Neural correlates of categorical perception in learned vocal communication

Jonathan F Prather1, Stephen Nowicki1,2, Rindy C Anderson2, Susan Peters2 & Richard Mooney1

The division of continuously variable acoustic signals into discrete perceptual categories is a fundamental feature of vocal communication, including human speech. Despite the importance of categorical perception to learned vocal communication, the neural correlates underlying this phenomenon await identification. We found that individual sensorimotor neurons in freely behaving swamp sparrows expressed categorical auditory responses to changes in note duration, a learned feature of their songs, and that the neural response boundary accurately predicted the categorical perceptual boundary measured in field studies of the same sparrow population. Furthermore, swamp sparrow populations that learned different song dialects showed different categorical perceptual boundaries that were consistent with the boundary being learned. Our results extend the analysis of the neural basis of perceptual categorization into the realm of vocal communication and advance the learned vocalizations of songbirds as a model for investigating how experience shapes categorical perception and the activity of categorically responsive neurons.

One way that perception enables adaptive behavior is by grouping variable stimuli into classes. An impressive example of this synthetic capacity of the nervous system occurs when continuous changes in a stimulus parameter are perceived as discrete perceptual categories1. Categorical perception has a prominent role in processing acoustic signals that are important to social communication2, including vocalizations made by humans3,4, monkeys5, rodents6, birds7 and frogs8,9. In spoken English, for example, the distinction between the speech sounds that elicit the phonemes /ba/ and /pa/ stems from a difference in delay time between the initial sound made by opening the lips and the onset of vocal fold vibration (voice onset time). Voice onset time varies considerably from individual to individual, but this variation is perceived categorically by native English speakers recognizing a categorical boundary between /ba/ and /pa/ at ~ 12 ms3,4. Presumably, categorical processing of speech more generally facilitates comprehension by generating perceptual constancy in the face of individual variation in many dimensions, including vocal pitch, timbre and tempo.

Despite the prominent role of categorical perception in the processing of speech and other natural signals, the neural mechanisms underlying this cognitive ability remain poorly studied. The fact that some animals have been shown to perceive human speech sounds categorically, even when they do not learn their own species-typical vocalizations5, suggests that the capacity for categorical perception reflects some fundamental and innate brain mechanism. At the same time, however, the underlying mechanism must also be sensitive to experience because the perceptual boundaries that define speech sounds vary across native speakers of different languages3,4, indicating that the boundaries, as with other features of speech, are learned. What is lacking is a model system in which the neural underpinnings of categorical perception can be studied for a natural, learned communication system. Here we asked whether sensorimotor neurons that are known to be important for learned vocal communication in songbirds could also mediate categorical perception of learned vocalizations.

Swamp sparrows (Melospiza georgiana) are especially appropriate for examining the neural events underlying categorical perception of learned vocalizations for several reasons. First, as with all songbirds studied to date, swamp sparrows learn their song notes by imitation10, a feature of human speech that is otherwise rare among animals11. Second, behavioral experiments have shown that male swamp sparrows use categorical perception to distinguish fundamental acoustic elements in their species-typical vocal repertoire2. Swamp sparrow songs comprise repeated groups of 2–5 ‘notes’ (Fig. 1a), which are composed of short pure-tonal frequency sweeps, with note categories differing primarily in duration, bandwidth and rate of change in frequency12. Previous work has shown that males use categorical perception to distinguish between note types that, similar to the phones produced in speech, are produced with considerable variation by different individuals but are grouped into natural categories. Specifically, note types in categories I and VI share most spectral features, but differ along a duration continuum12,13, and individual male swamp sparrows in a wild population in New York perceived these differences categorically, with an estimated perceptual boundary at a duration of 13 ms (Fig. 1b,c)7.

A third feature that advances swamp sparrows as a model for investigating neural activity underlying categorical perception is that their brains, as in other songbirds, contain a discrete set of specialized

---

1Department of Neurobiology, 101 Research Drive, Duke University Medical Center, Durham, North Carolina 27710, USA. 2Department of Biology, 125 Science Drive, Duke University, Durham, North Carolina 27710, USA. Correspondence should be addressed to R.M. (mooney@neuro.duke.edu).

Received 14 October 2008; accepted 1 December 2008; published online 11 January 2009; doi:10.1038/nn.2246
sensorimotor structures devoted to learning, producing and perceiving song.

Previously, we found that one of these structures, the nucleus HVC (Fig. 2a), contained a certain class of striatum-projecting neurons (HVCX cells; Fig. 2b,c) that respond to only one song type in the bird’s repertoire.

Three observations implicate song type-selective HVCX neurons in song perception. First, song perception is impaired by lesions to HVC15,16. Second, song perception is also impaired by lesions to the striatal portion of an anterior forebrain pathway into which HVCX cells project their axons. Finally, some HVCX cells show a sensorimotor correspondence that is reminiscent of mirror neurons in the monkey cortex, which are hypothesized to be important in perception.

Although HVCX cells appear to be well-suited to a role in perception, the extent to which their auditory properties are correlated with the bird’s perception of categorical differences among song features is unknown. Therefore, we used a lightweight chronic recording device to assess whether HVCX cells in freely behaving adult male swamp sparrows responded categorically to changes in note duration, known from behavioral experiments to be an important feature that distinguishes between note-type categories I and VI.

RESULTS

In a set of initial experiments, we recorded from antidromically identified HVC neurons in freely behaving adult male swamp sparrows and presented each bird’s song types through a speaker located near its perch. These recordings confirmed that HVCX neurons responded to a single song type in the bird’s repertoire (the primary song type) and revealed that robust auditory responses could be evoked in interneurons (HVCINT cells). In contrast, all of the cells of the other projection neuron type in HVC that we sampled, which innervates the song premotor nucleus RA (HVCRA cells), were unresponsive to auditory stimuli.

To test whether individual HVCX or interneurons express categorical responses to changes in the duration of notes in the primary song type, we first identified the primary song type to which a cell responded and then identified a note in the repeated syllable comprising that song type that was in either a category I (<13 ms) or a category VI (>13 ms) note type, as defined in previous studies.

We then replaced the category I or category VI note in the primary song type with a note that had similar spectral features, but had a different duration, following the same procedures used in prior behavioral studies (Fig. 3a), in which replacement notes were chosen that differed in their durations by equal increments on a logarithmic scale (see Methods for additional details). The resulting synthetic syllable was assembled into a trill to form a stimulus song, with all other notes, internote intervals and trill rate being identical to the natural song. Repeating this procedure using different replacement notes to create each stimulus, we presented to each cell a set of song stimuli comprising 5–11 variants of the primary song type, each differing only in the duration of a single replacement note in each trilled syllable, with durations spanning a range from notes that would be classified as category I to those that would be classified as category VI.

**HVCX neurons display categorical auditory responses**

Acoustic presentation of variants of the primary song type revealed that the auditory responses of HVCX neurons, but not interneurons, are highly sensitive to changes in note duration (Fig. 3). Most HVCX neurons responded robustly to song stimuli containing replacement

---

**Figure 1** Earlier behavioral tests with swamp sparrows reveal categorical perception of changes in individual song notes. (a) Male swamp sparrows sing a small repertoire of distinct song types composed of trilled syllables that comprise multiple notes (2–5 song types, each ~2-s duration). (b) Categorical perception in swamp sparrows was demonstrated previously by splicing out one note in the syllable (note C) and replacing it with a note with similar spectral characteristics, but different duration. Replacement note durations differed by equal increments on a logarithmic scale. (c) Behavioral testing revealed that swamp sparrows perceived strong differences, evident as robust expression of aggressive displays, when note transitions spanned an estimated boundary of 13 ms (stimulus set 2), but perceived little or no difference when transitions did not cross that boundary (stimulus sets 1 and 3, figure adapted from ref. 7), indicating that swamp sparrows employ categorical perception in their song discrimination. Error bars denote s.e.
notes with durations in the range of the target note category, but weakly or not at all to variants of the primary song type with replacement notes having durations outside of that category ($n = 19$ of $22$ HVCX cells, 5 birds). Robust responses were invariably evoked by stimuli containing replacement notes with durations that would fall unambiguously into the same category as the target note in the natural song (for example, note durations of 4 or 8 ms, target note duration of 7 ms, replacement notes near the categorical boundary discussed below; Fig. 3a). Following the criteria established in previous studies of categorical perception$^{7,25–27}$, we termed these cells as being categorically responsive (see Methods for additional details). Categorical responsiveness was evident in HVCX cells both within and across birds (within bird: $n = 4$ of $4$ cells, 1 bird, $P < 0.001$; across birds: $n = 19$ cells, 5 birds, $P < 0.001$; Mann-Whitney U test; Fig. 3b,c), regardless of whether the original target note was of the shorter category I type or of the longer category VI type (Fig. 3d).

Beyond the sheer abundance of categorically responsive HVCX neurons, two observations indicated that the HVCX neuronal population encodes note duration in a categorical manner. First, the three HVCX cells that did not respond categorically also did not show bell-shaped tuning curves to intermediate note durations, as might be expected if note duration is represented continuously (Supplementary Figure 1 online). Second, our dataset differed substantially from a model in which note duration is encoded linearly by HVCX cells (Supplementary Figure 2 online). In contrast to HVCX neurons, HVCINT cells responded similarly regardless of whether the replacement note was of the same category or a different category as the target note ($n = 16$ of $18$ cells, 5 birds, $P = 0.24$, Mann-Whitney U test; Supplementary Methods online), further establishing the auditory responses of HVCX neurons as categorical.

The HVCX response boundary marks a perceptual boundary Although the categorical responses of HVCX cells are consistent with their having some role in the categorical perception of note type on the basis of duration, we found it curious that the categorical boundary estimated from our neural data ($\sim 21$ ms) was different from the perceptual boundary that was previously reported for swamp sparrows on the basis of behavioral testing ($\sim 13$ ms)$^{7}$. This discrepancy suggests that the neural responses that we measured are not linked to perception of note duration or, alternatively, that the perceptual boundary in the birds that we studied differed from the previously reported values. All of the birds from which we collected neural data were obtained from a population in northwestern Pennsylvania, whereas the previous behavioral study examined a population from upstate New York, some 540 km distant$^{7}$. As with most songbirds, geographically distinct populations of swamp sparrows learn different song dialects$^{29}$, raising the possibility that the perceptual boundary separating category I and category VI notes may be influenced by learning and thus may vary across populations.

To test whether the discrepancy between our neural data and previous behavioral measures of a categorical boundary reflects a learned dialect difference, we first compared the acoustic characteristics of category I and category VI notes obtained from the Pennsylvania and
New York populations. We found that the distributions of note durations differed between these two populations, as would be predicted by a dialect difference in this phonological feature (Fig. 4a). We next asked whether population-specific differences in note durations are paralleled by population differences in song perception. We tested the categorical perceptual boundary for note duration in the Pennsylvania population, replicating the behavioral experiment carried out on the New York population in which habituation to repeated presentation of one song stimulus was followed by presentation of a test stimulus that responded to the same song type (n = 4, left, open blue squares indicate the cell in a). In contrast, categorical activity was not evident in any of the HVCINT neurons that responded to the same song type in the same bird (n = 3, right). In behavioral testing, this same bird responded to differences between stimuli that crossed the putative neurophysiological boundary but not to differences between stimuli that did not cross the boundary (Supplementary Fig. 4). Among all cells tested in all birds (n = 22 of 24 auditory responsive HVCX cells, 5 birds and 18 of 18 responsive HVCINT cells, 5 birds), categorical activity was consistently evident for HVCX neurons (n = 19 of 22 cells, 5 birds), with stronger response strengths being evoked by replacement notes in the same category as the note that was replaced. In contrast, categorical activity was very rarely evident in HVCINT neurons (n = 2 of 18 cells, 2 birds). Error bars denote s.e. (d) Categorical responses were observed in HVCX cells, regardless of whether the duration of note that was replaced was naturally short or long, and the location of the categorical boundary (black triangle) was consistent across HVCX cells and birds (21 ± 4 ms, mean ± s.d.; ten cells shown for clarity). A subset of ten categorically responsive HVCX cells (two birds) was tested further to probe the categorical boundary at a higher resolution (19, 22 and 25 ms; for example, thick blue line), revealing an estimated categorical boundary (20 ± 4 ms) that was very similar to that estimated using the full dataset (21 ± 4 ms, P = 0.47, unpaired t test).

Supplementary Fig. 4 online; see also Figs. 2b,c and 3a,b and Supplementary Fig. 5 online). Although these findings are limited to a single individual, they indicate a direct within-individual correspondence between neurophysiological data and behavioral data related to those that we measured in a field setting.

Note duration is the major determinant of HVCX responses

Although prior behavioral work implicated note duration as the primary salient acoustic feature for categorical perception, the stimuli used in that study and the initial set of stimuli used here varied note duration without controlling for changes in other covariant features of the manipulated note, namely the rate of frequency modulation or frequency bandwidth. Thus it remains to be established whether note duration, rather than frequency modulation or bandwidth, is the salient feature underlying categorical perception of category I and category VI notes by swamp sparrows and the categorical responses of their HVC neurons. Although the larger stimulus set that was required to distinguish between these alternatives is impractical to employ in field studies of wild sparrows, such measurements could be made for the responses of single neurons.

Figure 3 Manipulation of song note duration reveals categorical auditory responses of HVCX neurons. (a) Individual notes in the primary song type syllable (note C in top spectrogram) were replaced by another swamp sparrow song note with similar spectral characteristics but different duration. Auditory responses in this HVCX cell were strong among one group of replacement notes (4, 8 and 16 ms), but weak among another group (27 and 31 ms; third row, auditory response PSTH, 1-ms bin size; bottom row, stimulus syllables, gray boxes indicate replacement notes). (b) Categorical responses such as those in a were evident in all HVCX cells in this bird that responded to the same song type (n = 4, left, open blue squares indicate the cell in a). In contrast, categorical activity was not evident in any of the HVCINT neurons that responded to the same song type in the same bird (n = 3, right). In behavioral testing, this same bird responded to differences between stimuli that crossed the putative neurophysiological boundary but not to differences between stimuli that did not cross the boundary (Supplementary Fig. 4). Among all cells tested in all birds (n = 22 of 24 auditory responsive HVCX cells, 5 birds and 18 of 18 responsive HVCINT cells, 5 birds), categorical activity was consistently evident for HVCX neurons (n = 19 of 22 cells, 5 birds), with stronger response strengths being evoked by replacement notes in the same category as the note that was replaced. In contrast, categorical activity was very rarely evident in HVCINT neurons (n = 2 of 18 cells, 2 birds). Error bars denote s.e. (d) Categorical responses were observed in HVCX cells, regardless of whether the duration of note that was replaced was naturally short or long, and the location of the categorical boundary (black triangle) was consistent across HVCX cells and birds (21 ± 4 ms, mean ± s.d.; ten cells shown for clarity). A subset of ten categorically responsive HVCX cells (two birds) was tested further to probe the categorical boundary at a higher resolution (19, 22 and 25 ms; for example, thick blue line), revealing an estimated categorical boundary (20 ± 4 ms) that was very similar to that estimated using the full dataset (21 ± 4 ms, P = 0.47, unpaired t test).
To begin to determine whether note duration was a primary determinant of HVC\textsubscript{X} neuronal responses, we tested a subset of HVC\textsubscript{X} neurons (n = 10 cells, 2 birds; different subset than the cells used in high-resolution testing of boundary location) with variants of the primary song type containing computer-generated replacement notes. The use of these synthetic notes allowed us to systematically change note duration while holding constant either the note’s frequency modulation or frequency bandwidth (Fig. 5a–c). Syllables containing synthetic notes that closely replicated the features of the natural note that they replaced were just as efficacious as the natural syllable at eliciting responses from HVC\textsubscript{X} neurons (P = 0.47, Wilcoxon signed rank test; Fig. 5a). Furthermore, syllables containing synthetic notes with durations similar to those of the natural notes that they replaced, but with different frequency modulation or bandwidths, were also highly effective at eliciting responses from HVC\textsubscript{X} neurons (frequency modulation, P = 0.13; frequency bandwidth, P = 0.43; Wilcoxon signed rank test; Fig. 5b,c). Thus, these data indicate that neither frequency modulation nor frequency bandwidth has a primary role in influencing our results, and that note duration, the song feature previously implicated in affecting categorical perception of these note types\textsuperscript{7}, is the salient song feature in driving categorical responses in HVC\textsubscript{X} neurons.

**Categorical responses in sensorimotor HVC\textsubscript{X} neurons**

In a prior study in which we recorded both auditory and singing-related activity in HVC\textsubscript{X} neurons of swamp sparrows, we established that individual cells show a precise sensorimotor correspondence\textsuperscript{20}. In primates, neurons that display sensorimotor ‘mirroring’ are hypothesized to mediate perception of communication signals\textsuperscript{23}. Therefore, we investigated whether categorical responsiveness was evident in the same cells for which this sensorimotor correspondence had been established.

Categorical auditory responsiveness was observed in all HVC\textsubscript{X} neurons for which we also recorded singing-related activity (n = 5
Comparison of auditory and singing-related activity revealed that all of these neurons showed a precise sensorimotor correspondence20 (for example, see Supplementary Fig. 5). These results suggest that neurons displaying a form of sensorimotor mirroring convey perceptual information about song.

**DISCUSSION**

This study provides, to the best of our knowledge, the first evidence for neurons that encode perceptual information about a phonological feature of learned vocal behavior, specifically information about a categorical perceptual boundary. Furthermore, our neural results accurately predicted a previously unknown geographic dialectical difference in this perceptual boundary, one that we subsequently confirmed using behavioral tests with wild birds. Variation in this perceptual boundary across swamp sparrow populations strongly suggests that both categorical perception and categorical neural responses in sparrow populations are affected by experience. Finally, by identifying a cell type that expresses categorical activity, we provide insights into the neural circuitry for categorical perception of note duration.

Categorical perception has been shown to have a prominent role in a wide variety of communication systems3–9, including human speech3,4. Although categorical processing of both human and nonhuman vocalizations has been demonstrated in a variety of animals3,4,6,8, neural correlates of this phenomenon await identification. Neurophysiological studies in other modalities, however, have identified that neurons can reliably encode information about perceptual categories. For example, neurons in the monkey inferotemporal cortex respond in a categorical manner to facial features33, an important component of social signaling in primates, and brain imaging studies in humans have detected regions that show categorical sensitivity to differences in facial expression33. These studies strengthen the idea that categorical perception is mediated by categorically responsive neurons, and that the activity of those cells could be involved in communication. Our results identify neurons that respond categorically to learned vocal signals, extending the analysis of the neural basis of perceptual categorization to another major facet of communication with special relevance to human speech.

To our initial surprise, we found that the categorical response boundary for note duration that we measured from neurons differed markedly from the previously reported categorical perceptual boundary for this system using behavioral methods7. We found, however, that the response boundary that we determined from physiological recordings of HVCX cells accurately predicted a previously unidentified population difference in this perceptual boundary, indicating that the perceptual boundary varies across geographically distinct swamp sparrow populations and consistent with the view that, as is the case for human speech3,4, these boundaries can be influenced by early experience. The fact that bird songs vary geographically as a consequence of learning is widely documented3,4,33, but our data are, to the best of our knowledge, the first to suggest an effect of learning on the categorical perception of a natural signal other than speech. The observations that human infants across cultures show an initially similar ability to categorize speech sounds35 and that categorical boundaries of human speech can be detected by mammals that do not learn their vocalizations6 suggest that categorization of vocal signals may exploit innate neural mechanisms. However, the role of social experience in shaping human speech perception, including the perception of categorical boundaries, is also well documented3,4, indicating that the underlying neural mechanisms must also be pliable. Future studies in songbirds will be especially fruitful for addressing how acoustic and perhaps social experience shapes the development of categorical perception and categorically responsive neurons.

Our study localized categorical responses to HVCX projection neurons, which innervate a striatal pathway that lesion studies implicate in song perception and song learning15,16,18,22 and may be regarded as being analogous to mammalian corticostriatal neurons. Our finding that the auditory responses of these projection neurons are closely linked to song perception provides a physiological basis for understanding how lesions to this pathway could interfere with song recognition. Furthermore, our findings indicate that categorical responses are shown by HVCX neurons that display precise sensorimotor mirroring36, providing a link between mirror neurons and perception. In the songbird, categorically responsive sensorimotor neurons could enable auditory perception to directly guide subsequent behavior, as evident in countersinging in response to hearing a neighbor’s song20. Notably, neurons that represent learned visual categories have been detected in regions of monkey parietal cortex that have been implicated in motor planning and decision-making36. The expression of categorical responses in sensorimotor neurons in both songbirds and primates could point to an efficient means by which the nervous system determines when variations in a sensory stimulus warrant similar or different behavioral responses.

Establishing that categorical responses are expressed by striatal-projecting HVCX neurons, but not by interneurons, may also provide a clue as to how and where encoding of perceptual attributes occurs. Prior studies have shown that although HVCX neurons and HVC interneurons receive shared excitatory auditory input37,38, the auditory selectivity of interneurons more closely parallels the activity of auditory afferents to HVC37 and the highly selective auditory responses of HVCX neurons require inhibitory sculpting through interneurons39. The presence of categorical auditory responses in HVCX neurons and the absence of such responses in HVC interneurons suggest that categorical activity emerges through the local circuitry of HVCX and the processes of synaptic integration in HVCX cells. Future studies using intracellular recordings could test this idea by determining whether HVCX neurons receive broadly responsive excitatory input from which categorical responses are sculpted by local inhibition.

Precise mapping of the neuronal response boundary revealed that the action potential discharge of HVCX cells is exquisitely sensitive to changes in note duration on the millisecond time scale. Sensitivity to fine temporal features of song has been described for HVC neurons in other songbirds41, and intracellular and extracellular recordings from putative HVCX neurons reveal that their song-evoked auditory responses are strongly dependent on auditory context, such as sequences of notes or syllables42–44. This context sensitivity involves integration over hundreds of milliseconds42 and is thought to depend at least in part on circuit mechanisms that are local to HVCX38,39. Although our study cannot resolve whether categorical responsiveness to note duration arises in HVC, it does suggest that temporal aspects of song ranging from milliseconds to hundreds of milliseconds can be encoded by single neurons. The sensitivity of individual neurons to features over multiple timescales may constitute a strategy for optimizing encoding of complex stimuli that can vary in both local structure and global sequence.

**METHODS**

**Subjects.** The male swamp sparrows that were used in neural testing were collected as adults (age >1 year) from their breeding grounds. All field tests of song perception were performed in Crawford County, Pennsylvania during May and June of 2007. All experiments were performed in accordance with protocols approved by the Duke University Animal Care and Use Committee.
Song stimulus preparation. Exemplars of each song in the bird’s repertoire (2–5 song types) were recorded and digitized to be used as stimuli and in the construction of additional stimuli to test categorical responsiveness. Song notes of the type implicated previously in swamp sparrow categorical perception served as target notes for replacement (for example, note C in Fig. 3a). Each replacement note had similar spectral characteristics, but had a different duration than the target note (for example, Fig. 3a), and only one target note in the syllable was replaced in any stimulus. The possible durations of replacement notes were 4, 5, 7, 8, 16, 19, 22, 25, 27 and 31 ms. These values were chosen to replicate as closely as possible the methods used in a previous behavioral assessment of categorical perception in swamp sparrows and to resolve the location of the categorical boundary. The resulting synthetic syllable contained a replacement note in place of the target note, and all other notes and intervals were identical to the natural song. This synthetic syllable was assembled into a trill with a trill rate that was identical to the natural song and a total duration that closely approximated the duration of the natural song type (increasing or decreasing individual note durations changed the overall duration slightly). In a subset of cells, additional synthetic stimuli were presented in which either the rate of frequency modulation or the frequency bandwidth was controlled while the duration of the note was varied (Fig. 5a–c).

Experimental protocol. All experiments were performed using awake and freely behaving birds. Individual neurons were isolated using a microdrive device implanted under isoflurane anesthesia. All birds were allowed to recover for 3 d following implantation before recording began. Neurons in HVC were identified as HVCX, HVCXr or HVCYr cells either using antidromic stimulation methods or according to their electrophysiological and auditory response properties (Supplementary Methods). All neural data were amplified, filtered (band pass 500 Hz to 10 kHz) and digitized (25 kHz) to computer file (LabView) and all action potentials of individual units were discriminated using amplitude discrimination of the largest unit in a record (custom software) or discrimination on the basis of waveform characteristics (WaveClus). In both cases, single-unit isolation was verified using an interspike interval histogram to test for the presence of a refractory period.

When a single unit had been isolated and identified, stimulus presentation was immediately initiated (10-s quiet interval between each song presentation, stimuli presented in randomized order). Songs were played to the sparrow at 70 dB (peak r.m.s., A-weighted) through a speaker placed 20–35 cm away in the chamber (distance varied according to the bird’s location in the cage). Playback of the bird’s entire song repertoire, as well as their synthetic variants (see Supplementary Methods), was used to assess the auditory response of each neuron described above. Natural song types (unaltered from the original recordings) were played through a speaker placed inside the recording chamber to assess the auditory selectivity of each neuron (that is, to identify the primary song type). Extracellular recordings of action potentials in response to these stimuli were collected from 29 individual HVCX (5 birds) and 18 HVCXr units (same 5 birds).

Quantification of auditory response. Auditory activity in HVCX neurons was taken as significant if the response to any natural song stimulus exceeded the mean ± 5 s.d. of the baseline firing rate. Auditory activity in HVC interneurons was tested for significance using response strength (R < 0.05), a metric that also compares the response and baseline conditions. The responses of HVCX neurons and HVCXr cells were normalized using the strongest response of each cell to any member of the set of synthetic stimuli, enabling comparison of auditory responses across cells (for example, Fig. 3b) and birds (for example, Fig. 3c).

Neural assessment of categorical responsiveness. Noting that the dataset of auditory responses in HVCX cells tended to include transitions from strong responses (normalized responses closer to 1) to little or no responses (normalized responses closer to 0), we used interpolation to compute the note duration at which the auditory response crossed 0.5. In cases in which interpolated data crossed 0.5 multiple times (for example, Fig. 3d), the transition point in the cell was represented by the mean of the note durations corresponding to those multiple crossing points. This method revealed a putative boundary in the auditory responsiveness of these cells (~21 ms) and this value was used to test whether cells expressed categorical responses to stimuli on either side of this boundary. In off-line analysis, neural responses were compared to well-established criteria to determine whether responses were categorical in nature (Supplementary Methods). In short, responses were said to be categorical if activity was similar among members of a group of stimuli, but different between groups of stimuli, indicating a greater sensitivity to stimulus category than to the physical properties of the stimulus.

Behavioral assessment of categorical perception. Behavioral methods (Crawford County, Pennsylvania) closely paralleled a previous study of categorical perception in a New York population of swamp sparrows. Song stimuli were created in the same manner described for neurophysiological testing (see above) and were presented in pairs (4.8 ms, 8.16 ms, 16.32 ms; one pair of songs to each bird; presentation sequence randomized) through a speaker placed inside of the bird’s breeding territory. Perception of song features was assessed by quantifying territorial wing-wave responses to song stimuli in a habituation/disruption procedure that was previously employed to test auditory perception in swamp sparrows.

Note: Supplementary information is available on the Nature Neuroscience website.

AUTHOR CONTRIBUTIONS J.E.P. collected and analyzed neural and behavioral data and wrote the manuscript. S.P. supervised the project, collected behavioral data and edited the manuscript. R.C.A. collected and analyzed behavioral data. S.P. collected and analyzed songs and created acoustic stimuli. R.M. supervised the project and wrote and edited the manuscript.

Published online at http://www.nature.com/natureneuroscience/ Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/