



## Sensory constraints on birdsong syntax: neural responses to swamp sparrow songs with accelerated trill rates

J. F. Prather<sup>a</sup>, S. Peters<sup>b</sup>, R. Mooney<sup>c</sup>, S. Nowicki<sup>b,c,\*</sup>

<sup>a</sup> Department of Zoology and Physiology, University of Wyoming, Laramie, WY, U.S.A.

<sup>b</sup> Department of Biology, Duke University, Durham, NC, U.S.A.

<sup>c</sup> Department of Neurobiology, Duke University Medical Center, Durham, NC, U.S.A.

### ARTICLE INFO

#### Article history:

Received 29 August 2011

Initial acceptance 4 November 2011

Final acceptance 29 February 2012

Available online 17 April 2012

MS. number: A11-00683R

#### Keywords:

auditory

birdsong

development

HVC

*Melospiza georgiana*

sensorimotor

swamp sparrow

syntax

vocal learning

Both sensory and motor mechanisms can constrain behavioural performance. Sensory mechanisms may be especially important for constraining behaviours that depend on experience, such as learned birdsongs. Swamp sparrows, *Melospiza georgiana*, learn to sing by imitating the song of a tutor, but they fail to accurately imitate artificial tutor songs with abnormally accelerated trills, instead singing brief and rapid trills interrupted by silent gaps. This 'broken syntax' has been proposed to arise from vocal-motor limitations. Here we consider whether sensory limitations exist that could also contribute to broken syntax. We tested this idea by recording auditory-evoked activity of sensorimotor neurons in the swamp sparrow's brain that are known to be important for the learning, performance and perception of song. In freely behaving adult sparrows that sang songs with normal syntax, neurons were detected that exhibited precisely time-locked activity to each repetition of the syllable in a trill when presented at a natural rate. Those cells failed to faithfully follow syllables presented at an accelerated rate, however, and their failure to respond to consecutive syllables increased as a function of trill rate. This 'flickering' auditory representation in animals performing normal syntax reveals a central constraint on the sensory processing of rapid trills. Furthermore, because these neurons are implicated in both song learning and perception, and because auditory flickering began to occur at accelerated trill rates previously associated with the emergence of broken song syntax, these sensory constraints may contribute to the emergence of broken syntax.

© 2012 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Behavioural performance can be constrained by a variety of factors. Some behaviours may be shaped by the physical properties of the peripheral motor system, or by physiological limitations of the central motor networks that pattern the behaviour (Lauder & Liem 1989). Other behaviours, especially those that are learned, depend heavily on sensory experience and sensorimotor integration. Therefore, the constraints that shape the performance of learned behaviours could also arise from a limited ability of the central nervous system to accurately encode sensory experience relevant to learning. Juvenile songbirds learn to sing by copying the song of an adult tutor in a process that depends heavily on auditory experience and vocal rehearsal (reviewed in Mooney et al. 2008). Despite the songbird's remarkable capacity for vocal learning, even the best singers are constrained in their imitative ability (Ryan & Brenowitz 1985; Nowicki et al. 1992). When tutored with songs containing features exceeding species-typical norms, birds are

often unable to accurately copy those songs (Podos 1996; Podos et al. 1999). A possible explanation for this inaccurate copying is that the properties of the vocal and respiratory musculature, or of the central circuits important to song patterning, limit the bird's ability to reproduce an otherwise accurately encoded auditory representation of the abnormal model (Podos 1996). Alternatively, learning could be constrained by the brain's inability to accurately encode an auditory representation of a song model with abnormal acoustic features (Marler 1976). In this study, we recorded the auditory responses of neurons in freely behaving adult swamp sparrows, *Melospiza georgiana*, to ask whether sensory limitations constrain the representation of abnormal song models that are difficult for young sparrows to copy.

Swamp sparrow songs consist of a multinote syllable trilled continuously at a steady rate. Young swamp sparrows tutored with manipulated songs composed of artificially accelerated trills often fail to copy those songs accurately, producing instead versions of those songs that have been modified in ways that make them apparently easier to produce (Podos 1996; Podos et al. 1999). Sparrows tutored with song models containing highly accelerated trill rates subsequently sing songs in which they accurately

\* Correspondence: S. Nowicki, Department of Biology, Duke University, Box 90338, Durham, NC 27708-0338, U.S.A.

E-mail address: [snowicki@duke.edu](mailto:snowicki@duke.edu) (S. Nowicki).

reproduce both the acoustic features of the song syllable and match the accelerated trill rate, but only in short bursts that are separated by silent gaps, resulting in what is referred to as 'broken song syntax'. Broken song syntax is highly atypical of normal swamp sparrow songs, and it remains unclear what constrains the copying of accelerated trills during the song-learning and development process in such a way as to produce this outcome. Because birds can accurately copy the accelerated model in brief epochs, Podos and colleagues concluded that the central nervous system must contain an accurate auditory representation of the song, and speculated that broken syntax arises because of peripheral motor constraints (Podos 1996). Here we consider the possibility that the sparrow's brain is limited in its ability to accurately encode a sensory representation of songs with artificially accelerated trill rates.

If the production of broken song syntax is influenced by the inability of the central nervous system to accurately encode tutor songs with accelerated trills, then this encoding constraint may be most clearly evident at the point where auditory information about song models is integrated with vocal motor commands (Nottebohm et al. 1976; Hahnloser et al. 2002; Nick & Konishi 2005; Prather et al. 2008, 2010; Roberts et al. 2010). In that light, we recorded the activity of individual neurons in the swamp sparrow's HVC (used here as proper name), a sensorimotor brain structure where individual neurons are active during both singing and listening to song (McCasland & Konishi 1981; Prather et al. 2008), and where auditory-evoked activity has been linked to song learning and perception (Nick & Konishi 2005; Prather et al. 2009; Roberts et al. 2010). Specifically, we sampled the activity of HVC neurons that project to the striatal Area X (i.e. HVC<sub>X</sub> neurons), which in swamp sparrows exhibit auditory-evoked action potential activity that is precisely time-locked to each syllable in a natural song trill (Prather et al. 2008, 2009). If the properties that constrain the production of broken syntax are exclusively motor-related, then HVC<sub>X</sub> auditory responses should accurately encode continuous syllables in an accelerated trill. Alternatively, if the emergence of broken syntax also reflects auditory-processing constraints, then HVC<sub>X</sub> cells may be limited in their ability to follow accelerated trills. We compared the responses of HVC<sub>X</sub> neurons to auditory presentation of natural swamp sparrow songs versus variants of those songs in which the trill rate had been accelerated. Our previous studies have shown that these cells respond in a precisely time-locked fashion to each syllable presented at a normal trill rate (Prather et al. 2008, 2009). In contrast, HVC<sub>X</sub> neurons failed to faithfully follow syllables presented at an accelerated rate, and failure to respond to consecutive syllables increased as a function of the trill rate. Notably, gaps in neural responses began to form at accelerated trill rates associated with the emergence of broken song syntax in previous behavioural studies. These data were recorded in birds that were naïve to accelerated trills and that were not themselves performing broken syntax, eliminating motor performance of broken trills as a possible influence on this failure to respond faithfully to songs with accelerated trills. Together, these data reveal a previously unknown limitation on the central representation of accelerated trills that could limit the bird's ability to learn these songs.

## METHODS

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.

### *Animal Collection and Song Stimulus Preparation*

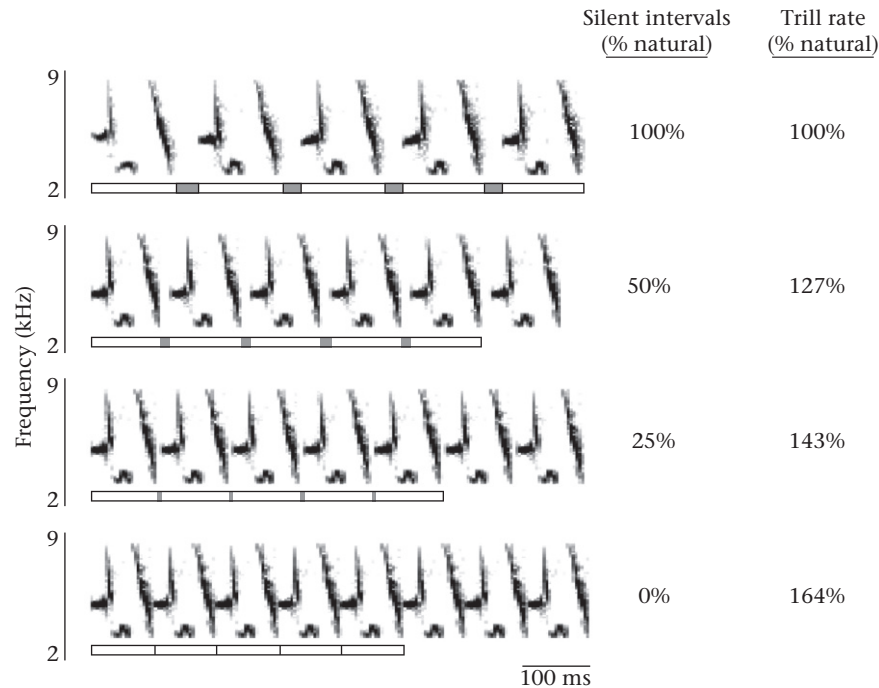
Swamp sparrows were caught with mist nets as adults (age >1 year) on either their winter grounds in Orange County, NC,

U.S.A., or their summer breeding grounds in Crawford County, PA, U.S.A. Birds were housed in cages (46 × 22 × 26 cm) individually throughout their time in the laboratory, both before and during experimentation, and were provided seed and water ad libitum and given a regular supplement of mealworms. Males ( $N = 6$ ) were identified either by external morphology (breeding season) or by molecular marker techniques (out of season), and females were released (methods described in Prather et al. 2008). Prior to implantation of the stimulus and recording devices, birds were subjected to gradually lengthening photoperiod (1 h/week, 9:15 h to 15:9 h light:dark cycle) meant to simulate the onset of the spring breeding season, when swamp sparrows sing most robustly. This change in photoperiod, combined with a subcutaneous implant of testosterone, T (20 mm length of silastic tubing, inner diameter 1.47 mm, outer diameter 1.96 mm, packed with crystalline T; see Nowicki & Ball 1989), was sufficient to induce the birds to sing. Birds were recorded in a semianechoic chamber (Industrial Acoustics Model AC-1, 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 (two to five song types). Although the exact age of birds we used for these experiments was unknown, all songs were performed in the 'crystallized' song syntax that typifies adult swamp sparrow song (Marler & Peters 1982), indicating birds were at least 1 year old. Exemplars of each song type were digitized (25 kHz) and saved onto a computer hard drive (Signal, Engineering Design, Belmont, MA, U.S.A., LabView, National Instruments, Austin, TX, U.S.A.) to be used as stimulus songs.

Song stimuli consisted of natural song types and synthetic variants 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. Following methods used to generate accelerated trills in previous studies, songs were digitally edited to create synthetic variants of the natural song type (Podos et al. 1999). Individual notes were not altered in this process, but the silent intervals between notes and syllables were compressed to 50%, 25% or 0% of their respective durations in the natural song type, resulting in a slight compression of the syllable duration and an acceleration of the overall trill rate (Fig. 1, see Supplementary Fig. S1 for WAV files of songs). Individual synthetic syllables were concatenated to generate trilled stimuli with total song duration like that of natural song types (typically ~2 s).

### *Microdrive Implantation Surgery*

Neurons were sampled using a miniaturized micromanipulation device in awake and freely behaving birds (Fig. 2a) (Fee & Leonardo 2001; Prather et al. 2008, 2009). Several days prior to implantation, birds were transferred from their housing cage to a cage within the recording chamber, a sound-attenuating box (Industrial Acoustics, Bronx, NY, U.S.A., Model AC-1) where they would reside throughout experimentation. During implantation, adult male swamp sparrows were anaesthetized using isoflurane (inhalation, 1–3% in 100% O<sub>2</sub>) 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 and the sites of its projections into the avian striatum (Area X) and the robust nucleus of the arcopallium (RA) (Reiner et al. 2004) (HVC is typically 0.3 mm anterior and 2.3 mm lateral of the bifurcation of the midsagittal sinus, Area X is typically 5.7 mm anterior and 1.9 mm lateral, RA is typically 0.7 mm posterior and 2.3 mm lateral, head angle 30° below zero in all cases). A small craniotomy (approximately 300 × 300 μm) was made in the inner leaflet over Area X, and a small custom-made bipolar stimulus electrode



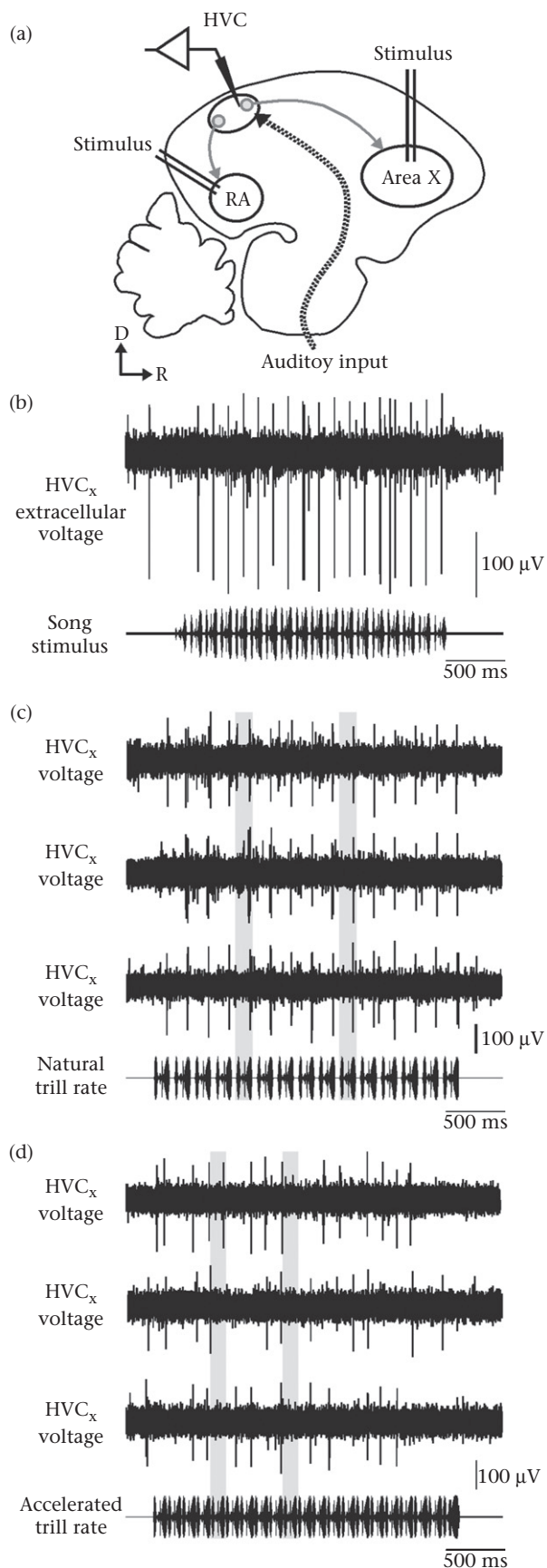
**Figure 1.** Auditory responses of individual neurons were tested using song types in the repertoire of individual adult swamp sparrows and variants of those song types in which the trill rate had been accelerated. An example of a natural song type (top row) contains periods of song syllables (white boxes below each spectrogram, five syllables marked in each row) separated by periods of silence (grey boxes below each spectrogram). In variants of that song type with accelerated trill rate (lower three rows), the periods of sound production were unaltered, but the silences between individual notes and between each syllable were compressed by a specific percentage, resulting in a corresponding acceleration of the trill rate (columns to the right of each spectrogram). The maximal trill rate acceleration for this song type was 164% of the natural trill rate, corresponding to complete compression of all internote and intersyllable intervals.

(J. F. Prather) 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 procedure was repeated to place a bipolar stimulus electrode in the RA. With both stimulus electrodes firmly in place, another small craniotomy was made directly over the HVC. The HVC was located by passing brief (100  $\mu$ s) current pulses through the stimulating electrodes in Area X and the RA and using a sterilized extracellular electrode (Carbostar 1, Kation Scientific, Minneapolis, MN, U.S.A.) 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 the HVC. Across experiments, there was subtle variation of the location of recording electrodes within the HVC. Histological verification of recording sites revealed that our recordings spanned the full anterior–posterior and medial–lateral extent of the HVC, and the depth of the recording electrodes spanned the full depth of the HVC in each bird. Across these sites, we found no evidence of anatomical specialization within the HVC, which would have been evident as a relation between electrode position and auditory response characteristics. The microdrive was secured to the skull using dental cement (microdrive  $\sim$ 1.2 g including dental cement, birds  $\sim$ 16 g), and the incision site was closed using surgical skin adhesive (Vetbond, 3M Animal Care Products, St Paul, MN, U.S.A.). The bird was monitored closely until it was fully recovered, typically within 15 min. After the recording session was complete (1–5 weeks), the bird was deeply anaesthetized with equithesin, perfused transcardially with 0.9% saline followed by 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  $\mu$ m).

### Experimental Protocol

Birds were allowed to recover for 3 days following the implantation procedure before electrophysiological recording began. During recording, microdrive electrodes were slowly advanced into the HVC while weak electrical stimulation was delivered to the stimulus electrodes in either Area X or RA (100  $\mu$ s pulses,  $\sim$ 100  $\mu$ A). The boundaries of the HVC could be reliably identified by observing where antidromic activity was evident. Once an electrode was positioned in the 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, National Instruments). 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 (Wave clus; Quiroga et al. 2004). In both cases, single unit isolation was verified using an interspike interval histogram to test for the presence of a refractory period (Prather et al. 2008; Fujimoto et al. 2011). Our investigation focused on HVC<sub>X</sub> neurons, and individual HVC<sub>X</sub> units were identified using antidromic stimulation via the electrodes placed in Area X or RA (five cells, three birds), or by characteristic features of their action potentials and auditory responses (nine cells, three birds) (Prather et al. 2008, 2009).

When the action potential of an HVC<sub>X</sub> single unit had been isolated and identified, playback of each song type in the bird's repertoire was initiated immediately (10 s quiet interval between each song presentation, stimuli presented in randomized order). Songs were played to the sparrow at 70 dB (peak RMS, 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



**Figure 2.** Extracellular recordings revealed reliable responses of HVC<sub>x</sub> neurons to individual syllables at natural trill rates, but reliability decreased with accelerated trill

bird's entire song repertoire was used to assess the auditory response selectivity of individual HVC<sub>x</sub> neurons, with each cell responding robustly to one song type ('primary song type' for that cell, e.g. Fig. 2b) and responding little or not at all to other songs (Prather et al. 2008). With the primary song type identified for a given neuron, artificially accelerated variants of the primary song type were presented to investigate the responses of HVC<sub>x</sub> neurons to trill rate acceleration like that associated with the expression of broken song syntax (Fig. 2c, d).

Many of the HVC<sub>x</sub> neurons described here (14 cells, six birds) were characterized using not only their responses to accelerated trills, but also their sensorimotor correspondence evident during singing (three cells, one bird) (Prather et al. 2008) and their categorical auditory responses related to song perception (eight cells, four birds) (Prather et al. 2009). In some cases (three cells, one bird), all three parameters were characterized in one and the same cell. This study seeks to gain insight into a behaviour in which the bird must perceive accelerated trills and translate them into a sensorimotor context in order to achieve imitation. Although studying birds using a head-fixed anaesthetized preparation could have facilitated recording from a larger number of cells (Mooney et al. 2001; Prather et al. 2010), it was essential to investigate the response to accelerated trills in the awake state in which HVC<sub>x</sub> activity is also linked to perception (Prather et al. 2009) and in which HVC<sub>x</sub> cells express a precise sensorimotor correspondence (Prather et al. 2008). Although testing these multiple auditory and singing-related properties in the present study introduced limitations in the number of cells, establishing such links in one and the same cell is essential to understanding the degree to which the auditory responses described here could potentially influence perception and imitation.

#### Data Analysis

Action potentials from individual HVC<sub>x</sub> neurons were discriminated and compared against features of the auditory song stimulus (Prather et al. 2008, 2009). Song features were discerned using spectrograms generated in MatLab (Mathworks, Natick, MA, U.S.A.), and all analyses were performed in Matlab using custom software (J. F. Prather). 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; e.g. Figure 1 in Prather et al. 2008). Because swamp sparrow songs are trills consisting of 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; e.g. Figure 1 in Prather et al. 2008). The onset of song or of each syllable was defined as the first time when a song note more than 10 dB louder than background could be detected. In

rate. (a) HVC<sub>x</sub> neurons were identified using antidromic stimulation in Area X, and HVC<sub>x</sub> cells could be easily distinguished from HVC cells that projected to premotor structures (robust nucleus of the arcopallium, RA) or that projected locally with the HVC (interneurons, omitted for clarity). All classes of HVC neurons receive auditory input (dotted arrow; parasagittal schematic of the swamp sparrow brain at the level of the HVC; D = dorsal, R = rostral) (Prather et al. 2009). (b) HVC<sub>x</sub> auditory responses to natural song types were phasic and typically evoked by only one song type in the bird's repertoire ('primary song type' for that cell; top: raw data trace; bottom: oscillogram of song stimulus) (Prather et al. 2008, 2009). (c) HVC<sub>x</sub> cells typically responded to each syllable in primary song types with natural trill rate (7.3 syllables/s, shaded areas highlight individual syllables; top three rows: responses of one HVC<sub>x</sub> neuron to its primary song type; bottom row: oscillogram of the song stimulus) (Prather et al. 2008, 2009). (d) Response of the same cell as in (c) was much less reliable when the primary song type trill rate was accelerated (10.3 syllables/s, data arranged as in (c)).



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 was compared against the background firing rate when no stimulus was present, and the mean background rate plus 5 SD was taken as the threshold for significance (Prather et al. 2008). If the value in any bin in the peristimulus time histogram exceeded that threshold (accounting for bin size), the auditory response was deemed significant. Results obtained in this manner were in good agreement with visual assessment of the efficacy of an auditory stimulus. In quantification of HVC<sub>X</sub> responses to artificially accelerated versions of the primary song type for that cell, measurements were normalized by the response of the same neuron to the natural song type used to create the accelerated stimulus, which facilitated comparison of responses across different cells and birds. Specifically, the degree of trill rate acceleration was expressed as a percentage of the trill rate of the natural song type, responses to accelerated stimuli were expressed as a percentage of the cell's response to the corresponding natural song type, and single-syllable response latencies for accelerated stimuli were computed as the difference between the latency of the peak response to the accelerated stimulus versus the latency of the peak response to the natural song type (expressed in milliseconds or as a percentage of syllable duration). The response strengths evoked by reverse playback of the primary song type in each HVC<sub>X</sub> neuron were averaged, and that mean was taken as a standard above which response strengths were considered valid.

## RESULTS

### *HVC<sub>X</sub> Response Strength Decreased but Response Latency Was Unchanged with Accelerated Trill Rates*

Previously we recorded from HVC<sub>X</sub> sensorimotor neurons and found that they respond regularly and reliably to each syllable in the primary song type when it is presented at a natural trill rate (Prather et al. 2008, 2009). Here we asked whether the activity of those cells would be altered when trill rates were artificially accelerated (in 14 cells (six birds), including three cells (one bird) that expressed a sensorimotor correspondence: Prather et al. 2008; in eight cells (four birds) with activity related to song perception: Prather et al. 2009; and in three cells (one bird) in which sensorimotor correspondence and activity related to perception were both characterized in the same cell). Using stimuli comprising accelerated trills of each neuron's primary song type in which internote and intersyllable silences were shortened, but the notes themselves were unchanged, we observed that the responses of individual HVC<sub>X</sub> neurons diminished with progressively more accelerated trills. In the representative cell shown in Fig. 3a, b (same cell as in Fig. 2b–d), responses to the natural song type were phasic and quite robust, evident as large peaks in the histogram of whole-song responses (Fig. 3a, panel 1). With progressively accelerated trills, however, those histogram peaks became smaller (Fig. 3a, panels 2–4). This decrement in histogram peaks could indicate that the number of action potentials generated in association with each syllable had decreased, or that the response had remained strong but had become temporally dispersed. To distinguish these possibilities, we examined the latency of those auditory responses in relation to each syllable in the song (Fig. 3b). The latency of the peak response differed at most by 5.5 ms across the natural and accelerated conditions (5.1% of natural syllable duration), revealing that the temporal precision of the auditory response was not disrupted as the trill rate was increased. Therefore, this HVC<sub>X</sub> neuron

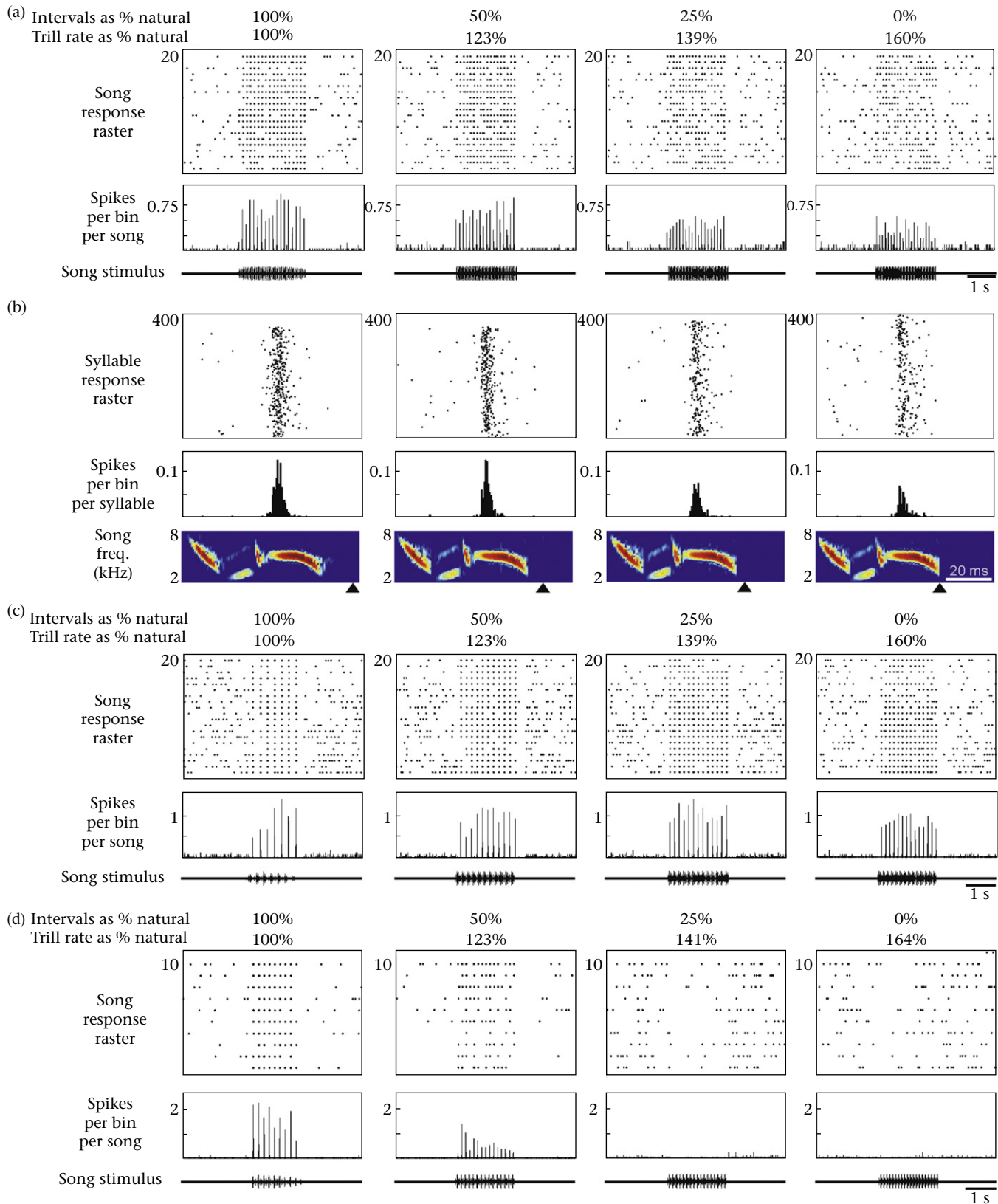
became progressively less responsive to individual syllables of the primary song type when syllables were presented at accelerated trill rates.

The cell described in Fig. 3a, b is representative of the typical response of HVC<sub>X</sub> neurons to accelerated trills. Two other cells illustrate the range of responses detected, where the fidelity of response to individual syllables in accelerated trills was much better (Fig. 3c) or much worse (Fig. 3d). The HVC<sub>X</sub> neuron in Fig. 3c maintained a strong response even as the trill rate was maximally accelerated, indicating that the auditory activity of some HVC<sub>X</sub> neurons could accurately represent both the phonological and the temporal features of accelerated trills. In contrast, the response of the neuron shown in Fig. 3d was strongly decremented even at moderately accelerated trill rates, and no response was evident at maximally accelerated trill rates, indicating that the auditory response of some HVC<sub>X</sub> neurons failed beyond even moderately accelerated trill rates. Notably, the neurons shown in Fig. 3c, d were sampled from the same bird (different primary song types), demonstrating that these disparate patterns of HVC<sub>X</sub> response to accelerated trills coexist within an individual. Furthermore, such differences were also evident even among cells that responded to the same primary song type (e.g. square symbols in Fig. 4). These results reveal that although HVC<sub>X</sub> neurons maintain the temporal precision of their response to specific features of syllables in accelerated trills, individual neurons differ in their ability to respond to each syllable in an accelerated trill.

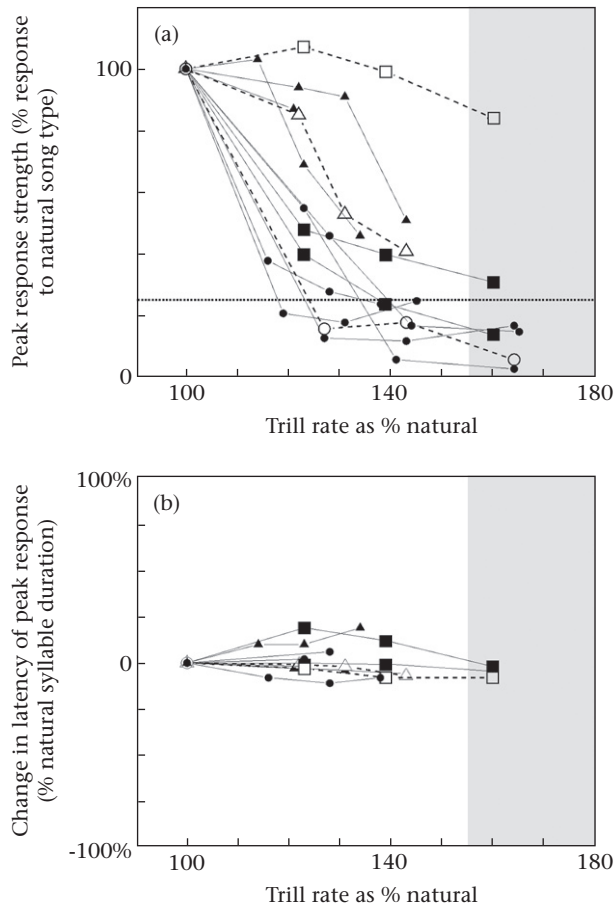
Across the population of all HVC<sub>X</sub> neurons that we tested ( $N = 14$  cells, six birds), responses were strongly decremented with progressively accelerated trill rates, with most cells exhibiting a response roll off between 8 and 10 Hz (Fig. 4a, Supplementary Fig. S1). In the response of each neuron, both the peak response strength and the mean response strength were similarly decreased with progressively accelerated trill rates (Spearman rank correlation:  $r_s = 0.91$ ,  $P < 0.001$ ), supporting our previous observation that action potentials generated in response to accelerated trills were fewer in number but not temporally dispersed. The stability of response latency was confirmed by comparing the responses of each neuron at the natural and accelerated trill rates (Fig. 4b), revealing no systematic change in response latency as a function of trill rate acceleration ( $P = 0.47$ , slope of linear regression compared to zero). Together, these results reveal that HVC<sub>X</sub> neurons responded to each syllable in an accelerated trill with action potentials that remained correlated with the same spectrotemporal features as in the natural song, but action potentials were less likely to be evoked by accelerated trills. Thus, auditory-evoked action potentials in HVC<sub>X</sub> neurons are precisely time-locked to specific acoustic features of the song syllable in both the natural trill and the accelerated trills, but the probability of an action potential occurring for any particular syllable declines with accelerated trill rates.

### *HVC<sub>X</sub> Neurons Failed to Respond to Consecutive Syllables in Accelerated Trills*

The temporal pattern of HVC<sub>X</sub> responses to accelerated trills resembled the temporal pattern of broken song syntax in birds tutored with accelerated trills. Broken syntax songs consist of abnormally brief trills of several syllables in which both the phonology and the accelerated trill rate are accurately reproduced, but those brief trills are interrupted by gaps of silence that are the equivalent of several syllables in duration (Podos 1996; Podos et al. 1999). When presented a song with accelerated trill rate, HVC<sub>X</sub> cells were less likely to respond to individual syllables, although when they did respond, they fired action potentials at the same time in the syllable as when presented with natural songs (e.g. Figs 3, 4). This diminished response to accelerated trills



**Figure 3.** Strength of auditory response in individual HVC<sub>x</sub> neurons typically declined with progressively more accelerated trill rates, but different cells within the same bird could express different types of response ( $N = 14$  cells, six birds). (a) Response of this cell, also shown in Fig. 2b–d, was typical of the HVC<sub>x</sub> population, with decreasing responses at higher trill rates (top: response raster for each song presentation; middle: histogram of aggregate auditory response; bottom: oscillogram of song stimulus). (b) Response of the same cell as in (a) retained its precise relation to specific song features, revealing that reduced peaks in (a) reflected a lower probability of response to each syllable (top: raster of responses to each syllable; middle: histogram of aggregate auditory response; bottom: spectrogram of song syllable). Exceptions to this typical response were also evident. (c) A different cell from a different bird maintained its response strength even at maximally accelerated trills. This was the only cell in which such maintenance of response was observed. (d) A different cell from the same bird as in (c) expressed strong decrement in response to accelerated trills. Differential responses of cells in the same bird (c, d) reveal that the HVC<sub>x</sub> population response comprises a diversity of responses to accelerated trills.



**Figure 4.** Across all HVC<sub>x</sub> neurons tested for responses to accelerated trills ( $N = 14$  cells, six birds), the strength of auditory response became weaker with progressive acceleration of trill rates, but cells retained their relation to specific features within the song syllable. (a) Responses of individual neurons became weaker as trill rates accelerated, with responses becoming less than that evoked by reversed presentation of the primary song type (horizontal dotted line) at even moderately accelerated conditions (triangles indicate antidromically identified HVC<sub>x</sub> cells; open symbols indicate the cells shown in Fig. 3; square symbols illustrate the diversity of responses among three cells that responded to the same primary song type in the same bird). At trill rates where broken syntax emerged behaviourally ( $>156\%$  of natural trill rate, shaded region), few HVC<sub>x</sub> neurons expressed a reliable auditory representation of each syllable in those accelerated trills. (b) Despite changes in response strength, activity of HVC<sub>x</sub> neurons was generally unchanged in its temporally precise association with specific features of the song syllable (symbols as in (a)).

generated epochs in which there were no auditory-evoked action potentials, and those silent periods were typically the equivalent of several syllables in duration (e.g. accelerated cases in Fig. 3a). To investigate this ‘flickering’ phenomenon further, we examined the response of each neuron to serial syllables in the natural and accelerated versions of the primary song type. In response to natural songs, HVC<sub>x</sub> neurons typically generated action potentials in association with many or all syllables in a song (Fig. 5a, panel 1), with rare failures to respond distributed throughout the song duration (Fig. 5b, panel 1). In contrast, responses to accelerated trills were characterized by significantly more failures to respond (Fig. 5a, panels 2–4) throughout the song (Fig. 5b, panels 2–4), significantly shorter sequences of serial responses and significantly longer sequences of serial failures (Figs 5c, 6). Together, these findings reveal a ‘flickering’ response of HVC<sub>x</sub> neurons to accelerated trills reminiscent of the short bursts of accurately imitated syllables separated by gaps of silence in the production of broken syntax.

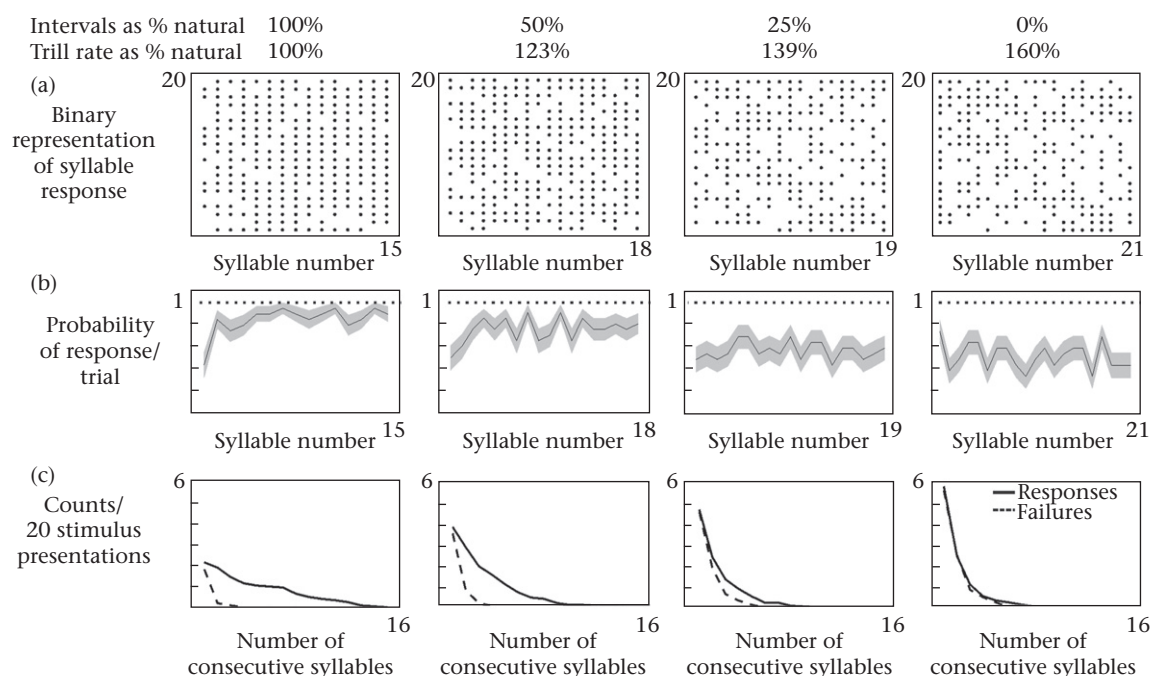
## DISCUSSION

In awake and freely behaving swamp sparrows, HVC<sub>x</sub> neurons express temporally precise responses to specific features of song syllables, and individual neurons typically generate action potentials in response to this feature each time the syllable is produced in the trill (Prather et al. 2008, 2009). In response to songs with artificially accelerated trill rates, HVC<sub>x</sub> neurons continued to maintain the temporally precise relation between auditory-evoked action potentials and specific features of each syllable. Importantly, however, individual cells typically failed to respond to some or many syllables in the accelerated trill, and failure to respond to consecutive syllables became more common with increasingly accelerated trill rates, declining markedly for trill rates above 8–10 Hz. This pattern resulted in a ‘flickering’ auditory representation of accelerated trills, where brief strings of syllables were evident when a response occurred separated by brief gaps in which the cell failed to respond to individual syllables. This pattern of auditory response is reminiscent of the pattern of broken song syntax produced by sparrows tutored on song with artificially accelerated trills, in which short bursts of several accurately imitated syllables are separated by gaps of silence (Podos 1996; Podos et al. 1999).

### A Possible Central Constraint on Song Imitation

Previous work on broken song syntax in swamp sparrows has implicated motor constraints as the primary factor contributing to the emergence of this atypical vocal behaviour (Podos 1996). Podos and colleagues noted that swamp sparrows producing broken syntax songs as a result of being tutored with accelerated trill rate models could accurately imitate the acoustic features of individual syllables even when the models from which they had learned had trill rates far beyond the species norm. Furthermore, those birds could reproduce those syllables at the rapid rate of the model, but only in short bursts punctuated by silent gaps (Podos 1996; Podos et al. 1999). They interpreted the accurate imitation of both syllable structure and trill rate as evidence that the birds’ perception and memory of the accelerated-trill song model was not impaired. Consequently, they suggested that the stereotypic gaps in broken syntax copies of accelerated models reflect a motor constraint encountered during song development, most likely as a consequence of peripheral mechanisms not being able to sustain the demands of producing such rapid songs. In support of the idea that motor constraints are important in that behaviour, broken syntax emerged primarily in the cases of the most accelerated song models, with song models accelerated to a lesser extent often reproduced with accurate phonology and in a continuous trill, but performed with slightly decreased trill rate relative to the model or with the loss of one note in the repeated syllable (Podos 1996; Podos et al. 1999, 2004). These observations led Podos et al. (2004) to speculate that broken syntax emerges as a result of young birds’ difficulty reproducing the accelerated models that they have memorized, calibrating those goals against the motor limits of their song performance and yielding an adult song with broken song syntax.

Consistent with this interpretation, northern mockingbirds, *Mimus polyglottos*, are apparently incapable of fully imitating the long trains of rapid trills performed by canaries and also express broken song syntax in their attempts to imitate those trills, with some evidence that learning is limited by underlying motor constraints (Zollinger & Suthers 2004). Measures of respiration during broken syntax reveal that mockingbirds do not perform the ‘mini-breaths’ that canaries routinely perform during such prolonged and rapid trills (Hartley & Suthers 1989), and instead fully



**Figure 5.** Response failures introduced gaps in the neural representation of songs with accelerated trill rates ( $N = 14$  cells, six birds). Responses of one and the same representative cell are shown in each panel. (a) The natural song type typically evoked a response to each syllable, evident as generally contiguous rows in each raster plot. The reduced probability of response to individual syllables in accelerated trills was evident as gaps in each row (different syllable numbers were presented at each trill rate to preserve the total stimulus duration). (b) Response failures occurred throughout the song duration (black line: mean; grey area: SE of response at each syllable in the song; horizontal dotted line: expectation if the cell responded to every syllable). (c) Response failures tended to occur across consecutive syllables, also evident as gaps in the responses in (a), and the length of those failures became longer as trill rate became more accelerated.

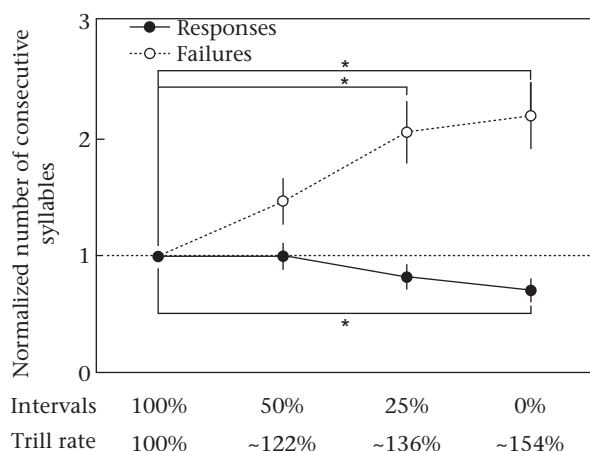
inspire periodically during the trill to respire and maintain the pressure head required for song production (Zollinger & Suthers 2004). These deeper inspirations result in the pauses within the trill that characterize broken song syntax in mockingbirds, raising the possibility that similar respiratory demands underlie the sporadic pauses evident in broken song syntax in swamp sparrows.

Our results suggest that a central neural constraint may equally well contribute to the expression of broken syntax in swamp sparrows, by revealing a limitation of the auditory processing

properties of neurons known to be key in the sensory-motor mechanisms underlying song learning and development. HVC<sub>x</sub> neurons were capable of responding to individual syllables in an accelerated trill, and the relation between neural activity and specific features within each syllable was preserved. Therefore, the phonology of individual syllables remains accurately represented by HVC<sub>x</sub> neurons, even when detected as part of an accelerated trill. However, the neural representation of song syntax became less accurate with progressively accelerated trills, with neural responses containing gaps of several syllables in which no action potentials occurred. Together, these data indicate that even though acoustic features of individual syllables in abnormally rapid trills can be encoded accurately in HVC, the syntax of rapid trills may fail to be represented faithfully at this level of the brain. In other words, at least some cases of broken song syntax could constitute accurate imitation of an inaccurately encoded song model. From weekly recordings of song performances, Podos et al. (2004) gave a broad brush stroke of the variation of trill patterns produced during song development. An investigation of a more detailed timeline of song ontogeny to compare the emergence and sequence of trill development for tutor songs comprising a range of accelerated rates (normal to very rapid) could give insight into the relative contributions of inaccurate auditory coding as well as constraints on production in this form of imperfect learning.

#### How Limitations in Auditory Processing in HVC Could Lead to the Development of Broken Song Syntax

Prior studies suggest several ways in which the auditory activity of HVC neurons, and especially HVC<sub>x</sub> cells, could be linked to song learning. First, synapses in the HVC are a potential site where auditory memories of tutor songs are stored, because auditory experience of a tutor song rapidly stabilizes and strengthens synapses onto



**Figure 6.** Across all HVC<sub>x</sub> cells tested for responses to accelerated trills ( $N = 14$  cells, six birds), the number of consecutive syllables with an associated auditory response became shorter (filled symbols, solid line), and the number of consecutive syllables with no response became longer (open symbols, dashed line) with progressively accelerated trill rate (means and SE are shown; \* $P < 0.05$ , Mann–Whitney  $U$  test; trill rate corresponding to each degree of interval compression varied across cells and is summarized by the mean trill rate across all cells).



HVC neurons, including HVC<sub>x</sub> cells, in juvenile birds lacking prior auditory experience of a tutor song (Roberts et al. 2010). Consistent with the idea that the HVC encodes tutor song memories, tutor songs can evoke auditory responses from HVC neurons in juvenile songbirds (Volman 1993; Nick & Konishi 2005), and robust auditory responses can be evoked in the adult sparrow HVC by tutor songs that have been either copied and rehearsed only transiently during juvenile life or, in some cases, apparently never copied at all (Prather et al. 2010). Second, the auditory activity of HVC<sub>x</sub> neurons in swamp sparrows, including eight of the 14 HVC<sub>x</sub> neurons studied here, is closely linked to song perception. Specifically, these neurons display a response boundary to changes in note duration that closely parallels categorical perceptual boundaries displayed by sparrows for this same acoustic feature (Prather et al. 2009). Furthermore, studies in other songbird species have shown that song perception is impaired by lesions placed in the HVC or in the downstream target of HVC<sub>x</sub> neurons (Area X) (Scharff et al. 1998; Gentner et al. 2000). In light of these two sets of findings, the flickering responses of HVC<sub>x</sub> neurons to songs with accelerated trill rate raise an intriguing alternative explanation for the emergence of broken song syntax. Specifically, flickering representations of accelerated trills may also be present in young sparrows as they experience accelerated-trill song models, and broken syntax may emerge at least partially because swamp sparrows encode memories of those songs as containing strings of individual syllables separated by gaps lasting several syllables in duration. These inaccurate auditory memories could in turn compromise subsequent attempts to imitate the vocal model, even in the absence of any motor constraints. Third, because auditory-responsive HVC<sub>x</sub> neurons also display motor-related activity during singing (Kozhevnikov & Fee 2007; Prather et al. 2008; Fujimoto et al. 2011) and exhibit a precise correspondence in their auditory and singing-related activity (Prather et al. 2008; Fujimoto et al. 2011), the impaired ability of HVC<sub>x</sub> neurons to faithfully follow accelerated trills may point to a central constraint in the sensorimotor activity of these neurons that could further interfere with motor aspects of imitation. Notably, the wild-caught birds that we studied were adults that sang songs with normal syntax, indicating that the flickering representation of accelerated trills is evident in birds previously naïve to accelerated trills and does not depend on the production of broken syntax. Our previous demonstration of a link between juvenile auditory experience and the activity of neurons in the adult HVC (Prather et al. 2010) leads us to posit that this pattern of response may also characterize the activity of HVC<sub>x</sub> neurons in naïve birds exposed to accelerated trills.

To further resolve whether and how limitations in auditory response properties of HVC<sub>x</sub> neurons lead to the emergence of broken song syntax, an important future direction will be to record the activity of HVC<sub>x</sub> neurons in juvenile swamp sparrows engaged in memorizing and then beginning to imitate accelerated-trill song models. Characterizing the activity of HVC<sub>x</sub> cells during the earliest exposure to accelerated-trill song models and tracking that activity throughout the emergence and crystallization of broken syntax will require recordings from very young sparrows collected from the wild as hatchlings and hand-reared in the laboratory. Our experience in attempting to record from very young birds indicates that the skull bones of those birds do not have sufficient density to enable the microdrive recording device to be securely implanted (J. F. Prather & R. Mooney, personal observation). Therefore, future tests of the ideas proposed here will benefit from more advanced recording techniques. In addition, a large population of hand-reared birds will be required, as only a small percentage of birds tutored with accelerated-trill song models display broken syntax (Podos 1996; Podos et al. 1999). Together, these factors will make for a very challenging set of future experiments, yet the insights gained from those outcomes will be essential to test the ideas

posited here and understand the basis of new forms of behavioural syntax.

### Possible Influence on Song Evolution

Our results suggest a mechanism by which constraints inherent in the central nervous system can contribute to the evolution of song syntax. Gradual changes in song structure have been documented in longitudinal studies of individual populations, mostly as differences in the acoustic structure of song elements ('phonology') attributed to errors in song copying, cultural drift, or changes due to selection (Thielke 1972; Slater 1986; Martens 1996; Derryberry 2009). However, differences across clades of birds (such as the genus *Melospiza*) are often characterized by syntactical as well as phonological differences, making it less clear how relatively large-scale changes in song syntax might arise and how these syntactical differences could emerge as the outcome of gradual changes in phonological structure. Broken song syntax represents a disjunct change in syntax and thus may provide a model for further investigation of how saltatory innovation in song syntax arises (Nowicki et al. 2001).

If the broken syntax phenomenon in swamp sparrow song is to serve as a model for saltatory change in the evolution of song syntax, then it is important to show that the trait can be transmitted to subsequent generations and that songs having this trait remain sufficiently functional as signals. The former criterion was demonstrated by Podos et al. (1999), who showed that young swamp sparrows copy model songs presented with broken syntax as readily as they copy songs with normal syntax, and birds that learn from broken syntax models subsequently reproduce them with a syntax matching the model. The functionality of broken syntax songs in male–male territorial interactions was tested in the field by Nowicki et al. (2001), who found them to be equally effective as continuously trilled songs in eliciting male aggressive response in the context of field playback experiments. Broken syntax songs also were effective in eliciting female sexual response as measured in laboratory copulation solicitation assays. Here, the response to broken syntax song is less than that to continuous trills, but significantly more than that to heterospecific songs (Nowicki et al. 2001). Thus, once broken syntax arose in a population, the trait could be transmitted culturally to the next generation, and those birds having songs with broken syntax could use them effectively as dual function signals. In support of a broadly relevant role for such processes, broken song syntax has been reported occasionally in the wild for many species that normally sing continuous trills (including *Spizella passerina*, *Pooecetes gramineus*, *Geospiza fortis* and *Certhidae olivacea*) (described in Podos et al. 1999), revealing that broken syntax is not simply a quirk of swamp sparrows or an artefact of hand rearing birds in captivity or from tape-recorded songs. Thus, the auditory properties described here may constitute a sensory constraint affecting imitation. Together with peripheral motor constraints on song performance, such properties and the broken syntax to which they contribute may provide a broadly relevant mechanism influencing the evolution of song syntax.

### Acknowledgments

We are grateful to Rindy C. Anderson for help with bird collection and to Duke University for support.

### Supplementary Material

Supplementary material for this article is available, in the online version, at [doi:10.1016/j.anbehav.2012.03.012](https://doi.org/10.1016/j.anbehav.2012.03.012).

## References

- Derryberry, E. 2009. Ecology shapes birdsong evolution: variation in habitat and morphology explains variation in white-crowned sparrow song. *American Naturalist*, **174**, 24–33.
- Fee, M. S. & Leonardo, A. 2001. Miniature motorized microdrive and commutator system for chronic neural recording in small animals. *Journal of Neuroscience Methods*, **112**, 83–94.
- Fujimoto, H., Hasegawa, T. & Watanabe, D. 2011. Neural coding of syntactic structure in learned vocalizations in the songbird. *Journal of Neuroscience*, **31**, 10023–10033.
- Gentner, T. Q., Hulse, S. H., Bentley, G. E. & Ball, G. F. 2000. Individual vocal recognition and the effect of partial lesions to HVC on discrimination, learning, and categorization of conspecific song in adult songbirds. *Journal of Neurobiology*, **42**, 117–133.
- Hahnloser, R. H., Kozhevnikov, A. A. & Fee, M. S. 2002. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*, **419**, 65–70.
- Hartley, R. S. & Suthers, R. A. 1989. Air-flow and pressure during canary song: direct evidence for mini-breaths. *Journal of Comparative Physiology A-Sensory Neural and Behavioral Physiology*, **165**, 15–26.
- Kozhevnikov, A. A. & Fee, M. S. 2007. Singing-related activity of identified HVC neurons in the zebra finch. *Journal of Neurophysiology*, **97**, 4271–4283.
- Lauder, G. V. & Liem, K. 1989. The role of historical factors in the evolution of complex organismal functions. In: *Complex Organismal Functions: Integration and Evolution in Vertebrates* (Ed. by D. B. Wake & G. Roth), pp. 63–78. New York: J. Wiley.
- McCasland, J. S. & Konishi, M. 1981. Interaction between auditory and motor activities in an avian song control nucleus. *Proceedings of the National Academy of Sciences, U.S.A.*, **78**, 7815–7819.
- Marler, P. 1976. Sensory templates in species-specific behavior. In: *Simpler Networks and Behavior* (Ed. by J. Fentress), pp. 314–329. Sunderland, Massachusetts: Sinauer.
- Marler, P. & Peters, S. 1982. Structural changes in song ontogeny in the swamp sparrow *Melospiza georgiana*. *Auk*, **99**, 446–458.
- Martens, J. 1996. Vocalizations and speciation of palearctic birds. In: *Ecology and Evolution of Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 221–240. New York: Comstock.
- Mooney, R., Hoese, W. & Nowicki, S. 2001. Auditory representation of the vocal repertoire in a songbird with multiple song types. *Proceedings of the National Academy of Sciences, U.S.A.*, **98**, 12778–12783.
- Mooney, R., Prather, J. F. & Roberts, T. 2008. Neurophysiology of birdsong learning. In: *Learning and Memory: a Comprehensive Reference* (Ed. by H. Eichenbaum), pp. 441–474. Oxford: Elsevier.
- Nick, T. A. & Konishi, M. 2005. Neural song preference during vocal learning in the zebra finch depends on age and state. *Journal of Neurobiology*, **62**, 231–242.
- Nottebohm, F., Stokes, T. M. & Leonard, C. M. 1976. Central control of song in the canary, *Serinus canarius*. *Journal of Comparative Neurology*, **165**, 457–486.
- Nowicki, S. & Ball, G. F. 1989. Testosterone induction of song in photosensitive and photorefractory male sparrows. *Hormones and Behavior*, **23**, 514–525.
- Nowicki, S., Westneat, M. W. & Hoese, W. 1992. Birdsong: motor function and the evolution of communication. *Seminars in Neuroscience*, **4**, 385–390.
- Nowicki, S., Searcy, W. A., Hughes, M. & Podos, J. 2001. The evolution of bird song: male and female response to song innovation in swamp sparrows. *Animal Behaviour*, **62**, 1189–1195.
- Podos, J. 1996. Motor constraints on vocal development in a songbird. *Animal Behaviour*, **51**, 1061–1070.
- Podos, J., Nowicki, S. & Peters, S. 1999. Permissiveness in the learning and development of song syntax in swamp sparrows. *Animal Behaviour*, **58**, 93–103.
- Podos, J., Peters, S. & Nowicki, S. 2004. Calibration of song learning targets during vocal ontogeny in swamp sparrows, *Melospiza georgiana*. *Animal Behaviour*, **68**, 929–940.
- Prather, J. F., Peters, S., Nowicki, S. & Mooney, R. 2008. Precise auditory-vocal mirroring in neurons for learned vocal communication. *Nature*, **451**, 305–310.
- Prather, J. F., Nowicki, S., Anderson, R. C., Peters, S. & Mooney, R. 2009. Neural correlates of categorical perception in learned vocal communication. *Nature Neuroscience*, **12**, 221–228.
- Prather, J. F., Peters, S., Nowicki, S. & Mooney, R. 2010. Persistent representation of juvenile experience in the adult songbird brain. *Journal of Neuroscience*, **30**, 10586–10598.
- Quiroga, R. Q., Nadasdy, Z. & Ben-Shaul, Y. 2004. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation*, **16**, 1661–1687.
- Reiner, A., Perkel, D. J., Mello, C. V. & Jarvis, E. D. 2004. Songbirds and the revised avian brain nomenclature. *Behavioral Neurobiology of Birdsong*, **1016**, 77–108.
- Roberts, T. F., Tschida, K. A., Klein, M. E. & Mooney, R. 2010. Rapid spine stabilization and synaptic enhancement at the onset of behavioural learning. *Nature*, **463**, 948–952.
- Ryan, M. J. & Brenowitz, E. A. 1985. The role of body size, phylogeny, and ambient noise in the evolution of bird song. *American Naturalist*, **126**, 87–100.
- Scharff, C., Nottebohm, F. & Cynx, J. 1998. Conspecific and heterospecific song discrimination in male zebra finches with lesions in the anterior forebrain pathway. *Journal of Neurobiology*, **36**, 81–90.
- Slater, P. J. B. 1986. The cultural transmission of bird song. *Trends in Ecology & Evolution*, **1**, 94–97.
- Thielke, G. 1972. On the origin of divergence of learned signals (songs) in isolated populations. *Ibis*, **115**, 511–516.
- Volman, S. F. 1993. Development of neural selectivity for birdsong during vocal learning. *Journal of Neuroscience*, **13**, 4737–4747.
- Zollinger, S. A. & Suthers, R. A. 2004. Motor mechanisms of a vocal mimic: implications for birdsong production. *Proceedings of the Royal Society B*, **271**, 483–491.