Striatal dopamine modulates song spectral but not temporal features through D1 receptors

Arthur Leblois[1,3] and **David J Perkel**[1,2]

[1]Department of Otolaryngology, University of Washington, Seattle, WA 98195, USA

[2]Department of Biology, University of Washington, Seattle, WA 98195, USA

Abstract

The activity of midbrain dopaminergic neurons and their projection to the basal ganglia (BG) are thought to play a critical role in the acquisition of motor skills through reinforcement learning, as well as in the expression of learned motor behaviors. The precise role of BG dopamine in mediating and modulating motor performance and learning, however, remains unclear. In songbirds, a specialized portion of the BG is responsible for song learning and plasticity. Previously we found that dopamine acts on D1 receptors in Area X to modulate the BG output signal and thereby trigger changes in song variability. Here, we investigate the effect of D1 receptor blockade in the BG on song behavior in the zebra finch. We report that this manipulation abolishes social context-dependent changes in variability not only in harmonic stacks, but also in other types of syllables. However, song timing seems not to be modulated by this BG dopamine signal. Indeed, injections of a D1 antagonist in the BG altered neither song duration, nor the change of song duration with social context. Finally, D1 receptor activation in the BG was not necessary for the modulation of other features of song such as the number of introductory notes or motif repetitions. Together, our results suggest that activation of D1 receptors in the BG is necessary for the modulation of fine acoustic features of song with social context while it is not involved in the regulation of song timing and structure at a larger time scale.

Keywords

zebra finch; songbird; basal ganglia; Area X; learning

Introduction

Cortico–basal ganglia (BG) circuits are thought to promote motor skill acquisition through reinforcement learning (Hikosaka et al., 2002; Graybiel, 2005). During reinforcement learning, individuals first explore their environment. Reinforcers shape this variable behavior until it converges on an optimum for a particular context. Repetition of the successful behavior (exploitation) then replaces exploration (Sutton and Barto, 1990; Ishii et al., 2002), and well-learned skills exhibit low trial-by-trial variability. However, trial-by-trial variability enabling motor exploration may continue to be necessary after learning to maintain the desirable outcome in the presence of external perturbations.

Correspondence should be addressed to: Arthur Leblois, UMR CNRS 8119, 45 rue des Saints Pères, 75006 Paris, France, Tel: +33 142863372 Fax: +33 149279062, arthur.leblois@parisdescartes.fr.
[3]Current address: Laboratoire de Neurophysique et Physiologie (UMR CNRS 8119), Université Paris Descartes, Sorbonne Paris Cité, Paris, France

Song learning and maintenance relies in part on the regulation of song variability for purposes of motor exploration (Scharff and Nottebohm 1991; Fiete et al., 2007; Tumer and Brainard, 2007). Birdsongs have a hierarchical structure spanning timescales from milliseconds to seconds: songs (seconds timescale) consist of repetitions of motifs (~1s each), themselves divided into syllables (10–100ms, Glaze and Troyer, 2007). Across song renditions, there can be variability in the spectral structure of syllables, in the temporal structure of the motifs, or in the organization of motifs into song. In adult birds, the social context modulates both the variability in the spectral and macroscopic song structure (Sossinka and Böhner 1980; Cooper and Goller 2006; Kao and Brainard 2006), and females prefer listening to the less variable directed song (Woolley and Doupe 2008).

Songbirds have a specialized portion of their BG, the anterior forebrain pathway (AFP, Fig. 1A), dedicated to song learning and plasticity (Nordeen and Nordeen 1997; Brainard and Doupe 2002). Changes in neural activity in the BG have been interpreted as representing the neural analog of explore-exploit behavior in mammals (Barnes et al., 2005). In songbirds, the AFP generates song variability that underlies vocal experimentation in juveniles (Ölveczky et al., 2005), and modulates song variability depending on the social context in adults (Kao et al., 2005; Kao and Brainard 2006). While changes in spectral variability clearly depend on AFP output (Kao and Brainard 2006; Ölveczky et al., 2005), the role of this circuit in the modulation of the macroscopic structure and tempo of song by social context remains controversial (Kobayashi et al., 2001; Ölveczky et al., 2005; Kao and Brainard 2006; Hampton et al., 2009; Stepanek et al., 2009; Thompson et al., 2011).

The dopaminergic input to the BG nucleus Area X appears to increase in male birds when they sing to females (Sasaki et al., 2006; Yanagihara and Hessler 2006), a social context associated with decreased BG-induced song variability (Kao et al., 2005). While dopamine (DA) likely triggers changes in the variability of the fundamental frequency of sub-syllabic elements displaying a clear harmonic structure through the activation of D1 receptors in Area X (Leblois et al., 2010), the role of BG DA in modulating spectral variability of other types of sub-syllabic elements as well as the macroscopic song structure remains unclear. Here, we quantified the effects of manipulations to the dopaminergic signal in Area X on the microscopic and macroscopic structure of zebra finch songs.

Material and methods

Animals

Adult male zebra finches (*Taeniopygia guttata*) were obtained from a commercial supplier and used in accordance with an animal protocol approved by the University of Washington Institutional Animal Care and Use Committee. Animals were housed under a 13/11h light/darkcycle with food and water available *ad libitum*. The four birds included in this study, also used in Leblois et al. (2010), were at least 90 days old when starting the experiment.

Surgical procedure

We infused drugs into the CNS using a combination of osmotic minipumps and cannulae, as in Meitzen et al. (2007) and Leblois et al. (2010). Animals were anaesthetized with 2% isoflurane and placed in a stereotaxic apparatus with an angle formed between horizontal and the line from the center of the ear bars to the tip of the beak of 64°. Anesthesia was maintained with 1% isoflurane for the duration of the surgery. Local anesthetic (1% lidocaine) was injected under the scalp, a small craniotomy was made above the midline reference point, the bifurcation of the midsagittal sinus, and a little mark was made at 6.5–7 mm anterior, 1.8 mm lateral from the reference point. The head angle was then changed to 0°, to place the line from the center of the ear bars to the tip of the beak horizontally.

Craniotomies were made around the two marks and two small bent cannulae (Alzet) were then lowered vertically to the surface of Area X (2.5–3 mm deep), at an angle of 90° to the horizontal line from the center of the ear bars to the tip of the beak and attached to the skull with dental cement. An osmotic minipump (Alzet Model 1002, length, 17 mm; diameter, 6 mm; filled weight, 0.5 g; 14 d delivery) filled with 100 µl of drug solution was then connected to the two cannulae with polyvinylchloride tubing and a Y distributor. The pumps were placed in a custom built backpack strapped to the bird's back using a harness made from surgical dressing. To mount theosmotic pump in the backpack, we used a 0.65 ml plastic microcentrifugetube (ISC BioExpress) filled with 250 µl of sterile saline. We threaded the output tube of the pump through a hole in the cap and inserted the pump snugly into the microcentrifuge tube until the lid could snap into place. We sealed the lid and tube extrusion hole using cyanoacrylate adhesive and Parafilm(Fisher Scientific).

Anatomy

At the end of experiments, animals were euthanized by overdose of pentobarbital and perfused transcardially with 0.9% saline followed by 4% paraformaldehyde. The brain was then removed, post-fixed in 4% paraformaldehyde for 24h, and cryoprotected in 30% sucrose. Fifty µm thick sections were then cut in the parasagittal plane using a freezing microtome, mounted on slides and stained with cresyl violet. Histological examination showed that the cannula tip was in Area X or adjacent to the edge of Area X in all four birds, and that any lesion due to cannula placement and drug flow was restricted to a region not exceeding 10% of the total volume of Area X. The typical distance from the cannula tip to the center of Area X was around 500 µm (the diameter of Area X), while the distance to the center of the next closest nucleus involved in song behavior, the lateral magnocellular nucleus of the anterior neostriatum (LMAN), was > 1 mm. The positive results of the manipulations reported by Meitzen et al. (2007) demonstrate that steroid hormones infused locally using the same apparatus as in the present study reached the intended target nucleus (in their case, HVC) at effective concentrations while not diffusing to another further nucleus (in their case, RA). Because steroid hormones are lipophilic and thus readily diffuse through brain tissue, we also expect the drugs used in the present study to act locally and will therefore assume that the effects of the D1 antagonists infused were mainly limited to Area X. A sample image of a section showing the track of the cannula used to infuse drugs into the behaving bird is shown in Leblois et al. (2010).

Song recordings

Birds were individually housed in sound isolation chambers (Acoustic Systems, Austin, TX) 7 days before and 20–40 days following the cannula implantation surgery. We recorded vocalizations using the sound-event triggered Syrinx recording software (John Burt, www.syrinxpc.com, sampling frequency of 22050 Hz). Song was captured in bouts (2–10 s bursts of continuous singing during which the motif may be repeated 1–10 times) and each song bout was saved as a time-stamped .wav file onto the computer hard drive. Spontaneous singing usually occurred throughout the day, with most song bouts produced in the first 3–4 h of light. In addition, we recorded vocalizations evoked by the presentation of a female by placing a female in the cage for 3–4 min, at intervals > 20 min. Such presentation was performed less than 6 times per day (alternating morning and afternoon female presentations on consecutive days) and for at least 3 days in each condition.

Following a week of acclimation to the sound-boxes, song recordings were performed during 5 consecutive preoperative days, including 3 days with female presentations. Postoperative singing was monitored continuously. Each postoperative condition (infusion of the D1-antagonist SCH23390 or saline) included 7 to 10 consecutive days, including 3 to 5 days with female presentations. Song production was reduced following cannula

implantation surgery, regardless of the specific condition considered. However, the song rate following the cannulation surgery apparently did not depend on whether the D1-antagonist SCH23390 or saline was infused into Area X; over the 6 days following surgery, the number of spontaneous song bouts was 184 and 230 in the two birds undergoing infusion of the D1-antagonist SCH23390, and 60 and 352 in the birds undergoing saline infusion.

Data analysis

Song sorting—Songs were sorted and analyzed using custom Matlab (MathWorks, Natick, MA) programs. We designed a program to sort motifs and songs from all sound files continuously recorded using the syrinx software. Briefly, the program detected putative motifs based on peaks in the cross correlation between the recorded sound file and a clean pre-selected motif. Such putative motifs were then sorted based on their spectral similarity with the pre-selected clean motif, using thresholds set by the experimenter. For motifs for which such analysis did not allow unambiguous distinction, an additional principal components analysis on the spectrograms of putative motifs allowed us to sort motifs from other sounds. This analysis allowed us to successfully sort >90% of the motifs sung by a bird on a given day (assessed by comparing hand sorting with the automated sorting by the program). For each extracted motif, we calculated 1) the envelope of the original wave signal (the signal was rectified, smoothed using a gaussian kernel on a 1000 point sliding window, and downsampled by a factor 100 for convenience), and 2) the spectrogram (fast Fourier transforms using 256-point Hanning windows moved in 128-point steps, see Fig. 1B).

Song, syllables and gaps duration—The duration of the motif was determined using the amplitude envelope of the song motifs. Each song consisted of one or several motifs, preceded by introductory notes and separated from each other by <100 ms of silence. The duration of the whole motif was determined as the time between the beginning of the first syllable of the motif and the end of the last one, and therefore included gaps and syllables. Additionally, we separated single syllables and gaps by applying a threshold for sound amplitude (for each bird, a single threshold was set by the experimenter to reliably partition motifs into whole syllables, as in Fig. 1C). This allowed us to compare the duration of syllables and gaps between conditions. Additionally and for comparison of our data with those of Thompson et al. (2011), we calculated the interquartile range (IQR, the range between the first and third quartile) of the motifs, syllables and gaps duration in each condition. The IQR is calculated as the difference between the values of first and last 4-quantiles, and therefore measures the width of the distribution of the 50% motifs/syllable/gap durations centered on their mean. By giving a measure of the spread of the center of a distribution, the IQR provides a robust measure of a variable range in the presence of small numbers of outliers, in contrast to the more commonly used coefficient of variation which is sometimes considerably distorted by a small number of outlying values. Note that the IQR is expressed in the same unit of measure as the variable being considered.

Spectral similarity—Each syllable in a motif was sub-divided into sub-syllabic elements. Each sub-syllabic element was defined as a uniform spectro-temporal unit (harmonic stack, frequency sweep, broad-band sound), and various sub-syllabic elements of a given syllable were usually associated with different sound intensities. The spectrograms of 20–100 randomly-selected, manually-checked renditions of a sub-syllabic element were calculated in each condition. Again, spectrograms were calculated by performing power spectra over 256-points (11.6 ms) Hanning windows moved in 128-point steps, with a frequency resolution of 43 Hz. Then, cross-correlations were run between all possible pairs of this subset. For each pair of syllable renditions, a cross-correlation index was calculated as the sum of the cross-correlation function between their two spectrograms, normalized by the

square root of the product of their autocorrelation function. The average cross-correlation index over all 10,000 pairs was called the "spectral similarity index" of the sub-syllabic element in that condition.

The spectral similarity index is based on cross-correlation measurements and may be influence by temporal variability in the signal. Indeed, the potential exists that differences in syllable length affect the cross-correlation values and thereby skew the spectral similarity index measured over many pairs in a given condition. Since we show that DA D1 receptor blockade does not alter temporal variability in song motifs and syllables (see Results), a systematic effect of temporal variability on spectral similarity between different pharmacological conditions seems unlikely. However, to rule out any confound, we replicated our findings concerning changes in spectral similarity between the various pharmacological conditions for the example sub-syllabic element shown in Fig. 4A by computing pair-wise cross-correlations after adjusting samples length through time warping (Anderson et al., 1996).

Additional features—A song bout was defined as a sequence of introductory notes and song motifs separated by less than 2s of silence (Sossinka and Böhner 1980). The number of motifs per song-bout was assessed over a subset of 100 randomly selected song files recorded in each condition. The number of introductory notes was also determined over the same sample of song files.

Statistics—The planned comparisons included, in control conditions, the effect of social context on the number of introductory notes and motifs per bout, motif, syllable and gap duration, and note similarity. In addition, we planned to compare the effect of social context with the presence and absence of the D1 antagonist (SCH 23390 v. saline groups). Numerical values are given as mean ± SD, unless stated otherwise. For comparison with the literature, values of these parameters in control conditions were compared between different social contexts using paired t-tests (Prism, GraphPad Software). When this assessment was made on a single note/motif from a single bird, for example purposes, we used an unpaired t-test across renditions. For these within-bird tests, the sample size was the number of renditions analyzed, leading to very many degrees of freedom in the test (>1000 renditions). To compare the combined effects of female presentation and drug condition on the various song parameters (spectral similarity, duration, number of motifs and introductory notes), we performed a two-way, repeated measures ANOVA (Prism, GraphPad Software). If the interaction between social context and drug condition was significant, we performed *post-hoc* t-tests comparing the effect of social context in the presence or absence of the D1 antagonist. These t-tests were Bonferroni corrected for the number of comparisons made. For each t-test applied, we report the associated p-value (the probability of observing the given result, or one more extreme, by chance if the null hypothesis is true), the value of the test statistic (t), and the degrees of freedom of the test (df). A value of $p < 0.05$ was considered as a significant difference.

Results

Activation of D1 receptors decreases spectral variability

In a previous study, we showed that the activation of D1 receptors in Area X was responsible for the social context dependent modulation of the variability of the fundamental frequency of specific sub-syllabic elements called harmonic stacks (Leblois et al., 2010). To test whether the change in acoustic variability with social context through D1 receptor activation could be generalized to sub-syllabic elements that do not display the clear spectral structure of harmonic stacks, we made pairwise comparisons of renditions of a subset of

each type of sub-syllabic element using the spectrogram cross-correlation method (see Methods, Nelson and Marler, 1994). The average cross-correlation coefficient among pairs of spectrograms of the renditions of this element was called the "spectral similarity index". It allowed us to compare the acoustic variability in a set of renditions of each sub-syllabic element across different conditions.

Fig. 2A displays the results of such analysis applied to the first element of syllable 6 in the song of bird #4 (whose motif and syllable partition is depicted in Fig. 1). In the baseline condition, the average spectral similarity between renditions of this note was higher when the bird sang in the presence of a female (solid black line, average cross-correlation of 0.65 ± 0.09) than when he sang alone (dashed line, average cross-correlation of 0.55 ± 0.13, $p<0.001$, $t=-11$, $df=1752$, note here that the number of degrees of freedom reflects the number of notes produced in each condition). To assess within this individual animal whether this example note exhibited different variability in the different social contexts and drug conditions, we compared spectral similarity values using a two-way ANOVA. This test revealed a significant interaction between the presence of a female and infusion of the D1 antagonist SCH23390 ($F=4.24$, $df=1$, $p<0.05$). Post-hoc analysis revealed that the syllable spectral similarity was increased in the presence of a female when saline was infused (0.55 ± 0.1 alone versus 0.67 ± 0.09 in the presence of a female, $p<0.001$, $t=4.9$, $df=1614$, Fig. 2A), but not during infusion of the D1 antagonist SCH 23390 into Area X (0.50 ± 0.06 alone and 0.55 ± 0.1 in the presence of a female, $p=0.1$, $t=1.6$, $df=1619$). Given that differences in syllable length may affect the cross-correlation values and thereby skew the spectral similarity index measured over many pairs, we replicated the analysis of spectral similarity for this sub-syllabic element by computing pair-wise cross-correlations after adjusting syllable length through time warping (Anderson et al., 1996). Although the spectral similarity was increased after time warping in all pharmacological conditions, the differences in spectral similarity with social context was still present at baseline (0.67 alone versus 0.74 in the presence of a female, $p<0.001$, $t=-9.7$, $df=1752$), abolished by SCH 23390 infusion in Area X (0.64 alone versus 0.66 in the presence of a female, $p=0.2$, $t=-1.7$, $df=1642$), and recovered following saline infusion (0.68 alone versus 0.76 in the presence of a female, $p<0.001$, $t=5.0$, $df=1714$).

In this bird, we first compared, under baseline conditions, spectral similarity of all sub-syllabic elements sung in the presence or absence of a female. There was a significant increase in the spectral similarity of sub-syllabic elements (0.67 ± 0.07 versus 0.59 ± 0.08, $p=0.001$, $t=5.2$, $df=8$). We next compared the effect of the D1 antagonist SCH23390 on modulation of similarity by social context using a two-way, repeated measures ANOVA. There was a significant interaction between drug condition and social context ($p<0.01$, $F=12.7$, $df=1$). Post-doc analysis showed that the female had no effect on syllable similarity when D1 receptors were blocked (saline: 0.68 ± 0.1 versus 0.45 ± 0.13, $p<0.001$, $df=8$, $t=5.6$; SCH23390: 0.64 ± 0.06 versus 0.66 ± 0.1, $p=0.3$, $df=8$, $t=-1.0$, Fig. 2B). These values indicate that, for this animal, blockade of D1 receptors in Area X suppressed social modulation of syllable spectral variability.

Across all birds recorded under baseline conditions, the mean spectral similarity was significantly higher when the birds sang in the presence of a female (0.684 ± 0.03) than when they sang alone (0.638 ± 0.04, $p=0.01$, $t=-4.2$, $df=3$, Fig. 3C). Because the variance of spectral similarity among sub-syllabic elements of a single bird was comparable to that of sub-syllabic elements from different birds (Bartlett's test for equal variances, $p=0.6$ and $p=0.4$ for spectral similarities in the presence and absence of female, respectively), we pooled song elements from all birds for the comparison between different pharmacological conditions.

In a first analysis, we considered only sub-syllabic elements that were not considered in the previous publication (Leblois et al., 2010), i.e. all elements except harmonic stacks. There were 26 "non-stack" sub-syllabic elements (7 in bird #1, 4 in bird #2, 10 in bird #3 and 5 in bird #4, see Fig. 3D), which, on average, displayed a higher spectral similarity when sung in the presence of a female (0.68 ± 0.07) than alone (0.64 ± 0.08, p=0.001, t=-3.9, df=25) in baseline conditions. For the analysis of the spectral similarity after surgery under drug and vehicle infusion, we excluded the last element in the song of bird #3 which displayed very low spectral similarity in the saline condition, well below all other elements (see Fig. 3D), because it was frequently dropped and therefore displayed a low number of iterations. A two-way, repeated measures ANOVA on all 25 other "non-stack" sub-syllabic elements revealed a significant interaction between drug condition and social context (F=5.1, df=1, p<0.05). Further *post-hoc* tests revealed significant increase in spectral similarity in the presence of a female in saline condition, but not when D1 receptors were blocked in Area X (saline: 0.69 ± 0.09 versus 0.61 ± 0.12, p<0.01, t=3.0, df=24; SCH23390: 0.67 ± 0.07 versus 0.66 ± 0.08, p=0.4, t=0.94, df=24).

When considering all sub-syllabic elements, including the harmonic stacks previously analysed, results were very similar. While the average spectral similarity was higher in the presence of a female in baseline conditions (0.69 ± 0.07 versus 0.65 ± 0.07, p<0.001, t=-5.5, df=38), a two-way, repeated measures ANOVA revealed a significant interaction between drug condition and social context (F=7.3, df=1, p<0.01). Further post-hoc analysis revealed that the difference in spectral similarity with social context disappeared during infusion of the D1-antagonist SCH 23390 (saline: 0.69 ± 0.10 versus 0.61 ± 0.12, p<0.01, t=3.3, df=38; SCH23390: 0.67 ± 0.08 versus 0.66 ± 0.07, p=0.1, t=1.6, df=38).

Finally, we investigated the effect of the D1 antagonist on spectral similarity during singing in the presence of a female. The mean spectral similarity of syllable renditions sung in the presence of a female was significantly lower during infusion of the D1-antagonist SCH 23390 into Area X than in the baseline condition (0.69 ± 0.07 versus 0.67 ± 0.08, p=0.02, t=2.4, df=38, Fig. 3A). Due to two outliers displaying low spectral similarity in the presence of a female during saline infusion in Area X (Fig. 3B), the difference in spectral similarity between saline and SCH 23390 infusion did not reach significance. However, there was a significant difference between the two conditions when the two outliers were excluded (0.70 ± 0.07 versus 0.68 ± 0.07, p=0.02, t=2.1, df=36).

Altogether, our data show that the modulation of spectral similarity of all syllables with social context is abolished by the infusion of the D1 receptor antagonist SCH 23390 in Area X, even when focusing on non-harmonic sub-syllabic elements.

Transmission through D1 receptors is not necessary for social context-induced change in motif duration

The duration of song motifs varied widely across the birds used in this experiment (mean duration ranged from 0.5 to 1 s). However, each bird displayed a very narrow distribution of motif duration, with only about 2 % variation in a given social context (the standard deviation of motif duration ranged from 8 to 20 ms). In three birds, there was a significant decrease in the motif duration when the bird sang in the presence of a female relative to when he was singing alone before surgery (p<0.001 in all three birds). As an example, in bird #2, the motif duration was reduced in the presence of a female by 2% from 0.567 ± 0.01 s to 0.555 ± 0.01 s (p<0.001, t=13, df=5554, Fig. 4A). In the last bird, however, song motifs were slightly longer in the presence of a female (0.532 ± 0.01 s vs. 0.535 ± 0.01 s, 0.5% difference, p=0.04, t=-1.9, df=4270).

Because a previous study reported that most, if not all, social context modulation of song duration is due to a shortening of expiratory sounds, while inspirations remain constant in duration (Cooper and Goller, 2004), we compared the duration of syllables (usually produced during expiration) and gaps (thought to correspond to inspiration) in the two social contexts. In contrast, we observed that the duration of gaps was much reduced in the presence of a female in all three birds showing social context modulation of song duration (−17%, −7%, −9%), while syllable duration was only slightly and inconsistently modified by social context in these birds (1%, −1%, −2%). The discrepancy of our results with the previous study of Cooper and Goller (2004) may be due to the fact that gaps and syllables do not simply reflect inspiration and expiration respectively (Aronov et al., 2011).

Following infusion of the D1-antagonist SCH23390 into Area X, the duration of song motifs was still decreased in the presence of a female, and a two-way, repeated measures ANOVA revealed no significant interaction between drug condition and social context on motif duration (F=0.3, df=1, p=0.9). Although this test did not give a significant result, it had low power, and does not provide strong support for the null hypothesis, i.e. that area X D1 receptors play no role in social modulation of motif duration. However, in a given social context, the infusion of the D1-antagonist SCH23390 into Area X did not consistently change the motif duration, and motif duration changed >1% in only one social context in one of the four birds. The 3 birds showing a reduction in motif duration in the presence of a female before surgery still displayed a strong and significant modulation in the motif duration with social context (p<10^{-4} in all 3 birds) during infusion of D1 antagonist in Area X. In particular, in bird #2, song duration was still decreased in the presence of a female (0.57 ± 0.01 versus 0.56 ± 0.02, p<0.001, t=6.1, df=5741, Fig. 4A). Altogether, our data do not support a role for Area X D1 receptors in social modulation of motif duration.

Because a previous study showed that the difference in song duration between social contexts is highest for the first motif in a song bout (Cooper and Goller, 2006), we may have underestimated the effect of DA on song duration by looking at all motifs produced. We therefore repeated our motif-duration analysis focusing only on the first motif of each song bout in each condition, and found the same results. Indeed, first motif duration was still decreased by 2 ± 2 % in the presence of a female over all birds in baseline condition, with three out of four birds displaying a decrease in song duration with the female (Bird #1: p<0.001, t=13, df=397; Bird #2: p<0.001, t=8.9, df=3112; Bird #3: no significant change; Bird #4: p<0.001, t=5.1, df=2667). A two-way, repeated measures ANOVA revealed no significant interaction between drug condition and social context on the duration of the first motifs (F=0.1, df=1, p=0.98), and the difference with social context persisted during the infusion of the D1-antagonist SCH23390 in Area X (Bird #1: p<0.001, t=16, df=1255; Bird #2: p=0.002, t=3.1, df=3453; Bird #3: no significant change; Bird #4: p<0.001, t=3.5, df=113). Overall, when we limited our analysis to the first motif of each song bout, we still observed no effect of area X blockade of D1 receptors on social modulation of motif duration (Fig. 4C).

Finally, since a previous study revealed an influence of the AFP on the dispersion of syllables and gaps duration (Thompson et al., 2011), we investigated whether social context or D1 receptor manipulation in Area X modified the dispersion of sounds and silences duration in songs. Dispersion was measured using the inter-quartile range (IQR, see Methods) for each parameter, as in that previous study. We found that social context did not alter the dispersion of motif duration in baseline condition (IQR of 17 ± 9 ms with a female versus 18 ± 6 ms alone, n=4, p=0.9, t=−0.16, df=3), syllable duration (IQR of 10 ± 11 ms with a female versus 8 ± 2 ms alone, n=20, p=0.4, t=−0.89, df=19), or gap duration (IQR of 6 ± 6 ms with a female versus 6 ± 6 ms alone, n=16, p=0.6, t=−0.51, df=15) as measured using IQR. Moreover, a two-way, repeated measures ANOVA revealed no interaction

between drug condition and social context for the motif duration IQR (F=0.35, df=1, p=0.6), syllable duration IQR (F=1.9, df=1, p=0.2), or gap duration IQR (F=2.3, df=1, p=0.1). These statistical tests revealed no significant effect of drug or social context alone (results not show). Because the power of the tests was low, our results do not make a strong case in favor of the null hypothesis, that area X D1 receptors paly a role in regulating syllable and gap durations.

In conclusion, the duration of the song motif and its modulation with social context seemed unaffected by the selective blockade of D1 receptors in the BG nucleus Area X. We therefore conclude that D1 receptor activation in Area X may not be necessary for the modulation of song duration with social context.

Signaling through D1 receptors is not necessary for social context-modulation of other global song features

The number of motifs per bout was significantly increased when the bird sang in the presence of a female (Bird #1: p<0.001, t=4.1, df=399; Bird #2: p<0.001, t=22, df=3113; Bird #3: p<0.001, t=7, df=3769; Bird #4: p=0.02, t=2.3, df=3733). For example, bird #2 displayed an increase in the number of motifs per bout from 2.1 ± 0.7 motifs when singing alone to 5.9 ± 2.9 motifs in the presence of a female (Fig. 5A). Over all birds, the average number of motifs per bout increased in the presence of a female (from 2.2 ± 0.1 motifs per bout alone to 3.8 ± 1.7 motifs per bout in the presence of a female, Fig. 5C). A two-way, repeated measures ANOVA revealed no significant interaction between drug condition and social context on the number of motifs per bout when comparing infusion of the D1-antagonist SCH23390 and saline in Area X (F=1.1, df=1, p=0.3). Moreover, single t-tests highlighted no change in the number of motifs per song bout during the infusion of the D1 antagonist SCH23390 in Area X in each bird and social context (Bird #1: with female, p=0.1, t=0.4, df=124; Bird #2: with female, p=0.1, t=−2, df=28; without female, p=0.4, t= −0.84, df=4035; Bird #3: with female: p=0.1, t=1.8, df=79; Bird #4: with female: p=0.8, t= −0.24, df=18; without female: p=0.5, t=0.64, df=315), and only two birds had a slight but significant decrease in the number of motifs per bout in the drug condition as compared to saline when singing alone (Bird #1: 2 versus 2.2, p<0.001, t=8, df=3061; Bird #3: 2 versus 2.4, p<0.001, t=5.2, df=1495). Bird #2 still displayed an increase in the number of motifs per bout from 2.1 ± 0.8 to 4.5 ± 2.9 in the presence of a female (Fig. 5B, p<0.001, t=−32, df=6880). Over all birds, the number of motifs per bout was still greater in the presence of a female after SCH23390 infusion (from 2.0 ± 0.1 motifs per bout alone to 3.2 ± 1.3 motifs per bout in the presence of a female, Fig. 5C). Altogether, our results do not support a role for D1 receptors in Area X in modulating the number of motifs produced in a song bout.

The scenario was very similar concerning the number of introductory notes per motif (Fig. 5D–F). Indeed, at baseline, the number of introductory notes was significantly increased when the bird sang in the presence of a female in 3 of the 4 birds (p<0.01 in all 3 birds). One bird did not display a significant modulation of the number of introductory notes with social context. As an example, the number of introductory notes sung by bird #2 increased from 3.1 ± 0.6 to 6.4 ± 1.2 in the presence of a female (Fig. 5D, p<0.001, t=10, df=30). Over all birds, the average number of introductory notes was strongly increased in the presence of a female (from 2.4 ± 0.8 alone to 4.2 ± 2.2, Fig. 5F). A two-way, repeated measures ANOVA revealed no significant interaction between drug condition and social context on the number of motifs per bout when comparing infusion of the D1-antagonist SCH23390 and saline in Area X (F=2.3, df=1, p=0.2). We compared the number of introductory notes in each social context between saline and drug condition for each bird. None of the single t-tests revealed a significant change in number of introductory note with drug treatment (Bird #1: without female p=0.7, t=−0.6, df=52, with female p=0.3, t=1.2, df=44; Bird #2: without female p=0.4, t=0.8, df=37, with female p=0.6, t=−0.5, df=26; Bird #3: without female p=0.8, t=0.3,

df=35, with female p=1, t=0.1, df=35; Bird #4: without female p=0.1, t=1.6, df=28, with female p=0.4, t=−1, df=12). During infusion of the D1-antagonist SCH23390 into Area X, bird #2 still increased the number of introductory notes from 3 ± 1.1 to 6.7 ± 1.3 in the presence of a female (Fig. 5E, p=0.001, t=6.5, df=17), and the number of introductory notes was still increased from 2.4 ± 0.7 when singing alone to 4.4 ± 2.3 when singing to a female over all birds (Fig. 5F). Because each of the applied tests had high power, and all were far from reaching statistical significance, our data do not support a role for Area X D1 receptors in social modulation of the number of introductory notes preceding a song bout.

In conclusion, the modulation of global song features like the number of song motifs per bout or the number of introductory notes preceding the first motif appears not to be specifically influenced by signaling through D1-receptors in Area X.

Discussion

In the present study, we report that interfering with D1-type DA receptor signaling in the songbird basal ganglia nucleus Area X abolishes social context-related changes in song acoustic variability, not only in harmonic stacks, as reported previously for the same animals (Leblois et al., 2010), but also in other types of song elements. However, our results do not support a role of D1 receptor signaling in Area X in the modulation of song tempo and structure at larger time scales. Indeed, song duration, its change with social context, or other global song features such as the number of introductory notes and motif renditions in a song bout seem unaffected by D1 receptors blockade in Area X. Measure of acoustic similarity

We have shown in a previous study that the modulation of variability in the fundamental frequency of harmonic stacks with social context may be triggered by the activation of D1 receptors in Area X (Leblois et al., 2010). The measure of fundamental frequency of harmonic stacks is widely used as a measure of song variability (Kao et al., 2005; Kao and Brainard 2006; Sakata et al., 2008; Andalman and Fee 2009; Hampton et al., 2009; Stepanek and Doupe 2010). This measure however restricts the analysis of song variability to only a small subset of the song elements produced, especially in the case of zebra finches. In zebra finch songs, harmonic stacks are typically present in only 10–50% of the syllables (unpublished data). Another simple measure of spectral similarity between two renditions of a given song element is the cross-correlation index between their spectrograms (Nelson and Marler 1994). This variability measure may in principle be applied to any arbitrary part of a bird song: entire motifs, syllables or sub-syllabic elements (sometimes called notes). Because the song duration is variable and even the smallest song elements can be stretched in time (Glaze and Troyer 2007), the cross-correlation measure is expected to provide a more accurate picture of spectral similarity of song when applied to the finest time-scale song elements: the 10–100 ms sub-syllabic elements. For these reasons, we compared renditions of such sub-syllabic elements in the present study. In particular, we extended previous findings restricted to harmonic stacks about modulation of spectral variability of song with the social context (Kao et al., 2005) to all song elements.

Role of DA in modulating spectral variability

Lesion or inactivation of the AFP output nucleus LMAN substantially reduces song spectral variability, suggesting that the AFP regulates both developmental and contextual modulation of song spectral features (Kao et al., 2005; Ölveczky et al., 2005). Neurons in the AFP and in the ventral tegmental area of adult zebra finches show singing-related activity that is modulated by social context (Hessler and Doupe 1999; Yanagihara and Hessler 2006; Kao et al., 2008). Moreover, extracellular DA level in Area X is higher when adult zebra finches sing to a female than when they do not sing or sing alone (Sasaki et al., 2006). Together, these results suggest that the dopaminergic input to Area X may be modified depending on

the social context and could trigger changes in song spectral variability through its action on the AFP. In a recent study, we have shown that DA may mediate these changes in the variability of the fundamental frequency of harmonic stacks by acting on D1 receptors in Area X (Leblois et al., 2010). Here, we extend those previous results, showing that the modulation of spectral similarity between renditions of all sub-syllabic elements by social context also relies on the activation of Area X D1 receptors.

Role of DA in modulating song duration

We found that song duration is strongly modulated with social context, with an increased song tempo when the bird sings to a female, consistent with previous studies (Sossinka and Böhner 1980; Cooper and Goller 2006). Previous studies have shown the role of the AFP in the modulating song tempo (Williams and Mehta 1999; Brainard and Doupe 2002; Kao et al., 2006). We therefore would have expected that DA plays a role in decreasing song duration in the presence of a female. However, we found no evidence of any modification in the song tempo in either social context after manipulation of the dopaminergic signal in Area X. More importantly, the difference in song duration with social context persisted during this manipulation, suggesting that the modulation of song tempo with social context does not depend on the activation of D1 receptors in Area X. This was true even when we restricted our analysis to the first motif of each song bout, for which social context-related duration variations are more pronounced (Cooper and Goller 2006). Therefore, if the AFP plays a role in the modulation of song tempo with social context as proposed by Kao et al. (2006), it is probably not the activation of D1 receptors in Area X that triggers the underlying change in the related AFP signal. Consistent with the idea that two different mechanisms are involved, changes in song tempo following LMAN lesions occur on long time scales from days to weeks (Brainard and Doupe 2002), while the changes in spectral variability triggered by the dopaminergic input in the AFP occur over seconds or minutes. Moreover, LMAN inactivation induces an immediate decrease in song spectral variability, without altering song tempo (Stepanek and Doupe 2009).

Modulation of global song structure

In contrast to song spectral features, the global structure of a song bout, such as patterning of introductory elements and motifs, and the specific order of syllables in motifs, is thought to be controlled by the premotor nucleus HVC and its inputs (Vu et al. 1994; Long and Fee 2008), and less subjected to AFP modulation. More specifically, previous work suggested that the AFP is not required for the modulation of introductory elements and motif patterning (Kao et al., 2006) or syllable sequencing (Hampton et al., 2009) with social context. However, two different studies reported an influence of the AFP on the syllable sequence order, either in juvenile zebra finches (Ölveczky et al, 2005), or in adult bengalese finches, which display a more complex song sequence than zebra finches (Kobayashi et al., 2001). Our results, consistent with the reports of Kao et al. (2006), do not support a role of D1 receptor activation in Area X in adult zebra finches for the modulation of the number of introductory notes and motifs per song bout with social context.

Role of DA input to the BG in sensorimotor learning

Trial-and error motor learning, and in particular vocal learning, relies on the ability of the subject to differentially reinforce patterns of motor activity that produce better outcomes and/or to punish those that result in worse outcomes. Previous studies suggest that the higher variability of song of adult birds in the absence of a female reflects motor exploration underlying such trial-to-trial learning (Tumer and Brainard 2007; Andalman and Fee 2009). Moreover, Kojima and Doupe (2011) have recently shown that very similar social-context dependent modulation of song variability is present to a greater extent in juveniles, where such motor exploration is necessary for learning (Ölveczky et al., 2005). A neural correlate

of this behavioral variability modulation is found in the AFP; firing is reduced and less variable during directed singing in Area X and LMAN (Hessler and Doupe 1999; Kao et al., 2008). Immediate early gene expression in Area X is also differentially modulated (Jarvis et al., 1998). Moreover, lesion of the AFP output nucleus LMAN abolishes the social-context dependent variability in adults (Kao et al., 2005; Kao and Brainard 2006) and the exploratory variability in juveniles (Ölveczky et al., 2005). Consistent with our previous report (Leblois et al., 2010) but for another set of syllables from the same animals, we confirmed that spectral fluctuation in the song is modulated by the dopaminergic input to Area X. This modulation may allow DA to regulate exploration and exploitation in song behavior, consistent with previous reports of DA-regulating genes responsible for inter-individual differences in exploration and exploitation behaviors (Frank et al., 2009). On the contrary, social context-dependent changes in the duration and global structure of song are independent of DA. Our results further suggest that if DA regulates spectral fluctuations with social context to gate vocal exploration, other social context-related changes such as modifications in tempo and song structure are likely driven by different mechanisms and may not be related to song plasticity.

Acknowledgments

Acronyms

AFP	Anterior Forebrain Pathway
BG	Basal Ganglia
DA	Dopamine
IQR	Interquartile Range
LMAN	Lateral Magnocellular nucleus of the Anterior Neostriatum
RA	Robust nucleus of the Archistriatum

References

Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc Natl Acad Sci USA. 2009; 106:12518–23. [PubMed: 19597157]

Anderson SE, Dave AS, Margoliash D. Template-based automatic recognition of birdsong syllables from continuous recordings. J Acoust Soc Am. 1996; 100:1209–19. [PubMed: 8759970]

Aronov D, Veit L, Goldberg JH, Fee MS. Two distinct modes of forebrain circuit dynamics underlie temporal patterning in the vocalizations of young songbirds. J Neurosci. 2011; 31:16353–68. [PubMed: 22072687]

Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature. 2005; 437:1158–61. [PubMed: 16237445]

Brainard MS, Doupe AJ. What songbirds teach us about learning. Nature. 2002; 417:351–8. [PubMed: 12015616]

Cooper BG, Goller F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. J Neurophysiol. 2006; 95:3798–809. [PubMed: 16554509]

Fiete IR, Fee MS, Seung HS. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. J Neurophysiol. 2007; 98:2038–57. [PubMed: 17652414]

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat Neurosci. 2009; 12:1062–8. [PubMed: 19620978]

Glaze CM, Troyer TW. Behavioral measurements of a temporally precise motor code for birdsong. J Neurosci. 2007; 27:7631–9. [PubMed: 17634357]

Graybiel AM. The basal ganglia: learning new tricks and loving it. Curr Opin Neurobiol. 2005; 15:638–44. [PubMed: 16271465]

Hampton CM, Sakata JT, Brainard MS. An avian basal ganglia-forebrain circuit contributes differentially to syllable versus sequence variability of adult Bengalese finch song. J Neurophysiol. 2009; 101:3235–45. [PubMed: 19357331]

Hessler NA, Doupe AJ. Social context modulates singing-related neural activity in the songbird forebrain. Nat Neurosci. 1999; 2:209–11. [PubMed: 10195211]

Hikosaka O, Nakamura K, Sakai K, Nakahara H. Central mechanisms of motor skill learning. Curr Opin Neurobiol. 2002; 12:217–22. [PubMed: 12015240]

Ishii S, Yoshida W, Yoshimoto J. Control of exploitation-exploration meta-parameter in reinforcement learning. Neural Netw. 2002; 15:665–87. [PubMed: 12371519]

Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F. For whom the bird sings: context-dependent gene expression. Neuron. 1998; 21:775–88. [PubMed: 9808464]

Kao MH, Doupe AJ, Brainard MS. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. Nature. 2005; 433:638–43. [PubMed: 15703748]

Kao MH, Brainard MS. Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. J Neurophysiol. 2006; 96:1441–55. [PubMed: 16723412]

Kao MH, Wright BD, Doupe AJ. Neurons in a Forebrain Nucleus Required for Vocal Plasticity Rapidly Switch between Precise Firing and Variable Bursting Depending on Social Context. J Neurosci. 2008; 28:13232–47. [PubMed: 19052215]

Kobayashi K, Uno H, Okanoya K. Partial lesions in the anterior forebrain pathway affect song production in adult Bengalese finches. Neuroreport. 2001; 12:353–8. [PubMed: 11209949]

Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. Proc Natl Acad Sci U S A. 2011; 108:1687–92. [PubMed: 21220335]

Leblois A, Wendel BJ, Perkel DJ. Striatal dopamine modulates basal ganglia output and regulates social context-dependent behavioral variability through D1 receptors. J Neurosci. 2010; 30:5730–43. [PubMed: 20410125]

Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. Nature. 2008; 456:189–94. [PubMed: 19005546]

Meitzen J, Moore IT, Lent K, Brenowitz EA, Perkel DJ. Steroid hormones act transsynaptically within the forebrain to regulate neuronal phenotype and song stereotypy. J Neurosci. 2007; 27:12045–57. [PubMed: 17978046]

Nelson DA, Marler P. Selection-based learning in bird song development. Proc Natl Acad Sci U S A. 1994; 91:10498–501. [PubMed: 7937982]

Nordeen KW, Nordeen EJ. Anatomical and synaptic substrates for avian song learning. J Neurobiol. 1997; 33:532–48. [PubMed: 9369458]

Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol. 2005; 3:e153. [PubMed: 15826219]

Sakata JT, Hampton CM, Brainard MS. Social modulation of sequence and syllable variability in adult birdsong. J Neurophysiol. 2008 Apr; 99(4):1700–11. [PubMed: 18216221]

Sasaki A, Sotnikova TD, Gainetdinov RR, Jarvis ED. Social context-dependent singing-regulated dopamine. J Neurosci. 2006; 26:9010–4. [PubMed: 16943558]

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. 1991; 11:2896–913. [PubMed: 1880555]

Sossinka R, Böhner J. Song types in the zebra finch (Poephila guttata castanotis). Z Tierpsychol. 1980; 53:123–132.

Stepanek L, Doupe AJ. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. J Neurophysiol. 2010; 104:2474–86. [PubMed: 20884763]

Sutton, RS.; Barto, AG. Time derivative models of Pavlovian reinforcement. In: Gabriel, MR.; Moore, JW., editors. Learning and Computational Neuroscience: Foundations of Adaptive Networks. MIT Press; Cambridge, MA: 1990. p. 497-537.

Thompson JA, Basista MJ, Wu W, Bertram R, Johnson F. Dual pre-motor contribution to songbird syllable variation. J Neurosci. 2011; 31:322–30. [PubMed: 21209218]

Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature. 2007; 450:1240–4. [PubMed: 18097411]

Vu ET, Mazurek ME, Kuo YC. Identification of a forebrain motor programming network for the learned song of zebra finches. J Neurosci. 1994; 14:6924–34. [PubMed: 7965088]

Williams H, Mehta N. Changes in adult zebra finch song require a forebrain nucleus that is not necessary for song production. J Neurobiol. 1999; 39:14–28. [PubMed: 10213450]

Woolley SC, Doupe AJ. Social context-induced song variation affects female behavior and gene expression. PLoS Biol. 2008; 6:e62. [PubMed: 18351801]

Yanagihara S, Hessler NA. Modulation of singing-related activity in the songbird ventral tegmental area by social context. Eur J Neurosci. 2006; 24:3619–27. [PubMed: 17229110]
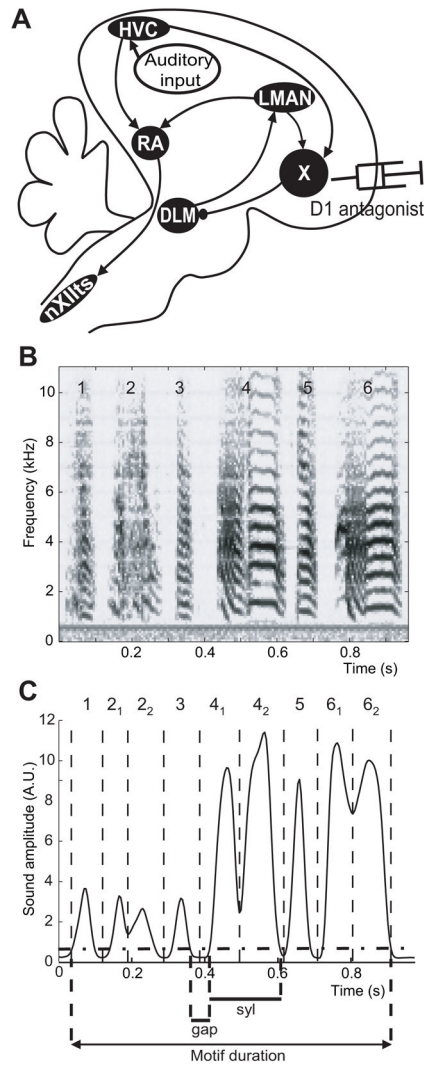
Figure 1. Example of a song motif divided into 6 syllables and 9 sub-syllabic elements
A) Schematic parasagittal representation of the song system. Dorsal is up, anterior is to the right. In the present study, The D1 receptor antagonist SCH23390 was infused into the BG nucleus Area X. B) Example of a song spectrogram of bird #4. This bird displayed six syllables, which are labeled on the top part of the spectrogram. C) Amplitude envelope of the song motif whose spectrogram is displayed in A. The envelope was used to divide syllables into sub-syllabic elements when applicable. Syllables 2, 4, and 6 of this bird where divided into two separate sub-syllabic elements each, because the sound amplitude reached two distinct maxima, with a decrease in amplitude between sub-syllabic elements. An amplitude threshold was chosen for each bird in order to determine the duration of motifs and the onset and offset of syllables, as indicated on the graph (horizontal dashed line).
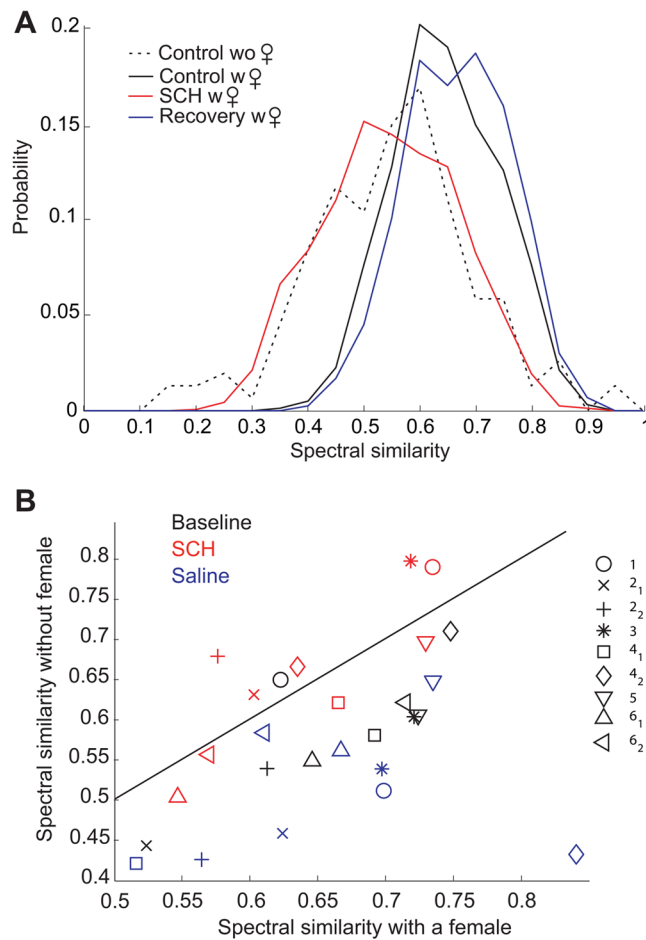
Figure 2. Infusion of the D1 receptor antagonist SCH23390 into Area X abolishes differences in spectral similarity due to social context: an example

A) Distributions of the spectral similarity index among pairs of renditions of the sub-syllabic element 6_1 of bird #4 at baseline (black lines) in the presence (solid line) and absence (dashed line) of a female, during the infusion of SCH23390 in Area X (red line) and during saline infusion into Area X (blue line), both in the presence of a female. Note that during SCH23390 infusion into Area X, the distribution of spectral similarity in the presence of a female is shifted toward lower similarity values, closer to their distribution when the bird was singing alone at baseline. B) Changes in spectral similarity with social context in all pharmacological conditions for each sub-syllabic element of bird #4. X-axis: average spectral similarity among pairs of renditions of a given sub-syllabic element in the presence of a female at baseline (black symbols), during SCH23390 infusion (red symbols) and during saline infusion (blue symbols) into Area X. Each sub-syllabic element is depicted with a different symbol (see Fig. 1 for the definition of each sub-syllabic element). Note that the spectral similarity among renditions of these elements was higher in the presence of a female at baseline and during saline infusion. On the contrary, during SCH23390 infusion in Area X, the difference between the two social contexts was damped and the red symbols are distributed around the diagonal.
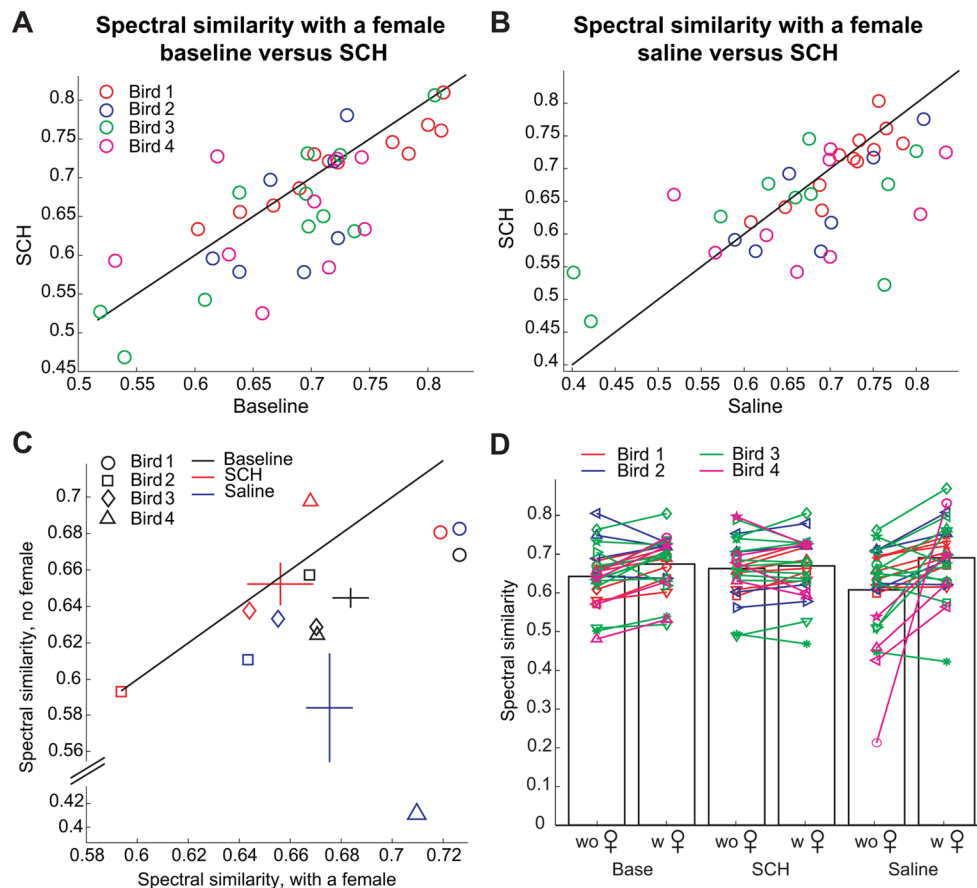
Figure 3. Infusion of the D1 receptor antagonist SCH23390 into Area X abolishes differences in spectral similarity due to social context: population data

A, B) Reduction in the spectral similarity of all sub-syllabic elements in the presence of a female under the infusion of SCH23990 in Area X as compared with the baseline condition (A) or with saline infusion in Area X (B). In A and B, each dot represents a sub-syllabic element, with an X-axis coordinate corresponding to its spectral similarity in the presence of a female before surgery, and a Y-axis coordinate corresponding to its spectral similarity in the presence of a female during the infusion of SCH23390 in Area X. Sub-syllabic elements drawn from different birds are represented with different colors. C) Context dependent changes in spectral similarity among all birds in three different conditions. Each bird is represented by a different symbol (circle: Bird #1, square: Bird #2, diamond: Bird #3, triangle: Bird #4), while each condition is represented by a different color (black: before surgery, red: during the infusion of SCH23390, blue: during saline infusion). X-axis coordinates correspond to the averaged spectral similarity in the presence of a female over all sub-syllabic elements. Y-axis coordinates correspond to the averaged spectral similarity in the absence of a female over all sub-syllabic elements. The crosses denote average spectral similarities across birds, with the length of the horizontal and vertical bars representing the SEM of the spectral similarities in the presence and absence of a female, respectively. Saline infusion followed SCH infusion in birds 1, 2 and 4. The difference in similarity with the social context observed at baseline was abolished during the infusion of SCH23990 into Area X, and tended to come back after saline infusion, although not significant. D) Context dependent changes in spectral similarity in three different pharmacological conditions among all non-harmonic sub-syllabic elements. In each paired graph, the left column refers to singing in the absence of a female, while the right column

refers to singing in the presence of a female. Each symbol represents a single sub-syllabic element. Left: before surgery; middle: during the infusion of SCH23390; right: during saline infusion. At baseline, the spectral similarity of the sub-syllabic elements was significantly increased when the birds sang in the presence of a female ($p<0.001$). This difference in similarity was abolished during infusion of SCH23990 into Area X, and returned after saline infusion ($p<0.01$).
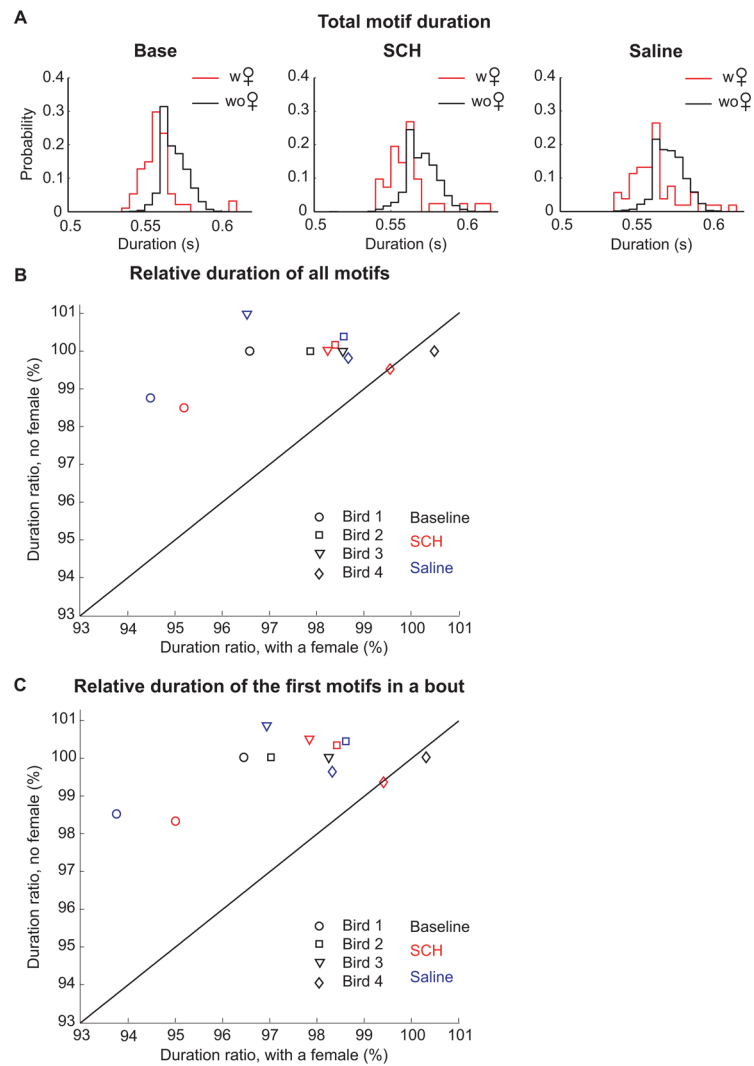
Figure 4. Infusion of the D1 receptor antagonist SCH23390 into Area X does not affect differences in song duration due to social context

A) Distribution of motif duration in bird #2 in the presence (red line) or absence (black line) of a female before surgery (left panel), during infusion of SCH23390 in Area X (middle panel) and during infusion of saline in Area X (right panel). The duration of song motifs was shorter in the presence of a female than when the bird sang alone in both control conditions and during the infusion of SCH23390 into Area X. B, C) Normalized duration of all song motifs (B) or in the first motif of each bout (C) in two different social contexts before, during and after infusion of SCH23390 in Area X. Each symbol represents the average normalized motif duration in one bird and in one pharmacological condition (black: baseline, red: infusion of SCH23390, blue: infusion of saline) in the presence of a female (x-axis) and alone (y-axis). In all cases, the average motif duration was normalized by the average motif duration in the absence of a female at baseline. Note that, except in bird 4, song motifs were shorter in the presence of a female, irrespective of the pharmacological condition. The difference in motif duration with social context was only marginally increased when restricting the analysis to the first motif of each bout.
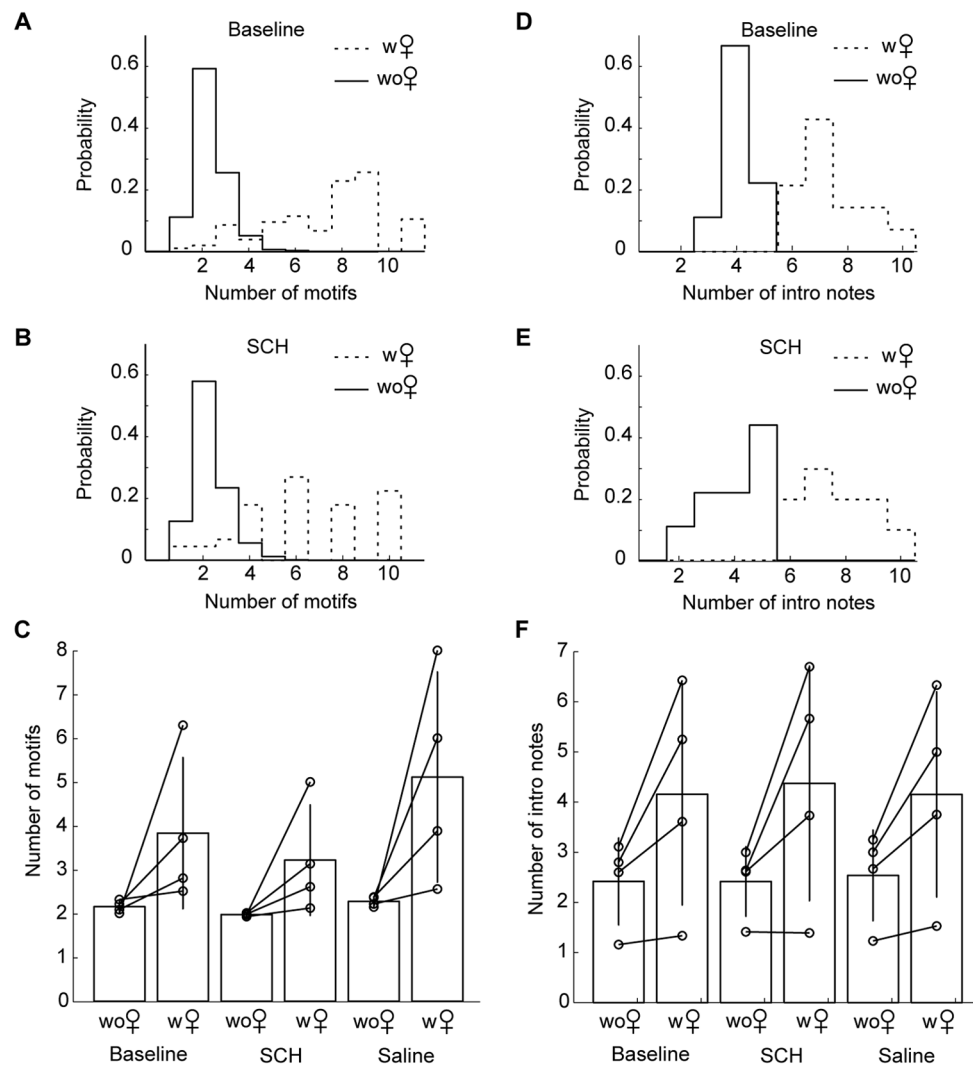
Figure 5. Infusion of the D1 receptor antagonist SCH23390 into Area X does not abolish differences in song structure due to social context

A) Distribution of the number of motifs per song bout in the presence (dashed line) or absence (solid line) of a female at baseline in bird #2. The number of motifs per song bout was strongly increased in the presence of a female. B) Distribution of the number of motifs per song bout in the presence (dashed line) or absence (solid line) of a female during the infusion of SCH23390 into Area X in bird #2. The number of motifs per song bout was still strongly increased in the presence of a female. C) Changes in the number of motifs per song bout with social context and pharmacological condition. In each paired graph, the left column refers to singing in the absence of female, while the right column refers to singing in the presence of a female. Left: before surgery; middle: during the infusion of SCH23390; right: during saline infusion. Over all birds, the number of motifs per song bout was strongly increased in the presence of a female, whether at baseline, during infusion of SCH23390 or saline in Area X. D) Distribution of the number of introductory notes per song bout in the presence (dashed line) or absence (solid line) of a female at baseline in bird #2. The number of introductory notes was strongly increased in the presence of a female. E) Distribution of the number of introductory notes per song bout in the presence (dashed line) or absence (solid line) of a female during the infusion of SCH23390 into Area X in bird #2. The

number of introductory notes was still strongly increased in the presence of a female. F) Change in the number of introductory notes with social context and pharmacological condition. Again, in each paired graph, the left column refers to singing in the absence of female, while the right column refers to singing in the presence of a female. Left: before surgery; middle: during the infusion of SCH23390; right: during saline infusion. The number of introductory notes was strongly increased in the presence of a female in all birds and all pharmacological conditions.