendritic calcium influx (*18*). Bursting in the RA is under the control of similar cellular mechanisms (*16*). The tight regulation of discharge timing in the RA suggests that information is coded in the temporal structure of the bursts. Temporal patterns of neural activity may convey information in other systems as well (*19*). An analysis of such patterning in relation to the behavioral requirements of the animal can lead to explication of the nature and function of the putative temporal codes.

REFERENCES AND NOTES

1. K. S. Lashley, in *Cerebral Mechanisms in Behavior*, L. A. Jeffress, Ed. (Wiley, New York, 1951), pp. 112–135; M. Konishi, in *Perception*, vol. VIII of *Handbook of Sensory Physiology*, R. Held, H. W. Leibowitz, H.-L. Teuber, Eds. (Springer-Verlag, Berlin, 1978), pp. 289–309; P. Marler and S. Peters, *Science* **198**, 519 (1977); P. Marler and R. Pickert, *Anim. Behav.* **32**, 673 (1984); H. C. Gerhardt, *J. Exp. Biol.* **74**, 59 (1978); R. R. Hoy and R. C. Paul, *Science* **180**, 82 (1973).
2. D. Ploog, *Brain Res. Rev.* **3**, 35 (1981); T. J. Seller, *Trends Neurosci.* **4**, 301 (1981).
3. K. Immelmann, in *Bird Vocalizations*, R. A. Hinde, Ed. (Cambridge Univ. Press, London, 1969), pp. 61–74; R. A. Zann, *The Zebra Finch* (Oxford Univ. Press, Oxford, 1996).
4. All procedures were approved by an institutional animal care committee. Under pentobarbital and chloral hydrate anesthesia, male zebra finches were implanted with headgear, including electrodes and electronics. After recovery, a bird typically participated in 6- to 8-hour recording sessions every 2 to 3 days. We achieved stable chronic recordings in HVc by using bundles of Isonel-insulated microwires (eight recording sites, seven birds) or a custom-built 1-g mechanical microdrive to simultaneously move four Pt-Ir electrodes (15 HVc recording sites in two birds, and 22 RA recording sites in four birds). There were no systematic differences noted in the activity patterns of HVc neurons recorded under the two conditions, and the data were combined in all analyses presented here. During a recording session, the bird was attached to a custom-built commutator by a flexible cable, permitting free exploration inside the cage. In some cases we collected data while zebra finches sang spontaneously; otherwise, females or mirrors were introduced into the adjacent half-cages to stimulate directed song [R. Sossinka and J. Böhner, *Z. Tierpsychol.* **53**, 123 (1980)]. During data analysis, female calls could be distinguished from male calls and from male song syllables. After a bird was killed with an overdose of pentobarbital, the position of the recording sites within the target nuclei was confirmed with standard frozen-section histology.
5. S. E. Anderson, A. S. Dave, D. Margoliash, *J. Acoust. Soc. Am.* **100**, 1209 (1996).
6. Single units were isolated (categorized) off-line according to established procedures [M. L. Sutter and D. Margoliash, *J. Neurophysiol.* **72**, 2105 (1994); M. S. Lewicki, *Neural Comp.* **6**, 1005 (1994)]. We constructed MAHs by aligning vocalizations and the related neuronal activity using the syllable onset or offset times, then binning the times of occurrence of action potentials. There were no systematic differences between onset and offset MAHs for either the HVc or RA data; we use onset MAHs in this report. The data shown here are from 40 well-isolated single units in the HVc and 23 well-isolated single units in the RA.
7. Pairs of MAHs were adjusted to the same duration with the use of time-warping decimation-interpolation techniques implemented using routines in the Matlab program (MathWorks, Natick, MA). The linear correlation coefficient r for two MAHs representing syllables x and y was calculated as

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}$$

where x_i and y_i represent individual histogram bin values, and $\bar{x}$ and $\bar{y}$ represent average bin values.
8. The MAHs were separated into two segments: one segment associated with shared note types, and one segment associated with dissimilar note types. The segmentation assumed that neuronal activity preceded the vocalizations by 40 ms for the RA recordings and by 50 ms for the HVc recordings. These values correspond to average multiunit latency preceding singing [J. S. McCasland and M. Konishi, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7815 (1981)].
9. Spike trains were brought into temporal registry by

cross-correlating the corresponding syllables, as follows. Introductory syllables were excluded because for the RA they are not associated with a precise temporal pattern of activity (Fig. 2A). Motif syllables comprising simple harmonic stacks were also excluded because they lack meaningful time-varying frequency modulation, which resulted in unreliable cross-correlations. This left acoustically complex syllables of motifs for analysis. The acoustic records were scored without reference to the associated spike trains to eliminate recordings with acoustic clutter (background cage noises, calls of females). Six additional syllables were eliminated on this basis because less than 10 acoustically uncontaminated exemplars were identified, preventing meaningful statistical analysis. For each of the 15 resultant syllable types (15 neurons, three birds), a "referent" syllable was chosen by manual inspection of spectrographs of the set of exemplar syllables. The spectrograph of each exemplar syllable was then cross-correlated with the spectrograph of the referent [C. W. Clark, P. Marler, P. Beeman, *Ethology* **76**, 101 (1987)]. We then adjusted the temporal registry of each spike train associated with each exemplar syllable, relative to the spike train associated with the referent syllable, by applying a shift in time (translation) based on the position of the peak in the correlation function. The time shifts were typically quite small ($\tau = 3.00 \pm 2.31$ ms, $n = 1436$ cross-correlations of referent and exemplar syllables), implying that the original manual segmentation was quite accurate; nevertheless, this shift significantly affected the temporal registry of spike trains.

10. An optimal translation of the spike trains was applied on a burst-by-burst basis to minimize the global difference in spike timing. That is, this procedure aligned each spike burst independent of the acoustics of associated notes. The acoustic procedure of (*9*) failed to further improve the temporal registry of spike bursts when applied on a note-by-note basis. This failure may be the result of the fine temporal resolution of RA neuronal discharge patterns overwhelming inherent limitations in time-frequency resolution in the calculation of spectrographs based on short-time Fourier transformations [G. D. Bergland, *IEEE Spectrum* **7**, 41 (1969)]. Nevertheless, it implies that we were not able to quantitatively demonstrate that the timing of each burst pattern was associated with the timing of each note. The RA exhibits a myotopic organization [D. S. Vicario, *J. Neurobiol.* **22**, 63 (1991)], hence the activity of RA neurons may be associated with activation of individual muscles or groups of muscles.
11. S. Nowicki and R. R. Capranica, *Science* **231**, 1297 (1986); F. Goller and R. A. Suthers, *Nature* **373**, 63 (1995); M. W. Westneat, J. H. J. Long, W. Hoese, S. Nowicki, *J. Exp. Biol.* **182**, 147 (1993).
12. E. E. Fetz, in *Movement Control*, P. Cordo and S. Harnad, Eds. (Cambridge Univ. Press, Cambridge, 1994), pp. 77–88.
13. H. Williams and D. Vicario, *J. Neurobiol.* **24**, 903 (1993).
14. See also E. T. Vu, M. E. Mazurek, Y.-C. Kuo, *J. Neurosci.* **14**, 6924 (1994).
15. N. Tinbergen, *Symp. Soc. Exp. Biol.* **4**, 305 (1950); *The Study of Instinct* (Oxford Univ. Press, Oxford, 1951).
16. R. Mooney, *J. Neurosci.* **12**, 2464 (1992).
17. M. Abeles *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8616 (1995); E. Vaadia *et al.*, *Nature* **373**, 515 (1995).
18. R. Llinás, *Cold Spring Harbor Symp. Quant. Biol.* **55**, 933 (1990).
19. C. E. Carr, W. Heiligenberg, G. J. Rose, *J. Neurosci.* **6**, 107 (1986); C. E. Carr and M. Konishi, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8311 (1988); E. Ahissar *et al.*, *Science* **257**, 1412 (1992); G. H. Recanzone, M. M. Merzenich, C. E. Schreiner, *J. Neurophysiol.* **67**, 1071 (1992); J. C. Middlebrooks, A. E. Clock, L. Xu, D. M. Green, *Science* **264**, 842 (1994); R. C. de Charms and M. M. Merzenich, *Nature* **381**, 610 (1996); W. Bair and C. Koch, *Neural Comp.* **8**, 1184 (1996).
20. To construct an eMAH, we defined a canonical song: the most common number of introductory syllables, the most common sequence of syllables within a motif, and the most common number of motifs. The eMAH was derived from concatenation of individual MAHs corresponding to each syllable type within its specific context of the canonical song—for example, all spikes corresponding to syllable E in the first motif or all spikes corresponding to syllable E in the second motif (Fig. 1B). Individual MAHs were calculated starting 50 ms before their corresponding syllable.
21. J. J. Gilpin manufactured the devices for chronic recording. A. S. Dave collected the data for one of the HVc birds. We thank A. S. Dave, S. E. Anderson, and J. A. Kogan, who provided valuable advice and assistance on aspects of the data analysis. M. Konishi and P. S. Ulinski provided useful critiques of the manuscript. Supported by a grant from the Whitehall Foundation (M91-05). A.C.Y. was supported by an NIH predoctoral fellowship (1 F31 MH10151).

26 April 1996; accepted 8 August 1996

How the Songbird Makes His Song

In human organizations, those at the top tend to set strategy while leaving the details to their underlings. The same thing seems to happen in the brain—at least in the brain of the male zebra finch as it controls the bird's singing. On page 1871, Albert Yu and Daniel Margoliash of the University of Chicago provide the most detailed look yet at how the finch's brain controls the bird's singing muscles, and they find a chain of command that might be familiar to any management-school graduate.

The work shows, Margoliash says, that the singing instructions are relayed down a hierarchy of brain regions, getting progressively more detailed as they go. The firing patterns of the neurons in the higher brain centers apparently specify the more complex components of the birds' songs, the syllables, which are collections of notes, while the patterns at the lower level in this neural chain of command define the basic sounds—most likely, the notes themselves.

While neuroscientists have long suspected that the brain has such hierarchical motor programs, there was little direct evidence for the hierarchy. This report "is something that people have been looking for," comments neuroscientist Eric Vu of the Barrow Neurological Institute in Phoenix, whose own earlier work on zebra finches had hinted at a similar result. "It shows that to perform more complex behavior, the brain follows a chain of command: Higher brain centers are responsible for more abstract [information], and lower brain centers fill out the details." As they traced this hierarchy, Yu and Margoliash also found additional support for the idea that the timing of the impulses from a nerve cell, not just the number of pulses in a given period of time, carries information (*Science*, 3 November 1995, p. 756).

BENJAMIN C. YU

Singing for science. Lightweight headgear, including implanted electrodes, doesn't stop his song.

The findings may also help neuroscientists understand speech production in humans, because human speech, like bird song, is modular. Whereas a song consists of notes strung together in syllables and then phrases, human speech consists of phonemes, basic utterances such as the sounds associated with particular letters, linked together in words, which then form sentences and paragraphs. If the firing patterns of neurons in different brain regions define the finch's notes and syllables, then some similar organization may be at work in the human brain's speech production centers, Margoliash and others suggest.

Until now, the most direct evidence suggesting hierarchical control of bird song came from work done in 1994 by Vu's team. At the time, other researchers had already implicated several regions in controlling bird song. Vu, and subsequently Margoliash and Yu, studied two of them: the HVc, a cluster of nerve cells sometimes called the song production center, and the robustus archistriatalis, which relays input from the HVc down toward the base of the brain, where the RA's nerve cells interface with those that activate the right muscles for making sounds.

For his study, Vu implanted tiny electrodes in either the HVc or the RA of various zebra finches and observed how sending an

electrical pulse down the electrode while the bird was singing disrupted the song. Pulses to the RA just caused the birds to make a small mistake—perhaps dropping a note. But a pulse to the HVc stopped their singing altogether for a moment, and the birds would start that part of the song over again.

As a result, Vu concluded that the HVc contributed at a higher level, perhaps by defining syllables or sequences of syllables, while the RA conveyed instructions about the notes needed to create syllables. But to actually prove that, neuroscientists needed to see the actual firing patterns of the neurons controlling bird song. And that is what Margoliash and Yu have provided in their current work.

For their experiments, Yu and Margoliash first outfitted 13 male birds with brain electrodes, projecting into either the HVc or the RA. These electrodes were in turn connected to a computer that recorded and analyzed the electrical impulses from individual neurons in the two brain centers as the birds sang. Because Margoliash also kept track of what sounds the birds made as various cells fired, the researchers could compare the activity of these different cells at the same point in the bird's song, which it repeats over and over again.

These comparisons showed, Margoliash says, that each HVc cell measured had a signature firing pattern that differed from cell to cell but corresponded to a specific syllable. These patterns weren't apparent at first, however, because of what Margoliash calls their "sloppiness." The firing patterns were variable both in the number of bursts in a series and the time between each series. But by lining these patterns up with the corresponding sound recordings, Yu and Margoliash found that the patterns did have subtle similarities, with some characteristic pauses and a recognizable, although not identical, series of bursts for each syllable. Eventually the researchers were able to predict the syllable to be sung just by looking at these patterns.

In the RA, in contrast, the firing patterns were much easier to discern, because the timing of firing varied very little from one round of singing to the next. And in this region of the brain, the patterns seem to correspond generally to individual notes rather than syllables. "As you get closer to the muscles, [the brain] is breaking [the message] down into smaller and smaller units," explains Allison Doupe, a neurobiologist at the University of California, San Francisco. "The precision of the timing [in nerve cells] gets more and more exact." Thus by the time the signal to sing reaches the muscles, that signal has been broken into very specific commands that synchronize each muscle contraction.

Yu and Margoliash have also shown, as Doupe puts it, that there is information not only in neuroscientists' traditional focus, "the neuron turning on and off, but also in the timing of when it goes on and off." In a given RA nerve cell, for example, the researchers found that two bursts of activity with almost no pause in between seem to lead to a different note than, say, three bursts with a slight pause after the first, or two bursts with a long pause between them.

And because Yu and Margoliash could look at the individual activity of several nerve cells at the same point in the generation of the song, they could also see how the combined temporal patterns of groups of cells can have specific meaning to the cells receiving this input. Somehow, the input from several HVc cells, with each firing according to its syllable-specific pattern, sums together to communicate to the RA cells what notes need to be generated for that particular syllable. "It's clear that the brain is a temporal pattern processor," says Margoliash.

Just accomplishing these kinds of measurements in birds was "quite a significant feat," says Doupe. Male finches sing only when they are at ease, and any recording system has to cope with their tendency to puff up their chests and hop about when they sing. Indeed, Margoliash says, he and Yu struggled for a long time to design a workable technique. "We had years where we got virtually no data," Margoliash recalls, until they finally designed a light and robust recording device.

The work should aid more than just future zebra finch studies, because it has possible implications for how the human brain controls speech. Indeed, neuroscientist John Middlebrooks of the University of Michigan, Ann Arbor, suggests that its implications may be even wider. He points, for example, to an intriguing tie between Yu and Margoliash's findings and his own studies of how the ear responds to sounds. Middlebrooks has found that when nerve endings in the ear are stimulated they send a very precise signal to the brain, much like the precise signal the RA sends to the throat muscles in a singing finch. Then as the sound's message travels to ever higher brain centers, it becomes ever more abstract, he notes—Yu and Margoliash's hierarchy in a new setting.

Moreover, Middlebrooks suggests that the way the brain executes the series of movements necessary to make sounds may prove to be the way the brain controls many of the body's activities. "[The findings] will have relevance to many aspects of human behavior," he predicts. If so, then the song of the zebra finch will indeed be music to the ears of neuroscientists.

–Elizabeth Pennisi