Neuron, Vol. 21, 359–371, August, 1998, Copyright ©1998 by Cell Press

Toward a Song Code: Evidence for a Syllabic Representation in the Canary Brain

Sidarta Ribeiro,* Guillermo A. Cecchi,† Marcelo O. Magnasco,† and Claudio V. Mello*‡
*Laboratory of Animal Behavior
†Center for Studies in Physics and Biology
The Rockefeller University
New York, New York 10021

Summary

We show that presentation of individual canary song syllables results in distinct expression patterns of the immediate-early gene *ZENK* in the caudomedial neostriatum (NCM) of adult canaries. Information on the spatial distribution and labeling of stained cells provides for a classification of *ZENK* patterns that (1) accords to the organization of stimuli into families, (2) preserves the stimuli intrafamily relationships, and (3) confers salience to natural over artificial stimuli, resulting in a nonclassical tonotopic map. Moreover, complex syllable maps cannot be reduced to any linear combinations of simple syllable maps. These properties arise from the collective response of NCM neurons to auditory stimuli, rather than from the behavior of single neurons. The syllabic representation described here may constitute an important step toward deciphering the rules of birdsong auditory representation.

Introduction

A central and challenging problem in neurobiology is how the brain represents complex objects. Insight into this issue is most likely to come from the investigation of natural objects that are relevant to behavior and reproductive success and thus play a significant role in brain evolution. One such object is birdsong, a complex stimulus of key importance for mating and territorial defense (Immelman, 1969; Brown, 1975; Marler and Peters, 1982; Godard, 1991; Kroodsma and Miller, 1996). Canary song, in particular, is highly structured, consisting of a succession of phrases that are formed by the rapid repetition of basic components or syllables, reminiscent of those in human speech (Nottebohm, 1972; Nottebohm and Nottebohm, 1978) (Figure 1A). There exist several breeds of domesticated canaries, each with its own characteristic syllabic repertoire (Güttinger, 1985; Mundinger, 1995).

The caudomedial neostriatum (NCM) is one of the main auditory processing stations in the songbird brain. NCM receives inputs from the auditory thalamus and from field L, the primary auditory region of the telencephalon, and is indirectly connected with regions that originate a descending auditory pathway (Vates et al., 1996; Mello et al., 1998); therefore, it is likely analogous to supragranular layers of the mammalian auditory cortex. Electrophysiological and 2-deoxyglucose uptake studies demonstrated responses to auditory stimuli, including song, in a vast zone of the avian caudal neostriatum, including NCM (Leppelsack and Vogt, 1976; Theurich et al., 1984; Müller and Leppelsack, 1985; Müller and Scheich, 1985; Scheich, 1991). Recent studies aimed more specifically at NCM show that the electrophysiological responses to novel songs habituate to repeated stimulus presentations in a song-specific and long-lasting manner (Chew et al., 1995; Stripling et al., 1997). This habituation has been postulated as a mechanism involved in the formation of auditory memories (Chew et al., 1996a, 1996b). NCM activation by song is also accompanied by a rapid and transient induction of the immediate-early gene *ZENK* (Mello et al., 1992), also known as *NGFI-A* (Milbrandt, 1987), *zif-268* (Christy et al., 1988), *egr-1* (Sukhatme et al., 1988), and *Krox-24* (Lemaire et al., 1988). *ZENK* induction in NCM is highest for same-species song, suggesting a tuning to species-specific features. Furthermore, this genomic response is modulated by experience. It is highest when birds are presented for the first time with same-species but unfamiliar songs and habituates when the same song is repeatedly presented, to recur upon introduction of another novel song (Mello et al., 1995).

The *ZENK* gene encodes a zinc finger transcription regulator whose induction occurs in association with neuronal depolarization and has been used as a marker for brain activation in a variety of physiological and behavioral conditions (Chaudhuri, 1997). *ZENK* induction is not synonymous with neuronal depolarization, but rather appears to reflect a specific type of electrophysiological activity related to neuronal and synaptic plasticity. For instance, its induction is associated with hippocampal long-term potentiation (Cole et al., 1989) and neuronal morphological changes after exposure to a complex environment (Wallace et al., 1995). Indeed, electrophysiological activation in the songbird brain is not always accompanied by *ZENK* expression (Mello and Clayton, 1994; Mello et al., 1995; Jarvis and Nottebohm, 1997; Mello and Ribeiro, 1998). Nevertheless, no evidence is presently available for *ZENK* induction in mature neurons in the absence of electrophysiological activation (see Discussion in Chaudhuri, 1997). It is therefore reasonable to assume that *ZENK* expression reflects, to a large extent, the underlying activation state of the neuronal populations studied.

To investigate the auditory representation of song, we took advantage of the *ZENK* induction phenomenon and mapped ZENK protein expression in the NCM of canaries (Waterslager breed) presented with various stimuli, including a whole song and several syllables that occur naturally in the repertoire of these birds. We studied the most frequent syllable types found in our aviaries; these cover three out of the four most common syllable types in this breed, according to a perceptual classification (Güttinger, 1985; see below). We grouped the syllables based on sonographic analysis into classes or families, namely whistles, combinations, and modulations (Figure 1A), and compared them with artificial stimuli. The resulting *ZENK* expression patterns were analyzed with

‡To whom correspondence should be addressed.

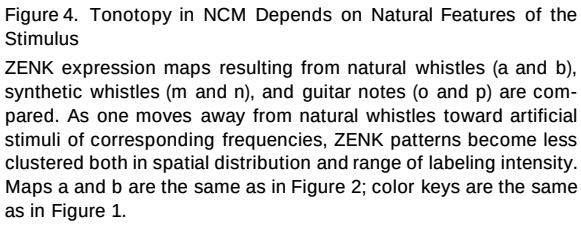

Figure 4. Tonotopy in NCM Depends on Natural Features of the Stimulus

ZENK expression maps resulting from natural whistles (a and b), synthetic whistles (m and n), and guitar notes (o and p) are compared. As one moves away from natural whistles toward artificial stimuli of corresponding frequencies, ZENK patterns become less clustered both in spatial distribution and range of labeling intensity. Maps a and b are the same as in Figure 2; color keys are the same as in Figure 1.

is progressively disrupted for synthetic whistles and guitar notes, providing a clear demonstration that frequency is not the only feature of a natural whistle being mapped in NCM.

Relationship between a Complex Stimulus and Its Components

A quantitative analysis of combinations allowed us to investigate how the component parts of a complex syllable contribute to the representation of that syllable. Given that the component whistles occurred equally often during the presentation of combinations, combination maps could be just the result of the summation of the patterns elicited by the component whistles. A comparison between a sequence map (Figure 6A) and the simple sum of the maps of its component whistles (Figure 6B) shows that this was not the case. The two maps have visible differences: for instance, the green area of activation in dorsal NCM (indicated by arrows in Figure 6) in the sum map is absent from the sequence map. When compared pixelwise using a similarity index (Chapman et al., 1996) that scores 1.0 for total identity and 0.593 for two random patterns (see Experimental Procedures), these maps yielded a score of 0.79, providing quantitative evidence that the two maps are dissimilar.

It is possible, however, that the component whistles provided different contributions to the patterns elicited by their combinations. We therefore tested whether combination patterns could be the result of a weighted sum of the component whistles. We generated a map in which the contributions of the two component whistles were weighted so as to minimize the difference between their sum map and the sequence map (Figure 6C). This weighted map, however, still differs from the sequence map (compare Figures 6A and 6C); its similarity to the sequence map (0.81) is just slightly higher than that of the simple sum map. This indicates that a combination is distinct from any linear superposition of the maps of its components. Further support for nonlinear processing of combinations is given by the fact that the patterns elicited by chords or sequences of the same whistles are different, and can be clearly discriminated by their degree of spatial clustering (Figure 5C). In addition, within the group of modulations, the map of a double sweep is different from the sum of the maps of its component sweeps (Figure 2, $l \neq j + k$; notice, for instance, how the blue/green area of activation in the ventralmost portion of NCM in k is absent from l).

Habituation Experiment

To further address the relationship between a combination and its component whistles, we explored the phenomenon of habituation of the *ZENK* gene response after sustained stimulation (Mello et al., 1995; Mello and Ribeiro, 1998). The rationale was to suppress the response caused by individual whistles from that seen when they are presented in combination, in order to reveal a combination-specific response. We exposed canaries to two whistles independently for a period sufficient to ensure habituation of ZENK protein expression (7 hr) and then presented a combination (sequence) of the same whistles for 30 min; controls continued to hear the two whistles independently for the same period (Figure 7A).

While controls showed extremely low ZENK levels throughout most of NCM, canaries habituated to whistles and then presented with the combination (sequence) showed robust ZENK activation, indicating a reinduction of gene expression by the last stimulus (Figure 7B). Interestingly, reinduction did not selectively reveal a subdomain of the activation in the sequence map. In fact, the most noticeable feature of the reinduction pattern is that labeled cells appeared mainly outside the region where the combination mapped prior to habituation (compare Figures 6A and 7B, bottom). These results reinforce the notion that a pattern resulting from a complex stimulus cannot be trivially explained by the patterns elicited by its components.

Principal Component Analysis (PCA)

To obtain an independent quantitative confirmation of the overall representation of song and syllables in NCM,

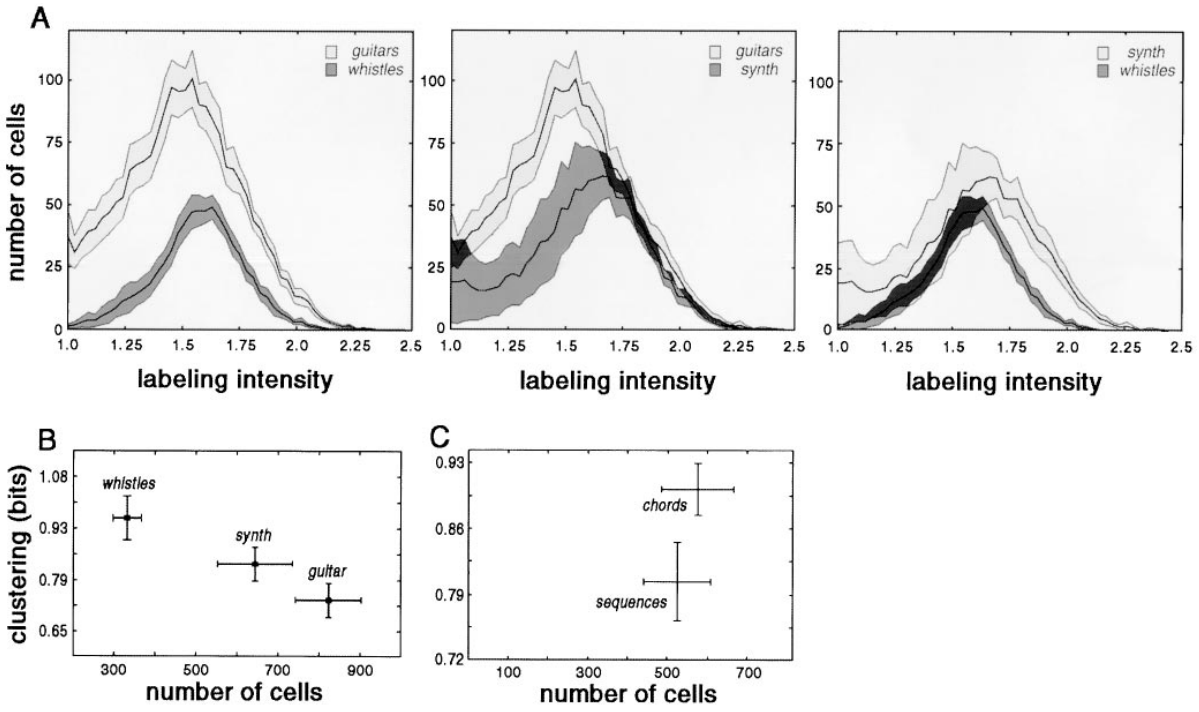

Figure 5. Quantitative Analysis of ZENK Expression Patterns

(A) Histograms of distributions of ZENK-labeled cells resulting from presentation of natural whistles, synthetic whistles, and guitar notes. For each distribution a central line indicates mean values of the number of cells with a given labeling intensity, while shaded areas represent the variance (±SEM); regions of overlapping distributions in pairwise comparisons are indicated by the darker gray. Notice that each stimulus elicited a distinct profile: the distributions of natural whistles and guitar notes are completely nonoverlapping ($p < 0.01$), while those of guitar notes and synthetic whistles are nonoverlapping in the low-to-middle range of labeling intensities ($p < 0.05$), and those of synthetic and natural whistles do not overlap in the upper range ($p < 0.01$).

(B) Clustering decreases and number of ZENK-labeled cells per section increases progressively from natural whistles to synthetic whistles and to guitar notes. Plotted are averages of the clustering indices of maps a and b (whistles), m and n (synth), and o and p (guitars) in Figure 4. Clustering in bits; error bars represent SEM.

(C) Sequences and chords activate comparable numbers of cells in NCM but can be discriminated by their degrees of clustering. Plotted are averages of the clustering indices of maps e and f (sequences) and g and h (chords) in Figure 2. Error bars represent SEM.

the ZENK expression patterns were subjected to Principal Component Analysis (PCA). PCA identifies "features" (components) that best discriminate patterns within a given set (Richmond and Optican, 1987), essentially by performing a high dimensional version of the linear best fit (see Experimental Procedures). Figure 8 depicts two different perspective projections onto a plane of the three-dimensional space defined by the first three components of the resulting PCA. The first component (x-axis) was related to the number of cells, the second (y-axis) to intensity distribution and rostrocaudal activation, and the third (z-axis) to dorsoventral activation. We found that these first three components were already enough to classify and organize the patterns into a structure concordant with that of the set of stimuli: different families of patterns lie at different depths/strata, without any disruptions of one family of patterns by other families. For instance, whistle patterns span a thin tube along the z-axis that does not intersect the subspaces defined by other families. Of particular

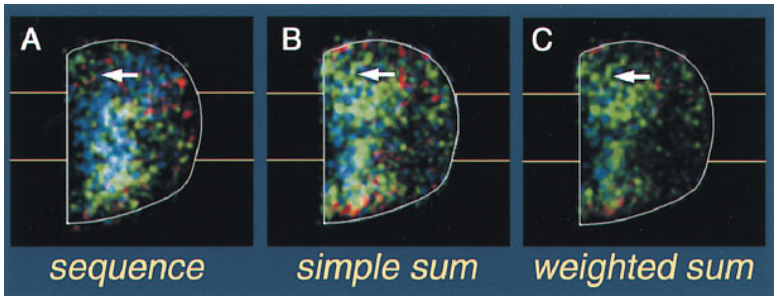

Figure 6. Contribution of the Maps of Two Whistles to the ZENK Expression Map of Their Combination as a Sequence

The map resulting from the presentation of a sequence ([A]; same as f in Figure 2) is compared to the simple sum (B) and weighted sum (C) of the maps of the corresponding component whistles (b and d in Figure 2). Notice that neither sum map is a good approximation to the sequence map; in particular, maps (B) and (C) contain a region of green activation in dorsal NCM (indicated by the arrows) that is not present in (A).

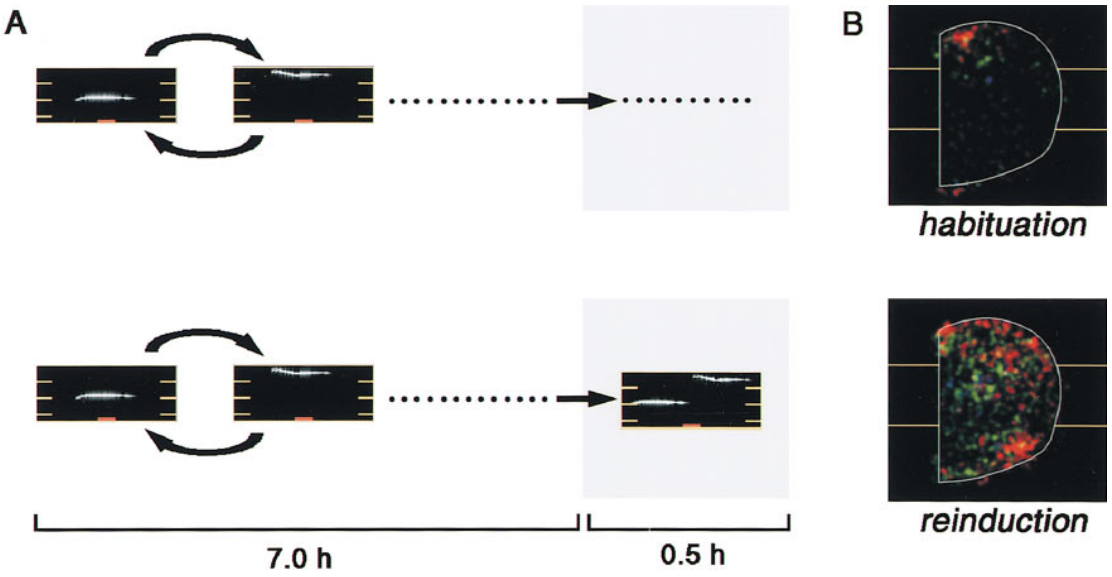

Figure 7. Habituation Experiment

(A) Experimental design: canaries were presented for 7.0 hr with two component whistles (same as b and d in Figure 2), alternated every 30 s; during an additional 0.5 hr period, controls continued to hear the whistles in alternation (upper row), while experimental animals were presented with the same whistles combined as a sequence (bottom); all animals were killed 60 min after the offset of stimulation.

(B) While ZENK expression is nearly suppressed in NCM after habituation to the component whistles (top), presentation of the sequence after habituation to the individual whistles causes ZENK reinduction (bottom); the latter is most pronounced in NCM regions that were not activated by presentation of the sequence without previous habituation to its component whistles (compare with Figure 6A). For both panels, color keys are the same as in Figure 1C.

interest is the fact that PCA sharply distinguishes whistle patterns from those of similar but artificial stimuli (synthetic whistles and guitar notes). In addition, the PCA arranged individual members orderly within their families, preserving stimuli kinship. Notice in this regard how the third component organized whistles and combinations according to their frequency, and modulations according to their ending frequency. For artificial stimuli, synthetic whistles were shifted in the z-axis with respect to whistles of the same frequencies but still aligned along this axis according to frequency, whereas this property was disrupted for guitars.

Discussion

The ZENK mapping methodology allows one to assess the activation levels and topographic distribution of large populations of song-responsive neurons. Particularly noteworthy is the fact that the experiments can be done in alert animals, without interference with their natural behaviors. Furthermore, ZENK levels in silence or after repeated stimulation are very low, yielding a high contrast for detecting responses to novel stimuli. In the present study, an average of 450 labeled neurons were detected per parasagittal plane, embedded in a lattice of $\sim$6000 nonlabeled neurons. The main disadvantages of the method are the poor temporal resolution and the fact that one bird is killed for each stimulus, limiting the number of stimuli that can be included in one study. In contrast, single unit electrophysiological studies can assess the response to several different stimuli and provide detailed information in the temporal domain but typically use anesthetized animals and investigate the activity of $<$40 neurons per animal (e.g., see Margoliash and Fortune, 1992; Doupe, 1997; Solis and Doupe, 1997; see also discussions in Capsius and Leppelsack, 1996; Margoliash, 1997). The information obtained with ZENK mapping is therefore complementary to that obtained with electrophysiology. This tradeoff proved invaluable for our purposes, as the ability to recognize and classify ZENK patterns resulting from different syllables strictly depended on analysis of the collective behavior of a large population of syllable-responsive neurons (see below).

As noted in the introduction, ZENK expression cannot be equated with electrophysiological activity. While immunostained regions in NCM most likely represent neuronal activation, areas devoid of ZENK could, in principle, have undergone neuronal depolarization without triggering the ZENK response. As suggested for other systems (Dragunow, 1996), the ZENK expression patterns we described may reflect a particular type of neuronal activity associated with plasticity, perhaps leading to the formation of auditory memories (Chew et al., 1995; Mello et al., 1995), but further work is needed to settle this issue. Indeed, the occurrence of ZENK induction associated with both learned and unlearned components of sexual behavior in the Japanese quail (Ball et al., 1997) indicates that caution needs to be exercised when trying to establish links between *ZENK* gene expression and learning. The interpretation of our present results, however, does not depend on determining the precise relationship between ZENK induction and electrophysiological activity or neuronal plasticity. In fact, as

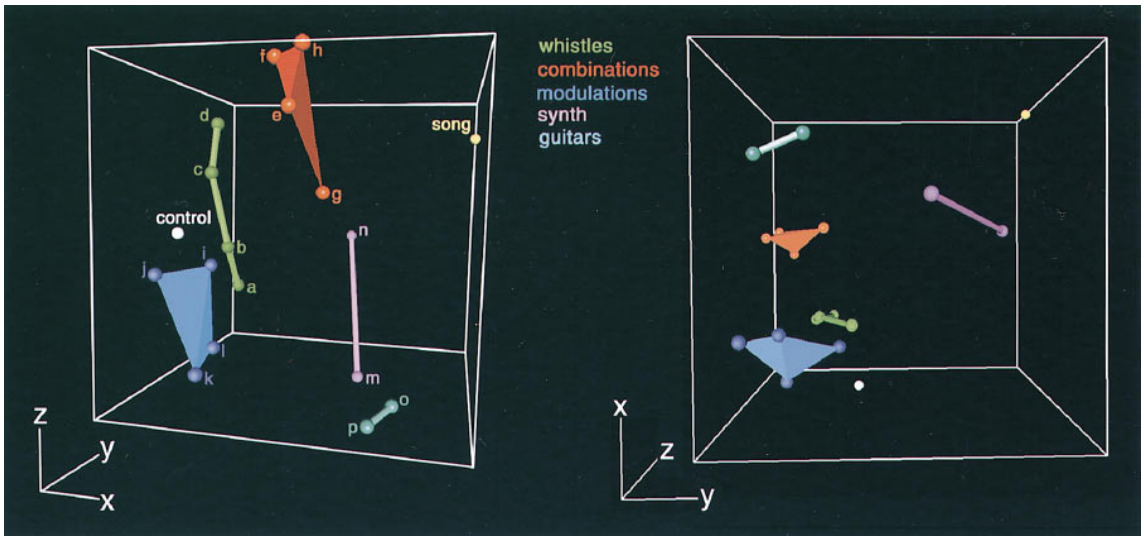

Figure 8. PCA of ZENK Expression Patterns Resulting from Natural and Artificial Stimuli

Shown are two perspectives of the space defined by the first three components of PCA (represented by the x-, y-, and z-axes, respectively); the components were normalized independently so that the coordinates span a unit cube. The distribution of ZENK patterns within this three-dimensional space mirrors the organization of the stimuli into different families; notice that individual patterns are aligned along the z-axis according to their frequency (natural and synthetic whistles) or ending frequency (modulations). The letters correspond to patterns shown in Figures 2 and 4.

discussed below, a quantitative analysis of the intrinsic organization of the patterns per se can provide some important insights into the functional organization of NCM.

Behavioral Responses of Female Canaries to Song

Behavioral studies have demonstrated that female canaries show sexual preference for own-breed songs, scored as the number of copulatory solicitation displays evoked by songs of various breeds, as well as other species' songs (Vallet et al., 1992). Female canaries also show sexual preferences for particular phrases within a song repertoire, mainly responding to short phrases with abrupt frequency modulations and short silences (Vallet and Kreutzer, 1995). We did not observe that fast modulations elicited stronger responses than other syllables assessed; in fact, modulations yielded the weakest ZENK responses in NCM. Thus, even though NCM appears to play a role in the auditory processing of syllables, other brain sites may show a more selective response to song elements with high sexual relevance. It is important to note, nonetheless, that the females studied here were not in the state of sexual responsiveness to males during which a preference for modulations has been described (Vallet and Kreutzer, 1995), and which is characterized by behaviors such as nest building and egg laying and by particular hormonal changes (Goldsmith et al., 1984; Wingfield and Goldsmith, 1990).

Rostral NCM Is Not a Classical Tonotopic Map

Different canary whistles activate clusters of ZENK-labeled neurons in rostral NCM; these clusters map along the dorsoventral axis in a graded manner, according to their frequencies. This finding fits well with the fact that rostral NCM receives afferents from the adjacent primary auditory telencephalic region, field L, which is tonotopically organized (Müller and Leppelsack, 1985; Heil and Scheich, 1991). It would thus seem that rostral NCM follows a classical tonotopic organization (Woolsey and Walzl, 1942; Reale and Imig, 1980), with one topographic axis directly mapping sound frequency. If this were the case, different stimuli of the same frequencies should have similar ZENK patterns. The comparison between natural whistles and artificial stimuli, however, argues against a simple tonotopic hypothesis. Synthetic whistles and guitar notes elicited widespread activation in NCM, in contrast to the highly clustered patterns obtained with natural whistles. Indeed, the degrees of clustering for natural whistles, synthetic whistles, and guitar notes are significantly different. Furthermore, the broadening of the spatial distribution of ZENK-labeled neurons was accompanied by a broadening of the labeling intensity distribution: while whistle patterns are dominated by intermediate (green) levels of activation, artificial stimuli patterns have considerable amounts of cells with low (blue) and high (red) ZENK levels.

The patterns resulting from artificial stimuli are not just stronger versions of the patterns resulting from natural whistles of corresponding frequencies. If this were the case, one would have expected obvious areas of peak activation in the former, corresponding to the clusters seen in the whistle maps. A close inspection of Figure 4, however, shows that guitar notes had no readily recognizable areas of peak activation; as for synthetic whistles, besides a considerable degree of activation in caudal NCM, activation in rostral NCM extends to regions that display little or no ZENK expression in the corresponding natural whistle patterns. This effect was quantitatively confirmed by PCA: the position along the axis that correlates with frequency is different for natural and

artificial stimuli of the same frequencies (compare the position along the z-axis of patterns a, m, and o and b, n, and p in Figure 8).

In summary, the specificity of the ZENK response is degraded as one moves away from the natural stimulus: while it is trivial to discriminate two whistles of different frequencies based on their ZENK patterns, that task is much more difficult for synthetic whistles and virtually impossible for guitar notes. These observations constitute the basis for our claim that the NCM does not behave as a classical tonotopic map: its topographic organization according to frequency is dominated by natural features of the stimuli. Interestingly, electrophysiological studies of field L neurons in the mynah bird (Hose et al., 1987) have demonstrated a tuning to pitches and rhythms typical of animal communication sounds. It is possible that similar mechanisms operate in NCM, but further work is necessary to establish the precise acoustic features to which NCM neurons are tuned.

Clustering as Information

Our clustering index provided a useful readout of the topographic relationships among labeled cells in the ZENK patterns. It is independent of the number of cells in the patterns (Okabe et al., 1992), reflecting only their relative spatial distribution. For instance, two patterns with a comparable number of cells may have substantially different clustering indices (see Figure 5C) and vice versa. Thus, the inverse relationship between clustering and number of cells seen in the comparison of natural and artificial stimuli (Figure 5B) does not reflect an intrinsic interdependence between these two measures, but rather the progressive change in the spatial distribution of labeled cells.

Our clustering index corresponds strictly to the Information Content of the spatial pattern of labeled cells (Shannon, 1948; Brillouin, 1989; see Experimental Procedures). An intriguing conclusion that can therefore be drawn from our analysis is that the expression patterns elicited by natural syllables contain more information, or salience, and that conversely those elicited by artificial stimuli have less information or salience, being hardly distinguishable from one another. Information Content cannot be equated with biological relevance or meaning, and it should be clear that we do not imply that patterns resulting from natural stimuli are more meaningful or relevant in any behavioral sense. This could, in principle, be said of the stimuli themselves but not necessarily of their corresponding ZENK expression patterns. The higher Information Content refers solely to the internal structure of the ZENK patterns, which in turn presumably reflects the functional organization of NCM. The higher clustering in the natural whistle patterns, as opposed to the artificial ones, provides a topographic parallel to results showing that spike trains elicited by naturalistic stimuli carry a high Information Content in their temporal structure (de Ruyter van Steveninck et al., 1997). We would therefore like to suggest that clustering, both in the temporal and spatial domains, is a relevant variable to consider when investigating brain representations.

Nonlinearity of Complex Stimuli: Building Syllabic Identity?

A recurring theme in our study was the systematic nonlinearity observed whenever we broke a stimulus into smaller (or simpler) pieces: the maps of natural whistles cannot be reduced to the maps of synthetic whistles, nor maps of combinations to those of individual whistles, or maps of double sweeps to those of individual sweeps. Hence, a complex syllable pattern is related to, but cannot be fully explained by, the patterns of its components. It is tempting to speculate that this nonlinear processing of discrete syllable features provides a mechanism for building up neural responses with unique identity within NCM. It is at present unclear whether such processing occurs somewhere along the ascending auditory pathway, within NCM itself, or as a result of interactions between NCM and its reciprocal targets.

The habituation experiment provided strong support for the notion of nonlinearity. It was conducted as an independent and direct assessment of whether a pattern elicited by a complex syllable can be explained by the responses evoked by its component parts. The results show clearly that this is not the case: one cannot abolish the response to a sequence by suppressing the responses to its component whistles. We therefore conclude that combinations are processed as distinct syllables, with unique identity. The fact that a sequence presented after habituation to its component whistles activated mostly regions that were not activated by the same sequence when presented without previous habituation indicates that the functional architecture of NCM can be modified by experience or context. A possible explanation for such reorganization, given the existence of GABAergic networks in the avian auditory telencephalon (Müller, 1988), is that presentation of individual whistles could have an inhibitory action on caudal NCM regions. According to this hypothesis, habituation of rostral NCM by repeated presentation of individual whistles would release the inhibition on caudal NCM, leading to the ZENK expression pattern observed in Figure 7B (bottom). Although further experimentation is necessary to elucidate the mechanisms involved, our data clearly indicate that syllabic maps are not the result of a static parcelling of NCM.

Another noticeable feature of the ZENK expression patterns elicited by different stimuli was that they partially overlapped. In consequence, particular NCM regions must participate in the representation of more than one syllable. Likewise, single neurons probably take part in the response to more than one stimulus. While the issue awaits confirmation, indirect evidence for the participation of single NCM neurons in the response to multiple song stimuli has been presented elsewhere (Mello et al., 1995).

The properties of nonlinearity and partial overlap suggest that syllabic representations in NCM involve sets of highly interacting cells, recruited from the total pool of NCM neurons; each stimulus would recruit a subset of the total neuronal population in a unique manner. According to this hypothesis, the response properties of individual cells cannot be meaningfully described without reference to the overall physiological state of the brain region in which they occur. Indeed, it should

be emphasized that nonlinearity and partial overlaps were properties we observed while analyzing the global patterns of activation and not the responses of individual cells that participate in them. Thus, these properties emerge from the collective behavior of neurons within NCM. Such a scheme is in line with the notion of ensemble coding (Lashley, 1950; Gross, 1992; Nicolelis et al., 1995; Deadwyler and Hampson, 1997; Joerges et al., 1997; Rolls et al., 1997). It also resonates with recently observed nonlinear combinatorial interactions of odor responses in the honeybee brain, postulated as possibly "crucial for the formation of singular codes for complex odor blends" (Joerges et al., 1997). In our study, we provide strong and quantitative evidence in favor of such a coding scheme in the brain of a higher vertebrate.

Toward a Song Code

The ZENK pattern elicited by a whole canary song (as in Figure 1C) is very complex both in spatial distribution of labeled cells and in the range of labeling intensities. This pattern presumably reflects the highly structured nature of song itself, but one cannot decode such a relationship based on that result alone. We took in our study a reductionistic route and dissected the song into basic components, or syllables. The results obtained provide a possible substrate or support for a syllabic auditory representation in NCM. Individual ZENK maps are representations of particular syllables and could be considered as the outputs of a syllabic code, in its turn understood as the rules by which the brain transforms the physical properties of a set of syllables into a set of representations. By analogy, one could say that a single syllabic map is like a letter in the alphabet, but not the alphabet, much less the language. Indeed, we have not conducted an exhaustive survey of the entire canary syllabic repertoire, nor do we claim to have uncovered all of the rules that govern the transformation of song syllables into their representation. We have rather deduced, from the intrinsic organization of syllabic maps, some properties of a syllabic code.

Our use of PCA constitutes a central part of our argument. As it was fed information only on the spatial distribution and labeling intensities of ZENK-labeled cells in NCM, the PCA provided an independent and unbiased confirmation that sufficient information is present in ZENK expression patterns to discriminate the auditory responses to various natural and artificial stimuli. Furthermore, the PCA was able to separate the patterns into distinct families that corresponded to our stimuli classification. Finally, the organization of stimuli within each family (such as the ranking of natural whistles according to frequency neighborhood or the distinction between chords and sequences within the combination family) was clearly preserved in the set of ZENK patterns in NCM. Thus, these patterns not only display syllabic identity but also a higher order kinship organization. The ability to perform such a classification task is a minimum requirement for an auditory brain region where a syllabic representation takes place and is met quite nicely by NCM.

It should be pointed out that the classification scheme we adopted for syllables is based on sonogram analysis, without any assumptions about a correspondence to perceptual categories. The purpose here was strictly analytical: to learn how the representation of a complex syllable is related to those of its component elements, it was important to separate syllables composed of multiple whistles (which we named combinations) from whistles that occur naturally without accompanying elements (named individual whistles). Our classification segregates chords from whistles of the same fundamental frequency, even though such syllables are placed in the same category by a classification based on a combination of sonogram analysis and human perception of canary songs (Güttinger, 1985). Conversely, our classification places together, within a single family, syllables such as sequences and chords, which are clearly perceived as distinct by humans. Our results demonstrate that the organization of syllabic maps in NCM mainly accords to such a sonographic classification, indicating that NCM is fundamentally responsive to spectral features of the stimuli. Interestingly, the patterns elicited by chords and sequences made of the same components can be distinguished by their degree of clustering. In that respect, chords are closer to individual whistles than to sequences (Figures 5B and 5C), in agreement with the fact that humans perceive chords as less similar to sequences than to whistles of the same fundamental frequency. Further experimentation will be necessary to elucidate the relationship between ZENK patterns in NCM and the bird's ability to perceive and discriminate syllables.

How do maps elicited by individual syllables relate to the complex pattern elicited by a whole song? Can different songs also be discriminated based on ZENK expression patterns in NCM? Is that information then made available to brain centers more directly involved in the bird's behavioral response to song? Our study has uncovered basic features of NCM's functional organization and provided tools that allow such questions to be addressed, taking us a step closer to deciphering the brain encoding of birdsong.

Experimental Procedures

Animal Groups

Our study consisted of 20 groups ($n = 4$ per group) comprised of 17 different stimuli and a silence control, plus two groups in the habituation experiment, for a total of 80 adult canaries (*Serinus canaria*, Waterslager breed). All animals were raised through adulthood in the aviaries of the Rockefeller University Field Center (Millbrook, NY), in contact with other birds of the same breed. Since the natural stimuli used were recorded from the same aviaries several years before the birth of the subjects studied, the birds can be considered experienced to the syllable types presented but naive with respect to the particular song and syllables utilized. We used females, which do not sing under our experimental conditions, to avoid ZENK induction by auditory feedback during singing.

Stimulation

Each bird was acoustically isolated for a day to minimize basal ZENK expression, and then presented with a particular stimulus. The whole song and all natural syllables were obtained from Waterslager canary repertoire. Synthetic whistles are pure sinusoidal tones whose amplitude envelopes were smoothed to resemble the elliptic shape of natural whistle envelopes. Guitars are digitally sampled guitar notes from which harmonics were filtered down to 30 dB below the intensity of the dominant frequency; the resulting intensity of harmonics during presentation of guitar notes was therefore 10 dB below ambient noise levels (see below). The song studied here

(Figure 1A) has a duration of 10 s. The average duration of other stimuli is as follows: 365 ms for natural whistles, chords, synthetic whistles and guitar notes; 720 ms for combinations; and 25–50 ms for modulations. These stimuli were presented in quick repetition for 10 s, to match the duration of the whole song; the intervals between consecutive renditions were 10–20, 25, and 50–100 ms, respectively, for the three groups above. The resulting rates of presentation were approximately those at which the different natural syllables occured in the canary song studied here; artificial stimuli were presented so as to match the protocol for natural whistles. Stimulation was delivered at a mean intensity of 70 dB sound pressure level (SPL), in 60 such blocks of 10 s, equally spaced by 20 s of silence (total duration of 30 min). Unstimulated controls were kept in silence (average ambient noise intensity of 50 dB SPL). In the habituation experiments, the birds were initially habituated to two individual whistles presented separately, in alternating blocks of 10 s each (separated by 20 s of silence) for 7 hr. One group was then presented with the same two whistles combined as a sequence for 30 min (also in a 10/20 s schedule), while the control group continued to be presented with the same two whistles separately.

Immunocytochemistry

Animals were killed and perfused with fixative 60 min after the end of stimulation, which is well within the time of peak ZENK protein expression for animals stimulated for 30 min (Mello and Ribeiro, 1998). For animals stimulated for 7 hr, this represents a time point when ZENK protein expression is habituated. The brains were dissected, frozen, and sectioned (20 μm) in the parasagittal plane. Particular care was taken to ensure that all brains were cut in the same plane. We used an anti-ZENK polyclonal antiserum (C-19, Santa Cruz Biotechnology, Santa Cruz, CA) and a previously described ICC protocol (Mello and Ribeiro, 1998). Cell labeling was the result of 3,3′-diaminobenzidine/nickel (DAB/nickel) precipitation over cell nuclei, where ZENK protein is localized. Staining variability was minimized by reacting all sections in a single batch and normalizing all cell labeling respective to background (see below).

Mapping

We analyzed sections between 1100 and 1300 μm from the midline; this particular level was chosen because it showed ZENK expression in response to all stimuli used. Human error in the identification of labeled cells was minimized by employing a computerized mapping system that applies a single and consistent set of criteria for labeling (Cecchi et al., 1998). Briefly, sections were scanned using a computer-controlled microscope setup; high resolution images were acquired through a CCD camera (3 pixels/μm) and automatically pasted to reconstitute a composite image of each entire NCM section. For each pixel, we measured the optical density (O. D.) resulting from DAB/nickel precipitation. Labeling intensity was defined as the ratio between the average O. D. of labeled nuclei and that of the background. Background was computed as the average O. D. of all tissue in NCM that was not recognized as labeled cells. The composite images were analyzed for recognition of labeled cell nuclei using criteria of size (3–10 μm diameter), eccentricity (ratio of elliptic radia $\geq$0.3), uniformity of labeling (variance of labeling over mean labeling $\leq$0.5), and shape (ratio between actual perimeter and perimeter of approximating ellipse $\leq$1.5). A lower cutoff for mean cell labeling intensity was set at 40% above background tissue labeling, equivalent to two STD above average background. The rostral boundary of NCM was defined by drawing a straight line along the major axis of field L2a on parasagittal sections; the dorsal, ventral, and caudal boundaries of NCM were naturally defined by the ventricular zone (Figure 1B). In order to obtain an anatomical reference for comparison across maps, a standardized NCM outline was generated by aligning the outlines of all 80 sections relative to each other and computing the average outline. All sections were then automatically aligned and scaled respective to this standardized NCM outline. The population of all cells studied (36,000 cells from 80 animals) was divided into three bins according to their labeling intensities, so that each bin contains a third of the total cell population (low, 1.40–1.62; medium, 1.62–1.77; high, 1.77–2.68). To obtain average density maps of each stimulus group, the set of points of each map was convolved with a Gaussian kernel of 50 μm radius, and an average map was generated for every four animals presented with the same stimulus. Information about labeling intensities (blue, green, and red bins) was processed independently and finally superposed. Brightness in the maps as finally displayed represents the absolute number of labeled cells per unit area, while hues display the relative proportions of cells with low, intermediate, and high amounts of ZENK protein.

Clustering Analysis

To measure the inhomogeneity of the spatial distribution of cells in each pattern, we computed the index of clustering given by Thiel's redundancy measure (T) (Okabe et al., 1992). For any pattern of scattered points (labeled cells in our case), T is defined as the negative of the difference between the Information Content (also referred to as negentropy; see Brillouin, 1989) of the given pattern (S) and that of a pattern with an equal number of points forming a regular grid (S_r), so that $T = -(S - S_r)$. For the calculation of Information Content, each pattern was subjected to a Delauney triangulation (Okabe et al., 1992), and the negentropy was computed as $S = (-\Sigma p_i \log p_i)/\ln 2$, where p_i is the area of each triangle (A_i) relative to the total area (A), i.e. A_i/A. This measure is independent of the number of cells and does not take into account their labeling intensities (high/medium/low channels); it thus is sensitive only to the relative spatial distribution of cells (Okabe et al., 1992).

Sum Maps

To generate a simple sum map (Figure 6B), we performed a pixelwise superposition of two whistle maps (Figure 2, b and d) and subtracted from it the silence map (Figure 1C). To generate the weighted sum map (Figure 6C), we took the same two whistle maps and computed the linear combination that best fits the sequence map. The fit is a least square approximation of the map $f' = (\alpha b + \beta d - z)$ to the map f, where b and d are whistle maps and f the sequence map in Figure 2, and z is the silence map in Figure 1C. The coefficients found were $\alpha = 0.58$ and $\beta = 0.36$. To quantitate the differences between the map of a sequence and the sum maps, we used a similarity index that quantifies the proximity or similarity, pixelwise, of two given maps (Chapman et al., 1996). It is computed as:

$$S = 1 - \sqrt{\frac{\sum_{i=0}^{N} (X_i - Y_i)^2}{N\Delta^2}}$$

where X_i and Y_i are the i^{th} pixel values, N is the total number of pixels, and Δ is the maximum value per pixel. This similarity index ranges from 0–1, corresponding to completely decorrelated maps and to identical maps, respectively. For comparison, the similarity index between two random maps (in which each i^{th} pixel value varies randomly) is 0.593.

PCA

To perform a PCA of the set of average maps, each map was parceled in space bins, defined by a 6×6 grid. Each space bin had three color bins, according to the labeling intensities of cells (see Mapping section above). This defined a $6 \times 6 \times 3$ vector space, so that each map was represented by a 108 dimensional vector containing coarse-grained information on spatial distribution and labeling intensity of cells. The value of each vectorial component was computed as the number of cells falling within the corresponding space-color bin.

Acknowledgments

We thank F. Nottebohm, A. Libchaber, and J. Hudspeth for their support, encouragement, and insightful comments; F. Doetsch, E. Jarvis, J. Hudspeth, C. Scharff, T. Wiesel, and three anonymous reviewers for their critical reading of this manuscript; E. Jarvis for sharing, at early stages of this project, data suggesting that song fragments elicit less ZENK expression than whole songs; D. Chialvo and R. Crist for fiery discussions; B. Sattie for technical assistance; Sandy Hills for secretarial help; and Daun Jackson, Helen Ecklund, and Sharon Sepe at The Rockefeller University Field Research Center for attentive animal care. This work was supported by Kluge and CNPq/Travel Fellowships (to S. R.), the NIDCD (grant #02853 to

C. V. M.), and the Mathers Foundation. S. R. dedicates his effort in this paper to the memory of the late Professor Luís F. G. Labouriau.

Received March 30, 1998; revised June 26, 1998.

References

Ball, G.F., Tiemçani, O., and Balthazart, J. (1997). Induction of the ZENK protein after sexual interactions in male Japanese quail. Neuroreport *8*, 2965–2970.

Brillouin, L. (1989). Life, thermodynamics and cybernetics. In Maxwell's Demon: Entropy, Information and Computing. H. Leff and A. Rex, eds. (Princeton, NJ: Princeton University Press), p. 101.

Brown, J.L. (1975). The evolution of auditory communication. In The Evolution of Behavior (New York: W.W. Norton), pp. 329–369.

Capsius, B., and Leppelsack, H.J. (1996). Influence of urethane anesthesia on neural processing in the auditory cortex analogue of a songbird. Hearing Res. *96*, 59–70.

Cecchi, G.A., Ribeiro, S., Mello, C.V., and Magnasco, M.O. (1998). An automated procedure for the mapping and quantitative analysis of immunocytochemistry of an inducible nuclear protein. J. Neurosci. Methods, in press.

Chapman, B., Stryker, M.P., and Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. J. Neurosci. *16*, 6443–6453.

Chaudhuri, A. (1997). Neural activity mapping with inducible transcription factors. Neuroreport *8*, III–VIII.

Chew, S.J., Mello, C., Nottebohm F., Jarvis E., and Vicario, D.S. (1995). Decrements in auditory responses to a repeated conspecific song are long-lasting and require two periods of protein synthesis in the songbird forebrain. Proc. Natl. Acad. Sci. USA *92*, 3406–3410.

Chew, S.J., Vicario, D.S., and Nottebohm, F. (1996a). A large-capacity memory system that recognizes the calls and songs of individual birds. Proc. Natl. Acad. Sci. USA *93*, 1950–1955.

Chew, S.J., Vicario, D.S., and Nottebohm, F. (1996b). Quantal duration of auditory memories. Science *274*, 1909–1914.

Christy, B., Lau, L., and Nathans, D. (1988). A gene activated in mouse 3T3 cells by serum factors encodes a protein with "zinc finger" sequences. Proc. Natl. Acad. Sci. USA *85*, 7857–7861.

Cole, A.J., Saffen, D.W., Baraban, J.M., and Worley, P.F. (1989). Rapid increase of an immediate early gene messenger RNA in hippocampal neurons by synaptic NMDA receptor activation. Nature *340*, 474–476.

Deadwyler, S.A., and Hampson, R.E. (1997). The significance of neural ensemble codes during behavior and cognition. Annu. Rev. Neurosci. *20*, 217–244.

de Ruyter van Steveninck, R.R., Lewen, G.D., Strong, S.P., Koberle, R., and Bialek, W. (1997). Reproducibility and variability in neural spike trains. Science *275*, 1805–1808.

Doupe, A. (1997). Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. J. Neurosci. *17*, 1147–1167.

Dragunow, M. (1996). A role for immediate-early transcription factors in learning and memory. Behav. Gen. *26*, 293–299.

Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird. Nature *350*, 228–229.

Goldsmith, A.R., Burke, S., and Prosser, J.M. (1984). Inverse changes in plasma prolactin and LH concentrations in female canaries after disruption and reinitiation of incubation. J. Endocrinol. *103*, 251–256.

Goller, F., and Suthers, R.A. (1996). Role of syringeal muscles in controlling the phonology of bird song. J. Neurophysiol. *76*, 287–300.

Gross, C.G. (1992). Representation of visual stimuli in inferior temporal cortex. Phil. Trans. R. Soc. Lond. B Biol. Sci. *335*, 3–10.

Güttinger, H.R. (1985). Consequences of domestication on the song structures in the canary. Behaviour *94*, 254–278.

Hartley, R.S. (1990). Expiratory muscle activity during song production in the canary. Resp. Physiol. *81*, 177–187.

Heil, P., and Scheich, H. (1991). Functional organization of the avian auditory cortex analogue. I. Topographic representation of isointensity bandwidth. Brain Res. *539*, 110–120.

Hose, B., Langner, G., and Scheich, H. (1987). Topographic representation of periodicities in the forebrain of the mynah bird: one map for pitch and rhythm? Brain Res. *422*, 367–373.

Immelman, K. (1969). Song development in the Zebra Finch brain and other estrildid finches. In Bird Vocalizations, R.A. Hinde, ed. (Cambridge: Cambridge University Press), 61–74.

Jarvis, E.D., and Nottebohm, F. (1997). Motor-driven gene expression. Proc. Natl. Acad. Sci. *94*, 4097–4102.

Joerges, J., Kuttner, A., Galizia, C.G., and Menzel, R. (1997). Representations of odors and odor mixtures visualized in the honeybee brain. Nature *387*, 285–288.

Kroodsma, D.E., and Miller, E.H. (1996). Ecology and Evolution of Acoustic Communication in Birds (New York: Cornell University Press).

Lashley, K.S. (1950). In search of the engram. Symp. Soc. Exp. Biol. *4*, 454–482.

Lemaire, P., Revelant, O., Bravo, R., and Charnay, P. (1988). Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells. Proc. Natl. Acad. Sci. USA *85*, 4691–4695.

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

Margoliash, D. (1997). Functional organization of forebrain pathways for song production and perception. J. Neurobiol. *33*, 671–693.

Margoliash, D., and Fortune, E.S. (1992). Temporal and harmonic combination-sensitive neurons in the zebra finch's HVc. J. Neurosci. *12*, 4309–4326.

Marler, P., and Peters, S. (1982). Sparrows learn adult song and more from memory. Science *213*, 780–782.

Mello, C.V., and Ribeiro, S. (1998). ZENK protein regulation by song in the brain of songbirds. J. Comp. Neurol. *393*, 426–438.

Mello, C.V., Vicario, D.S., and Clayton, D.F. (1992). Song presentation induces gene expression in the songbird forebrain. Proc. Natl. Acad. Sci. USA *89*, 6818–6822.

Mello, C.V., and Clayton, D.F. (1994). Song-induced ZENK gene expression in auditory pathways of songbird and its relation to the song control system. J. Neurosci. *14*, 6652–6666.

Mello, C.V., Nottebohm, F., and Clayton, D.F. (1995). Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. J. Neurosci. *15*, 6919–6925.

Mello, C.V., Vates, G.E., Okuhata, S., and Nottebohm, F. (1998). Descending auditory pathways in the adult male zebra finch (Taeniopygia guttata). J. Comp. Neurol. *395*, 137–160.

Milbrandt, J. (1987). A nerve growth factor–induced gene encodes a possible transcriptional regulatory factor. Science *238*, 797–799.

Müller, C.M. (1988). Distribution of GABAergic perikaria and terminals in the centers of the higher auditory pathway of the chicken. Cell Tissue Res. *252*, 99–106.

Müller, C.M., and Leppelsack, H.J. (1985). Feature extraction and tonotopic organization in the avian auditory forebrain. Exp. Brain. Res. *59*, 587–599.

Müller, S.C., and Scheich, H. (1985). Functional organization of the avian auditory field L: a comparative 2DG study. J. Comp. Physiol. *156*, 1–12.

Mundinger, P.C. (1995). Behaviour-genetic analysis of canary song: inter-strain differences in sensory learning, and epigenetic rules. Anim. Behav. *50*, 1491–1511.

Nicolelis, M.A.., Baccala, L.A., Lin, R.C., and Chapin, J.K. (1995). Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. Science *268*, 1353–1358.

Nottebohm, F. (1972). The origins of vocal learning. Amer. Nat. *106*, 116–140.

Nottebohm, F., and Nottebohm, M.E. (1978). Relationship between

song repertoire and age in the canary brain, *Serinus canarius*. Z. Tierpsychol. *46*, 298–305.

Okabe, A., Boots, A., and Sugihara, K. (1992). Spatial Tessellation Concepts and Applications of Voronoi Diagrams (New York: Wiley).

Reale, R.A., and Imig, T.J. (1980). Tonotopic organization in auditory cortex of the cat. J. Comp. Neurol. *192*, 265–291.

Richmond, B.J., and Optican, L.M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. J. Neurophysiol. *57*, 147–161.

Rolls, E.T., Treves, A., and Tovee, M.J. (1997). The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. Exp. Brain. Res. *114*, 149–162.

Scheich, H. (1991). Auditory cortex: comparative aspects of maps and plasticity. Curr. Opin. Neurobiol. *1*, 236–247.

Shannon, C.E. (1948). A mathematical theory of communication. Bell Sys. Tech. J. *27*. Reprinted in Shannon, C.E., and Weaver, W. (1949). The Mathematical Theory of Communication (Urbana, IL: University of Illinois Press).

Solis, M.M., and Doupe, A.J. (1997). Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. J. Neurosci. *17*, 6447–6462.

Stripling, R., Volman, S.F., and Clayton, D.F. (1997). Response modulation in the zebra finch neostriatum: relationship to nuclear gene regulation. J. Neurosci. *17*, 3883–3893.

Sukhatme, V.P., Cao, X., Chang, L.C., Tsai-Morris, C.H., Stamenkovitch, D., Ferreira, P.C.P., Cohen, D.R., Edwards, S.A., Shows, T.B., Curran, T., LeBeau, M.M., and Adamson, E.D. (1988). A zinc finger–encoding gene coregulated with c-*fos* during growth and differentiation, and after cellular depolarization. Cell *53*, 37–43.

Theurich, M., Muller, C.M., and Scheich, H. (1984). 2-Deoxyglucose accumulation parallels extracellularly recorded spike activity in the avian auditory neostriatum. Brain Res. *322*, 157–161.

Vallet, E.M., and Kreutzer, M.L. (1995). Female canaries are sexually responsive to special song phrases. Anim. Behav. *49*, 1603–1610.

Vallet, E.M., Kreutzer, M.L., and Richard, J.-P. (1992). Syllable phonology and song segmentation: testing their salience in female canaries. Behaviour *121*, 155–167.

Vates, G.E., Broome, B.M., Mello, C.V., and Nottebohm, F. (1996). Auditory pathways of caudal telencephalon and their relation to the song system of adult male zebra finches. J. Comp. Neurol. *366*, 613–642.

Wallace, C.S., Withers, G.S., Weiler, I.J., George, J.M., Clayton, D.F., and Greenough, W.T. (1995). Correspondence between sites of NGFI-A induction and sites of morphological plasticity following exposure to environmental complexity. Mol. Brain Res. *32*, 211–220.

Wingfield, J.C., and Goldsmith, A.R. (1990). Plasma levels of prolactin and gonadal steroids in relation to multiple-brooding and renesting in free-living populations of the song sparrow, Melospiza melodia. Horm. Behav. *24*, 89–103.

Woolsey, C.N., and Walzl, E.M. (1942). Topical projection of nerve fibers from local regions of the cochlea to the cerebral cortex of the cat. Bull. Johns Hopkins Hosp. *71*, 315–344.