# ARTICLES

# Precise auditory–vocal mirroring in neurons for learned vocal communication

J. F. Prather[1], S. Peters[2], S. Nowicki[1,2] & R. Mooney[1]

**Brain mechanisms for communication must establish a correspondence between sensory and motor codes used to represent the signal. One idea is that this correspondence is established at the level of single neurons that are active when the individual performs a particular gesture or observes a similar gesture performed by another individual. Although neurons that display a precise auditory–vocal correspondence could facilitate vocal communication, they have yet to be identified. Here we report that a certain class of neurons in the swamp sparrow forebrain displays a precise auditory–vocal correspondence. We show that these neurons respond in a temporally precise fashion to auditory presentation of certain note sequences in this songbird's repertoire and to similar note sequences in other birds' songs. These neurons display nearly identical patterns of activity when the bird sings the same sequence, and disrupting auditory feedback does not alter this singing-related activity, indicating it is motor in nature. Furthermore, these neurons innervate striatal structures important for song learning, raising the possibility that singing-related activity in these cells is compared to auditory feedback to guide vocal learning.**

To enable learned vocal communication, the brain must establish a correspondence between auditory and motor representations of the vocalization and use auditory information to modify vocal performance. Individual neurons that display a precise auditory–vocal correspondence could enable auditory activity to be evaluated in the context of the animal's vocal repertoire, facilitating perception. These neurons could also play an important role in vocal learning, because their motor-related activity could be compared with auditory feedback to modify vocalizations adaptively. Despite their potential importance to learned forms of vocal communication, including human speech, single neurons displaying a precise auditory–vocal correspondence have not been identified.

One major difficulty in identifying auditory–vocal attributes of individual neurons has been the challenge of recording from individual neurons in freely vocalizing animals. Another challenge in characterizing sensory and motor properties of neurons for learned vocalizations, such as human speech, is the dearth of suitable animal models. We overcame these challenges by using a lightweight chronic recording device[1] to sample neural activity in male swamp sparrows (*Melospiza georgiana*), a wild songbird that resembles humans in its dependence on auditory experience to learn its vocal communication signals[2–4]. Individual swamp sparrows sing only a few song types (range: 2–5 song types), each comprising a single trilled, multi-note syllable[5] (Supplementary Fig. 1a), simplifying exploration of the auditory and motor representations of the animal's vocal repertoire.

We focused our search in the telencephalic nucleus HVC, a structure necessary for singing[6] and normal song perception[7] and where high-level motor and auditory representations of birdsong have been detected[8–12]. HVC contains two distinct populations of projection neurons[13], including one (HVC$_{RA}$) that innervates song premotor neurons in the robust nucleus of the arcopallium (RA)[6] and another (HVC$_X$) that innervates a striatal region of the avian basal ganglia[14] (area X[6]) important to song learning and perception[15,16] (Supplementary Fig. 1b). Multiunit recordings from the HVC of awake songbirds have detected song-related auditory and motor activity[17], but whether single neurons display both types of activity remains

unknown. Furthermore, single neurons downstream of HVC, in the song premotor nucleus RA, exhibit similar patterns of singing-related and auditory activity, but auditory activity was evident only when the bird was asleep[18], making it difficult to reconcile this auditory activity with a possible role in communication. To test whether individual HVC neurons display similar patterns of auditory and singing-related activity, we recorded from identified projection neurons in the HVC of awake and freely behaving adult male swamp sparrows during auditory presentation of birdsong and during singing (Supplementary Fig. 1b).

## Auditory properties of identified HVC neurons

We probed auditory responses of identified HVC neurons by playing a variety of song stimuli, including the bird's own song types and songs of other swamp sparrows, through a loudspeaker near the bird's perch. A substantial subset of HVC$_X$ neurons (21 of 60 HVC$_X$ cells, 7 birds) responded robustly to song playback (Fig. 1), whereas HVC$_{RA}$ neurons were entirely unresponsive (16 HVC$_{RA}$ cells, 5 birds; Supplementary Fig. 1c). In a substantial proportion of responsive HVC$_X$ neurons (16 of 21 cells), auditory activity was selectively evoked by acoustic presentation of only one song type in the bird's repertoire, defined as the 'primary song type', and not by other swamp sparrow songs chosen at random (Fig. 1, Supplementary Fig. 1d, e). The primary song type varied among cells from the same bird, as expected given that each bird produces several song types. Because swamp sparrow song types consist of one syllable trilled many times (Supplementary Fig. 1a), action potential activity evoked throughout song presentation could be plotted as a response to many presentations of a single syllable (see below). This arrangement revealed that action potential activity in HVC$_X$ neurons occurred at a precise phase relative to syllable onset (s.d. of action potential latency: $18.34 \pm 14.33$ ms, or $15.05 \pm 10.85\%$ of syllable duration, $N = 21$ cells) and was both temporally sparse (action potentials per syllable in which a response occurred: $1.55 \pm 0.49$; action potential burst rate: $133 \pm 63$ Hz, $N = 21$ cells) and reliable (probability of activity per syllable: $0.64 \pm 0.18$, $N = 21$ cells). Thus,

---

[1]Department of Neurobiology, Duke University Medical Center, [2]Department of Biology, Duke University, Durham, North Carolina 27710, USA.

$HVC_X$ neurons display auditory responses highly selective in the stimulus domain, typically being activated by only one song type in the bird's repertoire. These auditory responses also are sparse in the time domain, occurring at a precise phase in the syllable of the effective song type.

### Individual neurons are active during listening and singing

To investigate whether auditory $HVC_X$ neurons also were active during singing, we relied on the tendency of swamp sparrows to countersing (5 birds, 555 cases of countersong)—this is a territorial singing behaviour triggered by presentation of either the bird's own songs or those of other swamp sparrows (Fig. 2a–c). We exploited this antiphonal behaviour to rapidly assess the auditory and singing-related activity of single neurons in the context of communication. Individual $HVC_X$ neurons could be active during both listening and singing (Fig. 2a–c; $N = 7$ cells, 3 birds). Moreover, the most robust singing-related activity in each $HVC_X$ cell occurred in association with the primary song type, as defined using auditory stimulus presentation (Supplementary Fig. 1e). Notably, the mean timing of singing-related activity, plotted relative to syllable onset, was the same as the mean timing of activity evoked by presentation of the same song type when the bird was not singing (Fig. 3). An auditory–vocal correspondence of this sort was observed in every $HVC_X$ neuron for which we were able to record activity during both singing and song playback. Further parallels between singing and auditory activity of $HVC_X$ cells were that action potential activity recruited during singing was reliable (probability of activity per syllable: $0.91 \pm 0.08$, $N = 7$ cells), persistent throughout the entire song (for example, Fig. 2a, c), and restricted to a limited phase relative to syllable onset (Fig. 3a–c; s.d. of action potential latency: $7.21 \pm 3.02$ ms, $N = 7$ cells). One difference in $HVC_X$ activity between the singing and listening states was that the singing-related activity involved short bursts of action potentials, whereas auditory-evoked activity typically consisted of single action potentials (Fig. 3a, b; action potentials per syllable: $1.27 \pm 0.21$ auditory, $2.89 \pm 0.67$ singing, $P = 0.004$; action potential burst rate: $148 \pm 64$ Hz auditory, $278 \pm 80$ Hz singing, $P = 0.01$; paired $t$-tests, $N = 7$ cells, 3 birds). In summary, $HVC_X$ neurons display highly similar, temporally precise

patterns of activity while hearing and singing the primary song type, suggestive of a precise sensorimotor correspondence (Fig. 3d).

### Singing-related activity is a corollary discharge

In a simple model of sensorimotor correspondence, motor-related activity should occur before sensory feedback elicited by the action. In this context, the similar action potential timing we observed in $HVC_X$ cells during singing and listening raises the possibility that the activity during singing was due to auditory feedback. Alternatively, singing-related activity may constitute a corollary discharge of the song motor activity, perhaps providing a motor estimation of auditory feedback[19]. Three observations indicated that $HVC_X$ activity during singing was motor-related corollary discharge.

First, we noted that background multiunit activity could increase before singing of any song type (for example, Fig. 2a, c) and, during this 'warm-up' period, the isolated $HVC_X$ cell's auditory responses to the primary song type were suppressed (Fig. 4a, Supplementary Fig. 2a; $N = 25$ occurrences, 5 cells, 2 birds). This suppression of auditory activity suggests that $HVC_X$ neurons switch from an auditory state to an auditory-insensitive motor state several hundred milliseconds
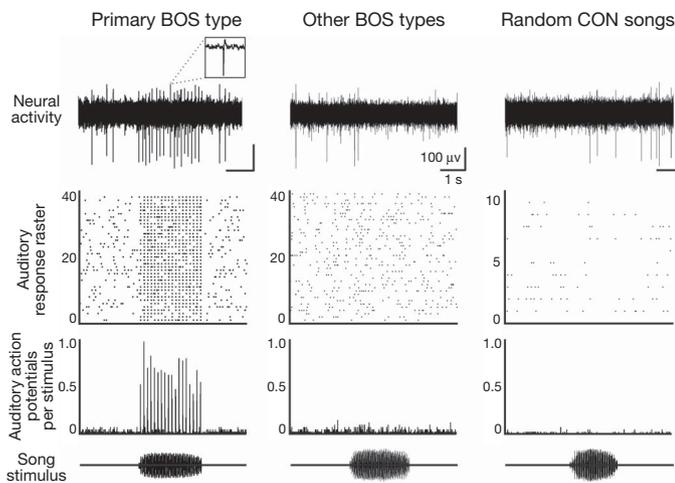


**Figure 1 | In freely behaving swamp sparrows, $HVC_X$ neurons respond selectively to one song type in the bird's repertoire.** A single song type in the repertoire of the bird's own songs (BOS) typically evoked an auditory response (left, the 'primary BOS type'), whereas other BOS types (middle) and randomly selected songs of conspecific birds (CON, right) were ineffective stimuli (top row, raw data recorded from an $HVC_X$ neuron during a single stimulus presentation; second row, response raster for multiple presentations; third row, peri-stimulus time histogram (PSTH), 10 ms bin size; bottom row: stimulus oscillogram). Audio files available as Supplementary Information.
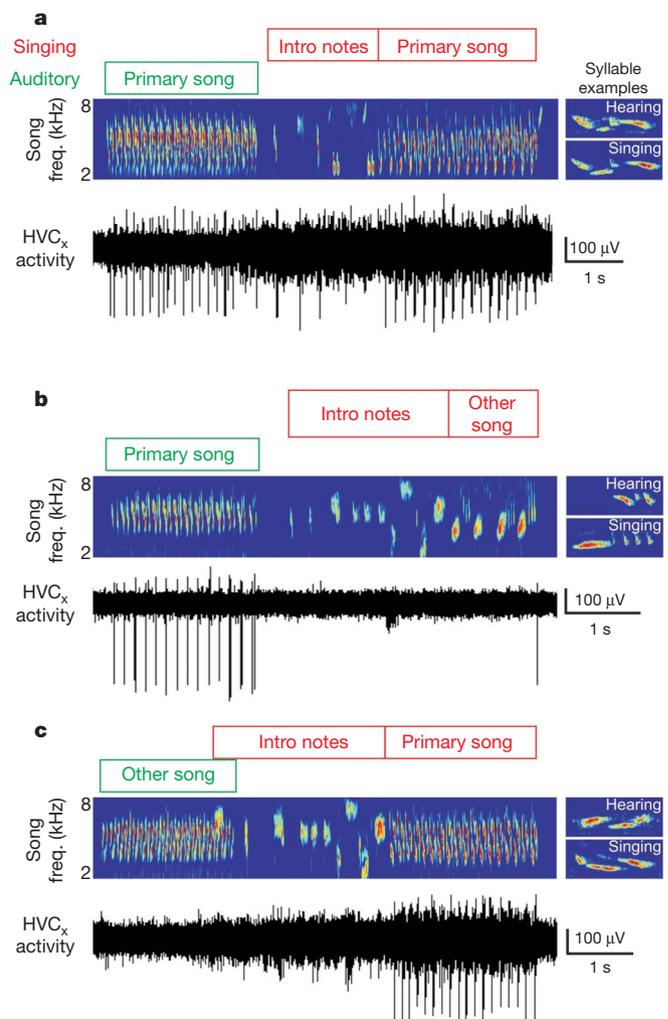


**Figure 2 | Countersinging in response to song presentation reveals auditory and singing-related activity of $HVC_X$ neurons in the context of communication. a**, In 'matched countersinging'[41], an $HVC_X$ neuron was active when the bird heard (green box, left) or sang (red box, right) the primary song type. **b, c**, In 'unmatched countersinging', another $HVC_X$ neuron was active when the bird heard (**b**) or sang (**c**) the primary song type but was silent as the bird sang (**b**) or heard (**c**) other song types. (In **a, b, c**: top, spectrogram of the acoustic signal; bottom, corresponding electrophysiological recording.) Audio files available as Supplementary Information.

before the onset of singing ($601 \pm 288$ ms, $N = 5$ cells, 2 birds). Furthermore, in cases where playback of the primary song type began immediately after the bird stopped singing, auditory-evoked activity remained briefly suppressed ($\sim$250 ms, much shorter than reported for HVC multi-unit activity in other species[17], Supplementary Fig. 2b). Second, there was often a 'secondary song type' for which singing-related activity in an $HVC_X$ neuron was evident, even though playback of that song type evoked no response from that cell (Supplementary Fig. 2c). Third, in several instances singing of either the primary or secondary song type overlapped playback ($N = 4$ cells, 3 birds), distorting auditory feedback. However, the singing-related activity pattern was unaffected by such distortion (Fig. 4b, d, e, Supplementary Fig. 2d, e; probability of neural activity: $P = 0.59$; mean latency: $P = 0.56$; action potentials per syllable: $P = 0.94$; paired $t$-tests, $N = 4$ cells, 3 birds). Furthermore, during the period of overlap, neural activity was locked precisely to features of the syllable being sung but not to the playback syllable, even when the two syllables were of the same type (Fig. 4b, d, Supplementary Fig. 2d, e; $N = 4$ cells, 3 birds). In contrast, recordings made in the absence of singing revealed that auditory activity normally evoked by presentation of the primary song type was strongly attenuated when a phase-delayed copy of the primary song type or another song

was simultaneously presented through another speaker (Fig. 4c–e, $N = 8$ cells, 4 birds). Together, these observations indicate that singing-related activity in $HVC_X$ cells is due to corollary discharge
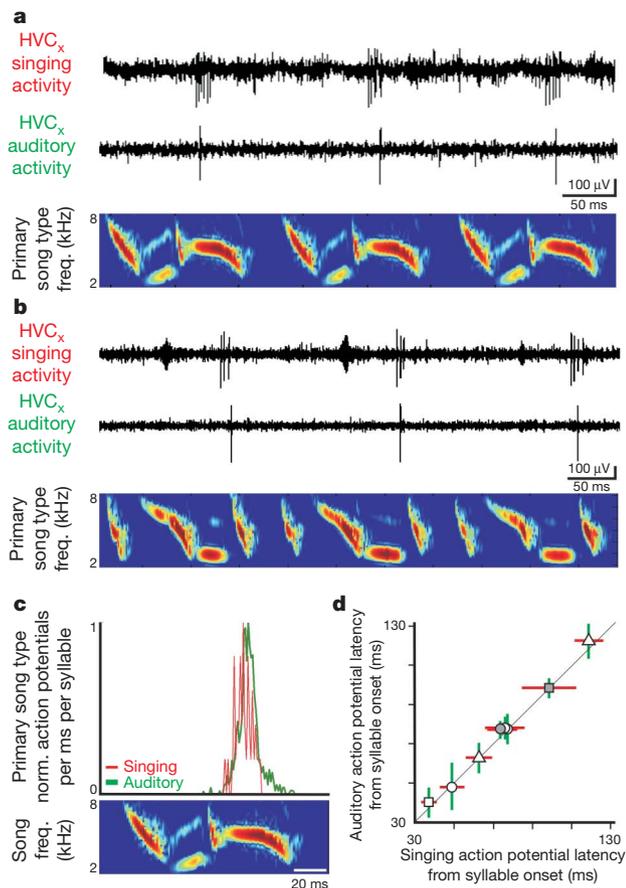


**Figure 3 | $HVC_X$ neurons exhibit a precise sensorimotor correspondence.** **a, b,** Singing-related and auditory activity (respectively top and middle row) in association with several syllables of the primary song type (shown as a spectrogram, bottom row). **c, d,** Action potential timing was quite similar in the singing and hearing states, both within and across $HVC_X$ cells. (In **c**, $P = 0.50$, paired $t$-test. In **d**, $N = 9$ song types, 7 cells, 3 birds, mean $\pm$ s.d.; shaded symbols, cells in **a, b**; regression: $P < 0.01$, $R^2 = 0.99$, slope = 1.05, intercept $= -2.90$; diagonal line represents identity.) Interestingly, two cells that were active in association with two song types in the bird's repertoire displayed a precise auditory–vocal correspondence for both song types (triangles and squares indicate paired song types; see also Supplementary Fig. 1d).
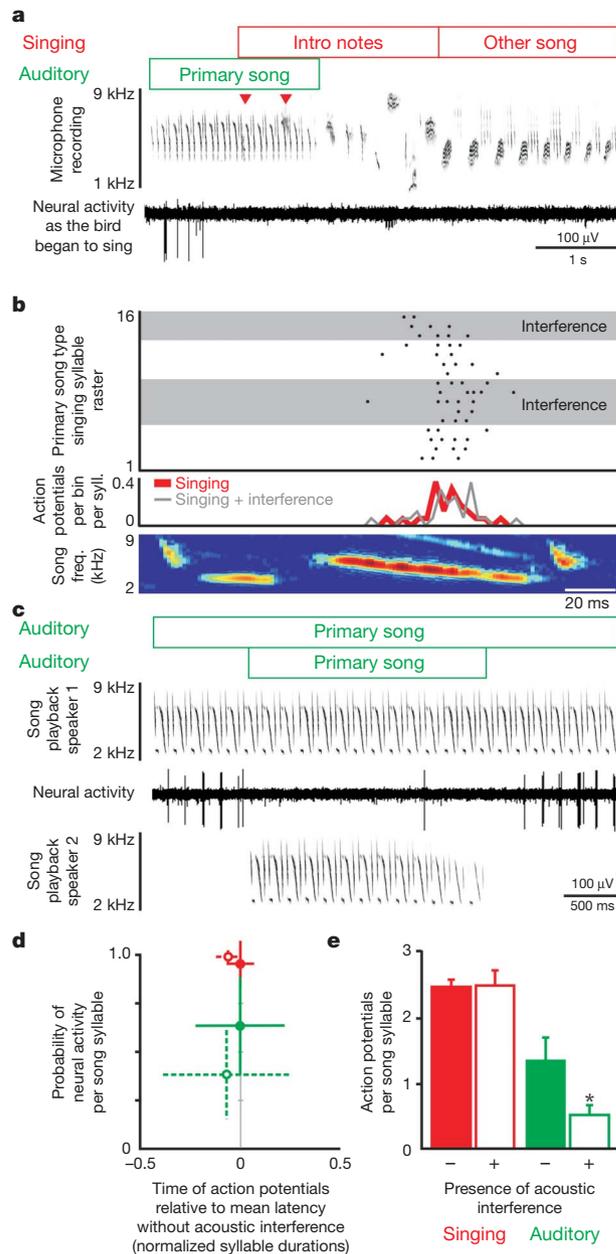


**Figure 4 | Action potentials in $HVC_X$ neurons during singing are a corollary discharge of song motor activity.** **a,** Auditory response to the primary song type was suppressed before and during singing (arrows, singing of introductory notes; top, spectrogram of microphone recording; bottom, raw data; see also Supplementary Fig. 2a). **b,** Distorted auditory feedback (DAF; shaded regions) as the bird sang the primary song type did not affect either the probability of occurrence ($P = 1.00$) or the timing of action potentials ($P = 0.52$, paired $t$-tests; second row, PSTH, 5 ms bin size; bottom, spectrogram of the vocalized syllable). **c,** Auditory response (middle) to the primary song type (top) was suppressed when the primary song type was played through a second speaker at a pseudorandom phase delay (bottom). **d,** Acoustic distortion strongly attenuated auditory activity but not singing-related activity (auditory, green, $P = 0.01$, $N = 8$ cells, 4 birds; singing, red, $P = 0.59$, $N = 4$ cells, 3 birds). Action potential timing was unaffected by distortion in either state (auditory, $P = 0.15$; singing, $P = 0.56$, mean $\pm$ s.d.; solid lines, control; dotted lines, distortion present). **e,** Acoustic distortion reduced the number of action potentials per syllable in the auditory state (green) but not in the singing (red) state (auditory, $P = 0.02$, $N = 8$ cells; singing, $P = 0.93$, paired $t$-tests in all cases; $N = 4$ cells, mean $\pm$ s.e.).

rather than an auditory feedback signal and that HVC$_X$ cells are gated to exist in purely auditory or motor states.

### Sensorimotor correspondence in another species

To investigate whether the sensorimotor correspondence seen in swamp sparrow HVC$_X$ neurons generalized to HVC$_X$ cells of other songbirds, we recorded from HVC$_X$ cells in Bengalese finches (*Lonchura striata domestica*). Adult Bengalese finch song is highly sensitive to distortion of auditory feedback[20,21], thus affording a more rigorous test of the idea that singing-related activity of HVC$_X$ cells is due to motor corollary discharge. As observed in swamp sparrows, HVC$_X$ cells in the awake Bengalese finch responded selectively to playback of the bird's own song (Supplementary Fig. 3a, $N = 16$ HVC$_X$ cells, 2 birds). These auditory responses were highly phasic, occurring in association with certain syllables in the song phrase (Supplementary Fig. 3a). In direct parallel with our observations in swamp sparrows, HVC$_X$ cells in Bengalese finches showed singing-related activity, and auditory and singing-related activities were aligned relative to syllable onset (Supplementary Figs 3a, 4a, $N = 6$ cells, 2 birds). This singing-related activity was unaffected by distorted auditory feedback (Supplementary Figs 3b, 4a–c, $N = 5$ cells,

2 birds), while auditory activity was strongly suppressed when two copies of the effective song phrase were presented simultaneously with variable phase offset (Supplementary Fig. 3c, $N = 6$ cells, 2 birds). Therefore, singing-related activity of HVC$_X$ neurons is dominated by motor corollary discharge in both swamp sparrows and Bengalese finches. Thus the capacity of HVC$_X$ cells to exhibit a precise sensorimotor correspondence and switch rapidly between auditory and motor states may constitute a general mechanism underlying learned vocal communication in songbirds.

### Auditory responses extend to other birds' songs

For HVC$_X$ neurons to facilitate communication, their sensory responsiveness must extend to other birds' songs. In initial experiments in swamp sparrows, we found that HVC$_X$ cells were unresponsive to other swamp sparrow songs chosen at random (Fig. 1). These conspecific songs may have failed to evoke responses because HVC$_X$ cells respond exclusively to self-generated vocalizations, or because the songs we chose lacked certain necessary features. Consistent with the idea that HVC$_X$ cells in swamp sparrows respond to specific features, all HVC$_X$ cells responded at a precise phase of the syllable presentation. Because note sequences are important features for some auditory HVC neurons[9,22], we presented artificial trilled syllables containing the primary song type notes in their natural or reverse order (Fig. 5a). Almost all HVC$_X$ cells tested in this manner (12/14 cells, 7 birds) responded to only the naturally occurring sequence, indicating that a sequence of at least two notes was necessary to elicit an auditory response. We then tested whether HVC$_X$ neurons would respond to other swamp sparrow songs containing note sequences similar to those in the primary song type. We found that a swamp sparrow song with a note sequence similar to that in the primary song type could drive auditory responses in HVC$_X$ neurons (Fig. 5b; $N = 14$ cells including 3 in which both singing and auditory data were collected, 7 birds; $N = 19$ effective stimuli). In some cases, a conspecific song could evoke a more robust response than that elicited by the primary song type (range: 1.00–1.64 conspecific response normalized to primary song type response; $N = 5$ cells, 4 birds). When exemplar syllables of the primary song type and an effective song of another sparrow were plotted relative to the average action potential latency for each syllable, the note sequences in the two syllables were aligned (Fig. 5b). Thus, the selective auditory responsiveness of HVC$_X$ cells extends to similar vocal sequences produced by other birds, making auditory–vocal HVC$_X$ neurons well suited to a role in communication.

The ability of HVC$_X$ neurons to respond to other birds' songs and to display an auditory–motor correspondence could facilitate vocal communication in two ways. First, when the sender's vocalizations activate the receiver's auditory–vocal HVC$_X$ neurons, those vocalizations could be compared to an internal representation of the receiver's vocal gestures, enabling perceptual categorization of songs in the context of the receiver's vocal repertoire[23]. Second, auditory activation of HVC$_X$ neurons by other birds' songs could provide a template for subsequent movement, enabling the animal to select a vocalization from its repertoire that matches songs of its neighbours. In many regards, auditory–vocal HVC$_X$ cells are similar to visual–motor 'mirror neurons' in the monkey frontal cortex[24–26] that are hypothesized to play a role in perception of communication gestures[27–29], including human speech[30,31]. In that light, the precise temporal alignment of auditory and vocal activity in HVC$_X$ cells suggests that auditory–vocal mirror neurons express an additional mode of sensorimotor correspondence not previously reported for visual–motor mirror neurons. An important remaining question is whether auditory activity in HVC$_X$ cells is related to the bird's perception of songs, as predicted for mirror neurons.

Beyond serving a perceptual role, auditory–vocal HVC$_X$ cells could have a role in vocal learning. During singing, HVC$_X$ neurons transmit song corollary discharge sufficiently delayed to mimic auditory feedback associated with the vocalization. This delay probably arises
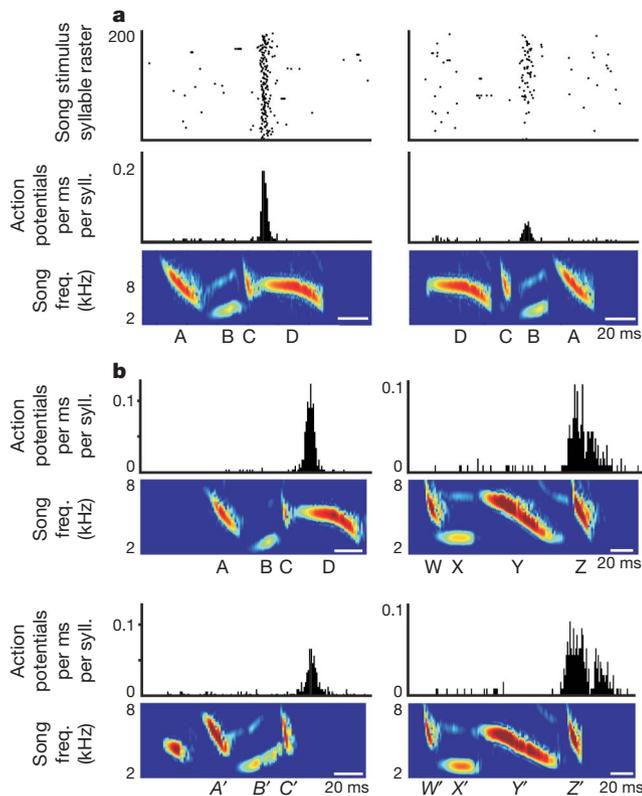


**Figure 5 | Swamp sparrow HVC$_X$ neurons respond to note sequences in the primary song type and to similar note sequences in other swamp sparrows' songs. a**, An HVC$_X$ neuron responded robustly to the primary song type with the notes in the natural sequence (left) but weakly or not at all when the notes were in the reverse order (right). Top row, syllable raster; middle row, PSTH, 1 ms bin size; bottom row, syllable spectrogram. **b**, HVC$_X$ neurons (left, cell 1; right, cell 2) responded to note sequences in the primary song type (top pair of histogram and spectrogram) and similar sequences in another (conspecific) sparrow's song (bottom pair). Histogram, PSTH, 1 ms bin size; Spectrogram, syllable spectrogram, notes labelled individually. (19 of 23 similar conspecific (CON) songs evoked a response; CON responses normalized to the primary song type response; effective stimuli, 0.87 ± 0.32; ineffective stimuli, 0.28 ± 0.16, mean ± s.d., $P < 0.01$, paired $t$-test, range of CON responses, 0–1.64; data not shown). Alignment of syllables in the primary song type (top, spectrogram) and effective conspecific song (bottom) using the mean timing of auditory activity revealed similar spectrotemporal features.

when song premotor activity of $HVC_{RA}$ cells is relayed by interneurons to $HVC_X$ cells[32]. Inhibitory interneurons in HVC help shape the temporally precise auditory responses of $HVC_X$ cells to song[10,33], suggesting that inhibitory synapses onto $HVC_X$ cells play an important role in establishing the observed sensorimotor correspondence. In both invertebrates[34] and vertebrates[35], corollary discharge of central motor commands can serve as an estimate of the anticipated sensory feedback. In HVC, this arrangement could provide a motor-based estimate of auditory feedback[19], with the useful outcome that differences between the motor estimate and the actual feedback could be used to guide song learning. If this model obtains in songbirds, then $HVC_X$ cells either transmit estimated feedback to a downstream comparator or are the site of comparison. In support of the idea that the comparator lies downstream of HVC, we observed that the singing-related activity of $HVC_X$ neurons was insensitive to distorted auditory feedback over acute timescales (Fig. 4); such insensitivity has also been described for $HVC_X$ cells during juvenile song learning[36]. Alternatively, $HVC_X$ cells may serve as comparators in which corollary discharge typically overwhelms auditory feedback signals, a mismatch that could facilitate song maintenance. Future studies can determine whether $HVC_X$ neurons are the site of auditory–vocal comparison by recording from those cells while presenting distorted auditory feedback over a timescale sufficient to induce vocal plasticity.

Finally, because $HVC_X$ neurons innervate striatal structures[14,37] important for song learning and perception[15,16], the coding strategy employed by $HVC_X$ neurons to represent vocal sequences may have implications for learning and perception of speech in humans. In the human brain, cortical neurons similar to $HVC_X$ auditory–vocal neurons could transmit speech-related auditory and motor information to striatal regions implicated in speech development[38,39]. Furthermore, auditory–vocal mirror neurons with properties similar to the $HVC_X$ cells described here could bind sensory and motor features of distinct vocal gestures, providing an efficient substrate for rapid decoding and encoding of speech[30,31].

## METHODS SUMMARY

**Song behaviour.** Birds' song types were recorded in a semi-anechoic chamber, digitized at 25 kHz and saved onto a computer hard drive to be used as stimulus songs. Individual note types were classified by S.P. using established criteria[5]. Similar songs were defined as those containing the same sequence of note categories as in the primary song type. Conspecific songs capable of driving auditory responses expressed a range of spectral similarity to the primary song type, as defined using cross-correlation of the two syllables (correlation value range: 0.17–0.78). Audio files of songs in Figs 1, 2 and 5 are available as Supplementary Information.

**Electrophysiological recordings and analysis.** Individual neurons were recorded extracellularly in awake and freely behaving birds. All HVC neurons from which both auditory and singing data were obtained were identified antidromically using stimulation in area X. Action potentials of individual neurons were discriminated by amplitude (custom software) or on the basis of waveform characteristics (WaveClus[40]), and unit isolation was verified by the presence of a refractory period in the interspike interval histogram. To assess auditory selectivity of an isolated neuron, the bird's own song types and other birds' songs were presented through a loudspeaker in the sound attenuating chamber in which the bird was housed. Singing-related activity was recorded along with the bird's vocalization. Rasters and histograms of action potential activity were constructed by aligning action potentials relative to the beginning of the associated song or syllable. Activity during song presentation or singing was compared against the cell's background firing rate using an activity histogram; any value exceeding the mean background rate plus 5 s.d. was deemed significant. Responses to other birds' songs were normalized to the response to the primary song type, with the criterion for effective stimuli being >0.5.

**Full Methods** and any associated references are available in the online version of the paper at www.nature.com/nature.

1. Fee, M. S. & Leonardo, A. Miniature motorized microdrive and commutator system for chronic neural recording in small animals. *J. Neurosci. Methods* 112, 83–94 (2001).
2. Marler, P. & Tamura, M. Culturally transmitted patterns of vocal behavior in sparrows. *Science* 146, 1483–1486 (1964).
3. Konishi, M. The role of auditory feedback in the control of vocalization in the white-crowned sparrow. *Z. Tierpsychol.* 22, 770–783 (1965).
4. Marler, P. & Peters, S. Sparrows learn adult song and more from memory. *Science* 213, 780–782 (1981).
5. Marler, P. & Pickert, R. Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Anim. Behav.* 32, 673–689 (1984).
6. Nottebohm, F., Stokes, T. M. & Leonard, C. M. Central control of song in the canary, *Serinus canarius*. *J. Comp. Neurol.* 165, 457–486 (1976).
7. Gentner, T. Q., Hulse, S. H., Bentley, G. E. & Ball, G. F. Individual vocal recognition and the effect of partial lesions to HVc on discrimination, learning, and categorization of conspecific song in adult songbirds. *J. Neurobiol.* 42, 117–133 (2000).
8. Hahnloser, R. H., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419, 65–70 (2002).
9. Margoliash, D. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* 3, 1039–1057 (1983).
10. Mooney, R. Different subthreshold mechanisms underlie song selectivity in identified HVc neurons of the zebra finch. *J. Neurosci.* 20, 5420–5436 (2000).
11. Yu, A. C. & Margoliash, D. Temporal hierarchical control of singing in birds. *Science* 273, 1871–1875 (1996).
12. Mooney, R., Hoese, W. & Nowicki, S. Auditory representation of the vocal repertoire in a songbird with multiple song types. *Proc. Natl Acad. Sci. USA* 98, 12778–12783 (2001).
13. Wild, J. M., Williams, M. N., Howie, G. J. & Mooney, R. Calcium-binding proteins define interneurons in HVC of the zebra finch (*Taeniopygia guttata*). *J. Comp. Neurol.* 483, 76–90 (2005).
14. Farries, M. A. & Perkel, D. J. A telencephalic nucleus essential for song learning contains neurons with physiological characteristics of both striatum and globus pallidus. *J. Neurosci.* 22, 3776–3787 (2002).
15. Scharff, C. & Nottebohm, F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* 11, 2896–2913 (1991).
16. Scharff, C., Nottebohm, F. & Cynx, J. Conspecific and heterospecific song discrimination in male zebra finches with lesions in the anterior forebrain pathway. *J. Neurobiol.* 36, 81–90 (1998).
17. McCasland, J. S. & Konishi, M. Interaction between auditory and motor activities in an avian song control nucleus. *Proc. Natl Acad. Sci. USA* 78, 7815–7819 (1981).
18. Dave, A. S. & Margoliash, D. Song replay during sleep and computational rules for sensorimotor vocal learning. *Science* 290, 812–816 (2000).
19. Troyer, T. W. & Doupe, A. J. An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J. Neurophysiol.* 84, 1204–1223 (2000).
20. Okanoya, K. & Yamaguchi, A. Adult bengalese finches (*Lonchura striata* var *domestica*) require real-time auditory feedback to produce normal song syntax. *J. Neurobiol.* 33, 343–356 (1997).
21. Woolley, S. M. & Rubel, E. W. Bengalese finches *Lonchura striata domestica* depend upon auditory feedback for the maintenance of adult song. *J. Neurosci.* 17, 6380–6390 (1997).
22. Lewicki, M. S. Intracellular characterization of song-specific neurons in the zebra finch auditory forebrain. *J. Neurosci.* 16, 5855–5863 (1996).
23. Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. Perception of the speech code. *Psychol. Rev.* 74, 431–461 (1967)].
24. Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain* 119, 593–609 (1996).
25. Rizzolatti, G. & Craighero, L. The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192 (2004).
26. Ferrari, P. F., Gallese, V., Rizzolatti, G. & Fogassi, L. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714 (2003).
27. Iacoboni, M. *et al.* Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol.* 3, e79 (2005).
28. Rizzolatti, G., Fogassi, L. & Gallese, V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Rev. Neurosci.* 2, 661–670 (2001).
29. Iacoboni, M. *et al.* Cortical mechanisms of human imitation. *Science* 286, 2526–2528 (1999).
30. Rizzolatti, G. & Arbib, M. A. Language within our grasp. *Trends Neurosci.* 21, 188–194 (1998).
31. Arbib, M. A. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* 28, 105–124 125–167 (2005).
32. Mooney, R. & Prather, J. F. The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *J. Neurosci.* 25, 1952–1964 (2005).
33. Rosen, M. J. & Mooney, R. Inhibitory and excitatory mechanisms underlying auditory responses to learned vocalizations in the songbird nucleus HVC. *Neuron* 39, 177–194 (2003).
34. Poulet, J. F. & Hedwig, B. The cellular basis of a corollary discharge. *Science* 311, 518–522 (2006).
35. Bell, C. C. An efference copy which is modified by reafferent input. *Science* 214, 450–453 (1981).

36. Kozhevnikov, A. A. & Fee, M. S. Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97,** 4271–4283 (2007).

37. Perkel, D. J., Farries, M. A., Luo, M. & Ding, L. Electrophysiological analysis of a songbird basal ganglia circuit essential for vocal plasticity. *Brain Res. Bull.* **57,** 529–532 (2002).

38. Vargha-Khadem, F., Gadian, D. G., Copp, A. & Mishkin, M. FOXP2 and the neuroanatomy of speech and language. *Nature Rev. Neurosci.* **6,** 131–138 (2005).

39. Lai, C. S., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413,** 519–523 (2001).

40. Quiroga, R. Q., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **16,** 1661–1687 (2004).

41. Hyman, J. Countersinging as a signal of aggression in a territorial songbird. *Anim. Behav.* **65,** 1179–1185 (2003).

## METHODS

**Swamp sparrows.** All procedures were in compliance with recommendations of Duke University Animal Care and Use Committee and state and federal regulations governing the capture and use of wild birds. Birds were caught with mist nets as adults (age > 1 yr) either on winter grounds in Orange County, North Carolina, or on their summer breeding grounds in Crawford County, Pennsylvania. Birds were housed individually throughout their time in the laboratory, both before and during experimentation. Birds were provided with seed and water *ad libitum* and were given a regular supplement of mealworms. Males were identified either by external morphology (breeding season) or by molecular marker techniques[42] (out of season), and females were released. Prior to implantation of the stimulus and recording devices, birds were subjected to gradually lengthening photoperiod (1 h week$^{-1}$ from 9:15 up to 15:9 L:D cycle) meant to simulate the onset of the spring breeding season, the only time of year when swamp sparrows sing robustly. This change in photoperiod, combined with a subcutaneous implant of testosterone[43], was sufficient to induce the birds to sing. Birds were recorded in a semi-anechoic chamber (recorded using a Sony TCM 5000 EV recorder and Shure SM57 microphone), and many examples of song (typically >100) were recorded from each bird to ensure that the bird's full repertoire was sampled (2 to 5 song types). Although the exact age of birds we used for these experiments was unknown, all songs were crystallized, indicating birds were at least 1 yr old. Exemplars of each song type were digitized (25 kHz) and saved onto a computer hard drive (SIGNAL and LabView software) to be used as stimulus songs. Song stimuli consisted of natural song types and synthetic variants (for example, reverse note order) of those song types from the experimental subject and conspecific birds. Natural song types (unaltered from the original recordings) were used to assess the auditory selectivity of each neuron. Digital editing was used to create synthetic variants of the primary song type in which the notes were arranged in either the natural or reverse order, using the same internote intervals as in the natural song. Copies of this syllable were then concatenated to form songs with the same intersyllable intervals and total song duration as in the natural song.

**Bengalese finches.** Procedures were generally the same as those described for swamp sparrows, except that birds were raised in our aviary (15:9 L:D cycle) housed in communal cages. Subject birds were adult males >155 days of age; males were distinguished from females by males' expression of song. Because Bengalese finch songs have variable syntax from one song bout to another, several variants of song from the subject bird were used to probe the auditory response of each neuron.

**Microdrive implantation surgery.** Neurons were sampled using a miniaturized micromanipulation device[1] in awake and freely behaving birds. Several days before implantation, birds were transferred from their housing cage to the recording chamber, a sound-attenuating box (Acoustic Systems) where they would reside throughout experimentation. During implantation, adult male swamp sparrows were anaesthetized using isoflurane (inhalation, 1–3% in 100% $O_2$) and placed in a stereotaxic device. A small incision was made in the skin overlying the skull, and the outer leaflet of bone was removed over HVC, area X and RA. A small craniotomy (approximately $300 \times 300 \mu m$) was made in the inner leaflet over area X, and a small custom-made bipolar stimulus electrode (J.F.P.) was inserted to the proper depth. The implant site was covered with a sterile film and the electrode was secured using dental cement. With the electrode in area X firmly secured, the head was repositioned and the same implant procedure was repeated to place a bipolar stimulus electrode in RA. With both stimulus electrodes firmly in place, another small craniotomy was made directly over HVC. HVC was located by passing brief (~100 μs) current pulses through the stimulating electrodes in area X and RA to generate antidromic activity in HVC, and the boundaries of HVC were defined using a sterilized extracellular electrode (Carbostar 1, Kation Scientific) to observe the extent of the region expressing the resultant antidromic 'hash'. The microdrive recording device was implanted so that the recording electrodes were initially positioned slightly dorsal of HVC. The microdrive was secured to the skull using dental cement (microdrive ~1.2 g including dental cement, birds ~16 g), and the incision site was closed using surgical skin adhesive (Vetbond). The bird was monitored closely until it was fully recovered, typically <15 min. After the recording session was complete (1–5 weeks), the bird was deeply anaesthetized with equithesin, perfused transcardially with saline and then 4% paraformaldehyde, and the brain was processed histologically. All electrode positions were verified at the end of each experiment using Nissl-stained sagittal sections (thickness 75 μm).

**Experimental protocol.** Birds were allowed to recover for three days following the implantation procedure before recording began. During electrophysiological recording, microdrive electrodes were slowly advanced into HVC while weak electrical stimulation was delivered to the stimulus electrodes in either area X or RA (100 μs pulses, ~100 μA). The boundaries of HVC could be reliably identified by observing where antidromic activity was evident. Once an electrode was positioned in HVC, the electrode was advanced very slowly so that antidromically-evoked action potentials of individual neurons could be identified. All neural data were amplified, filtered (band pass 500 Hz to 10 kHz), and digitized (25 kHz) to computer file (LabView).

Action potentials of individual units were discriminated using amplitude discrimination of the largest unit in a record (custom software) or discrimination based on waveform characteristics (WaveClus[40]). In both cases, single unit isolation was verified using an interspike interval histogram to test for the presence of a refractory period. Individual units were identified using antidromic stimulation via the electrodes placed in area X and RA or by their characteristic electrophysiological response properties, although all cells from which both auditory and singing data were obtained were identified antidromically. In antidromic identification, $HVC_X$ units displayed fixed-latency action potential responses to stimulation in area X but no response to stimulation in RA. In contrast, $HVC_{RA}$ units displayed fixed-latency action potential responses to stimulation in RA but not in area X. Each of these classes of projection neuron could be distinguished from HVC interneurons, which expressed variable-latency responses to stimulation in either RA[8] or area X and occasionally to stimulation at both sites.

When a single unit had been isolated and identified, song playback of each song type in the bird's repertoire 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), and a microphone in the chamber was used to record auditory stimuli and the bird's vocalizations. Playback of the bird's entire song repertoire, as well as songs of conspecific birds and synthetic variants of the bird's own song types were used to assess the auditory response of each neuron described in the main text. Auditory responsiveness to songs of other swamp sparrows was assessed using conspecific songs that contained some or all of the same sequence of note types[5] as in the syllable of the corresponding song in the bird's repertoire. Conspecific songs expressed a range of spectral similarity to the repertoire song, as defined using cross-correlation of the two syllables (correlation value range: 0.17–0.78), and all conspecific stimuli were selected before any neural recording.

We enforced the following criteria to qualify a neuron as suitable for further analysis: (1) action potentials must have been reliably distinguishable as belonging to only a single unit, (2) all song types in the bird's repertoire must have been presented as auditory stimuli, and (3) the bird must have sung at least once following implantation of the recording device and stimulus electrodes (this ensured that all birds were in roughly similar behavioural states). Extracellular recordings were collected from 60 individual $HVC_X$ units (7 birds) and 16 individual $HVC_{RA}$ units (5 birds, a subset of the 7 birds in which $HVC_X$ cells were sampled) that met these criteria.

Singing-related activity of $HVC_X$ neurons was recorded along with the song itself, using either a voice-triggered recording set-up or by evoking countersinging with song playback (see text). For each bird, these songs were compared against the exemplars recorded before surgery, and in each case we noted that song structure was unchanged, consistent with the crystallized song state. Neural activity associated with singing was recorded and compared against features of the song recorded through the microphone in the recording chamber. As swamp sparrow songs were highly stereotyped from one bout to the next, no time-warping of the data was necessary to permit comparison of data collected during singing and during auditory stimulus presentation. Because Bengalese finch electrophysiological data were compared at the level of one- or two-note sequences, stereotypy on that timescale was sufficiently good that time-warping was also unnecessary for those data. In short, time-warping of the data was not performed in any of the analyses reported here.

**Data analysis.** Action potentials from individual neurons were discriminated and compared against features of either the auditory song stimulus during passive playback or features of the song recorded during singing. Song features were discerned using spectrograms generated in Matlab (Mathworks). All analyses were performed in Matlab using custom software (J.F.P. and Stefan Nenkov).

Rasters and histograms of action potential activity were constructed by aligning discriminated action potentials to the song ('whole-song' analysis, 10 ms bin size in whole-song histograms, for example, Fig. 1). Because swamp sparrow songs consist of trilled syllables separated by brief quiet intervals, an additional technique was possible wherein the onset of each syllable was detected separately and used to align action potentials that occurred in association with each syllable ('single-syllable' analysis, 1 ms histogram bin size in single-syllable histograms, for example, Fig. 3c). In both whole-song and single-syllable analyses, the onset of song during presentation of auditory stimuli was defined as the time that the stimulus presentation began, as recorded in each computer file; onset of song

during singing was computed using the spectrogram of the microphone voltage recorded as the bird sang. The onset of song (or of each syllable following a brief quiet period) was defined as the first time when a song note >10 dB louder than background could be detected. In whole-song analyses, action potential latencies were assigned relative to the onset of the song; in single-syllable analyses, action potential latencies were assigned relative to the onset of each associated song syllable.

In both the whole-song and single-syllable analyses, action potential activity during song presentation or singing was compared against the background firing rate when no stimulus was present, and the mean background rate plus 5 s.d. was taken as the threshold for significance. If the value in any bin in the peri-stimulus time histogram exceeded that threshold (accounting for bin size), the auditory response was deemed significant. In assessment of auditory responses to con-specific songs, responses were normalized using the strength of response to the primary song type in each cell. Normalized responses greater than 0.5 were considered effective stimuli, and responses less than 0.5 were considered inef-fective. Results obtained in this manner were in good agreement with visual assessment of the efficacy of an auditory stimulus.

Audio files of songs in Figs 1, 2 and 5 are available as Supplementary Information.

42. Griffiths, R., Double, M., Orr, K. & Dawson, R. A. DNA test to sex most birds. *Mol. Ecol.* **7**, 1071–1075 (1998).
43. Marler, P., Peters, S., Ball, G. F., Dufty, A. M. Jr & Wingfield, J. C. The role of sex steroids in the acquisition and production of birdsong. *Nature* **336**, 770–772 (1988).