Timing during transitions in Bengalese finch song: Implications for motor sequencing

https://escholarship.org/uc/item/9h04q50p

Journal of Neurophysiology, 118(3)

0022-3077

Troyer, TW
Brainard, MS
Bouchard, KE

2017-09-01

10.1152/jn.00296.2017

Peer reviewed
Timing during transitions in Bengalese finch song: implications for motor sequencing

Todd W. Troyer, Michael S. Brainard, and Kristofer E. Bouchard

Department of Biology and Neurosciences Institute, University of Texas at San Antonio, San Antonio, Texas; Department of Physiology, University of California, San Francisco, California; Howard Hughes Medical Institute, San Francisco, California; and Biological Systems and Engineering, Lawrence Berkeley National Laboratory, Berkeley, California

Submitted 20 April 2017; accepted in final form 16 June 2017

Bengalese finch songs consist of strings of discrete syllables that can be organized into probabilistic sequences, with some syllables followed by two or more subsequent syllables (Okanoya 2004; Fig. 1). Such points of divergence may provide insight into the mechanisms controlling the “decision” to transition to one of several competing syllables (Bouchard and Brainard 2016; Bouchard et al. 2015; Jin 2009; Sakata and Brainard 2006, 2009; Sakata et al. 2008; Warren et al. 2012; Wittenbach et al. 2015). Conversely, convergence points where a syllable is preceded by two or more syllables enable the study of variations in song that depend on sequence context (Bouchard and Brainard 2013; Wohlgenuth et al. 2010).

Previously, Tachibana et al. (2015) showed that the length of the silent gap after syllables that were followed by a unique next syllable were significantly shorter than the intersyllable gaps at branch points. In a separate study, Matheson and Sakata (2015) showed that there was a negative correlation between the length of the silent gap between any given pair of syllables and the frequency of that transition relative to other possible transitions. Furthermore, increases in transition probability during early adulthood are associated with decreases in gap duration (Matheson and Sakata 2015), a relationship that has also been shown to hold in young zebra finches (Glaze and Troyer 2013).

One common explanation for the inverse relationship between transition probability and gap duration at divergence points is based on the idea of a “branched-chain” circuit, in which neurons active at the end of a branch syllable (divergence point) excite distinct neural populations that can trigger production of the possible subsequent syllables in the song sequence. The outcome of a winner-take-all competition among these populations determines the “winning” syllable (Jin 2009). Since a neural population receiving stronger synaptic drive is expected to both win the competition with higher probability and trigger the next syllable with shorter latency,
branched-chain connectivity has been posited as a possible neural mechanism leading to the observed negative correlation between transition probability and gap duration (Glaze and Troyer 2013; Matheson and Sakata 2015).

To investigate the relationship between syllable timing and syllable sequencing, we analyzed a large set of song recordings and compared the data with simulations of several models of competitive branching. Our results are presented in four sections. First, we reproduce and extend previous results demonstrating a negative correlation between transition probability and gap duration in Bengalese finches. Second, we analyze the relationship between timing and sequencing in several simple competitive branching models. We find that these models do not reproduce the strong negative correlation between transition probability and gap duration seen in our data. Third, we broaden our analysis of timing correlations and uncover a strong positive correlation in gap duration for transitions terminating at a given convergence syllable. Because transitions converging on a given syllable result from separate competitions, this novel result implies that the negative correlation between gap duration and probability cannot be explained by competitive mechanisms restricted to individual branch points. Finally, we pursue further analyses of the relationship between transition probability and gap duration at convergence points. The results suggest that the relationships between probability and gap timing may arise from a shortening in the latency to initiate syllables that are sung with higher probability.

MATERIALS AND METHODS

All analyses and simulations were performed using custom scripts written in MATLAB (MathWorks, Natick, MA). The code for generating all results and figures is included as a data supplement (Supplemental Material for this article is available online at the Journal website).

Data set. We analyzed song data from 31 Bengalese finches (Lonchuria striata domestica). Birds were provided with food and water ad libitum and maintained on a 14:10-h light-dark photocycle. Each bird was raised with a single adult male tutor. Data from some of these birds were also included in previous studies (Bouchard et al. 2015; Wittenbach et al. 2015). All procedures were performed in accordance with established animal care protocols approved by the University of California, San Francisco, Institutional Animal Care and Use Committee.

For sound recording, birds were housed individually in sound-attenuating chambers (Acoustic Systems, Austin, TX). An automated triggering procedure was used to record and digitize (44,150 Hz) several hours of singing from each bird. Song bouts were defined as continuous periods of singing separated by at least 2 s of silence. Song recordings were scanned to ensure that >50 bouts were obtained. Bengalese finch songs typically consist of 5–14 distinct acoustic elements, termed “syllables,” organized into probabilistic sequences (Fig. 1, A and D). Data were first filtered using a zero-phase, band-pass filter between 500 and 10,000 Hz. The signal was then squared and convolved with a 2-ms square window to determine the amplitude envelope (Fig. 1C). Songs were segmented into syllables by manual optimization of three thresholds. First, an amplitude threshold divided the song into segments of high amplitude. Then, subthreshold “gaps” of less than a threshold value were eliminated. Finally, any remaining periods of high-amplitude sound shorter than a threshold duration were eliminated. The threshold duration for gap elimination averaged 5.2 ms across birds (range 3–10 ms). The minimum duration for syllables averaged 19.2 ms across birds (range 9–30 ms). If a song bout contained ≥10 syllable iterations or was ≤2 s in total duration, it was not included.

The syllables from 12–54 song bouts per bird were hand labeled for subsequent analysis. This produced a data set of 79,803 syllables categorized into 301 distinct syllable types (900–6,704 syllables and 5–14 syllable types per bird). We determined the frequencies of pairwise transitions between syllables, yielding 78,919 transitions and 916 distinct transition types (886–6,629 transitions and 12–60 transition types per bird). For each syllable pair, we determined the duration of the gap between syllable offset and subsequent onset and determined the median gap length. To ensure that we had a reliable estimate of median gap length and to prevent infrequent transitions from skewing our analyses, we eliminated all transitions with five or fewer examples (343 transition types representing 607 total transi-

![Fig. 1. Probabilistic sequencing in Bengalese finch song. A: song spectrogram of 4 s within an 8.3-s-long song bout. B: raw acoustic signal. C. amplitude envelope. D: transition diagram showing syllable transitions. Thickness of line indicates transition frequency. E: gap length distributions for the two possible transitions following syllable k. Freq, frequency; amp, amplitude; prob., probability.](https://www.physiology.org/journal/jn)
populations drifting toward threshold: $d$ values. The third model considers mutual inhibition between two boundary is reached first. The bias toward one outcome is deter-
mined from a unit normal distribution ($\mathcal{N}(0, 1)$). The first model was a “race model,” in which two neural populations follow a random walk with drift: $dx_i/dt = W_i + \sigma_x$ for $i = 1, 2$, where $x_i$ is the state variable, $W_i$ is the drift rate (in Hz), $\sigma_x$ is a white noise process, and $\sigma$ is the noise level that determines the rate variance accumulation ($\sigma^2$ has units of state$^2$/ms). Simulations used a first-order Euler method: on each time step, $x_i$ is incremented by $W_i \Delta t + (\sigma \sqrt{\Delta t}) \eta_i$, where $\eta_i$ is drawn from a normal distribution ($\mathcal{N}(0, 1)$) with mean $0$ and variance $1$. State variables are con-
strained to the region $x_i \geq 0$. The first population to reach $x_i = 1$ wins the race. In all models, probabilities and timing statistics are deter-
mined from $N = 5,000$ trials.

The second model was the drift diffusion model (Ratcliff 1978). This model has a single state variable $x$ that represents the difference in activation in two populations. This variable follows a fixed latency so that the median decision time matched the target value exactly. We adjusted $\sigma$ by hand to ensure that the IQR matched the target value. For all models, we then found a level of drift asymmetry that gave a winning probability between 97.5 and 99%. Model behavior was determined from simulations that had drift weights equally spaced between the symmetric and extreme cases on a log scale.

Competitive branching models. Simulations were based on models of two alternative forced-choice decision tasks (Bogacz et al. 2006; Ratcliff and Smith 2004). The first model was a “race model,” in which two neural populations follow a random walk with drift: $dx_i/dt = W_i + \sigma_x$ for $i = 1, 2$, where $x_i$ is the state variable, $W_i$ is the drift rate (in Hz), $\sigma_x$ is a white noise process, and $\sigma$ is the noise level that determines the rate variance accumulation ($\sigma^2$ has units of state$^2$/ms). Simulations used a first-order Euler method: on each time step, $x_i$ is incremented by $W_i \Delta t + (\sigma \sqrt{\Delta t}) \eta_i$, where $\eta_i$ is drawn from a normal distribution ($\mathcal{N}(0, 1)$) with mean $0$ and variance $1$. State variables are con-
strained to the region $x_i \geq 0$. The first population to reach $x_i = 1$ wins the race. In all models, probabilities and timing statistics are deter-
mined from $N = 5,000$ trials.

The second model was the drift diffusion model (Ratcliff 1978). This model has a single state variable $x$ that represents the difference in activation in two populations. This variable follows a fixed latency so that the median decision time matched the target value exactly. We adjusted $\sigma$ by hand to ensure that the IQR matched the target value. For all models, we then found a level of drift asymmetry that gave a winning probability between 97.5 and 99%. Model behavior was determined from simulations that had drift weights equally spaced between the symmetric and extreme cases on a log scale.

Statistics. All regressions are reported with standard $P$ values and 10-fold cross validated $R^2$ ($R_{CV}^2$). Predicted outcome values for each 10% of the data were determined from a regression based on the other 90% of the data. $R_{CV}^2$ is equal to 1 minus the variance of the residual values divided by the total variance of the response variable. It quantifies the ability of the model to predict new data not used in training.

An $l_1$-regularized general linear model was used for multiple regression analysis (Lasso command in MATLAB with 10-fold cross validation). Parameter reliability was estimated by rerunning the analysis on data sets obtained from the original data set by random resampling with replacement ($N = 100$ resampled data sets). Data are reported as original Lasso estimates $\pm$ SE from the resampled estimates.

Several analyses focus on how transition-based parameters are grouped at convergence points. Consider two variables $X_{ai}$ and $Y_{ai}$, where $\alpha$ indexes the convergence syllables, $i$ indexes the $N_a$ transi-
tions converging on syllable $\alpha$, and there are $N$ total transitions. If we let $\bar{X}$ and $\bar{Y}$ be the mean of all transitions and $\bar{X}_a$ and $\bar{Y}_a$ be the mean values at convergence syllable $\alpha$, we can decompose the covariance between $X$ and $Y$ as follows:

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{\alpha} X_{ai} - \bar{X}(Y_{ai} - \bar{Y}) = \frac{1}{N} \sum_{\alpha} X_{ai} - \bar{X}_a(Y_{ai} - \bar{Y}_a) = \text{Cov}_{\text{within}}(X, Y) + \text{Cov}_{\text{between}}(X, Y)$$

where $\text{Cov}_{\text{within}}$ and $\text{Cov}_{\text{between}}$ are the between- and within-group co-
variances, respectively. If $X = Y$, then this is the usual decomposition of variance used in ANOVA. Dividing by the product of the SD of the two variables, $\sigma_x$ and $\sigma_y$, leads to a decomposition of the standard correlation coefficient, $r$, into a sum of two components, $r = r_{\text{between}} + r_{\text{within}}$:

$$r = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} = \frac{\text{Cov}_{\text{between}}(X, Y)}{\sigma_x \sigma_y} + \frac{\text{Cov}_{\text{within}}(X, Y)}{\sigma_x \sigma_y}$$

---

**Fig. 2.** Gap length and transition probability are negatively correlated. A: distribution of median gap lengths ($N = 573$). B: gap length vs. forward probability ($R_{CV}^2 = 23.41\%$, $P = 3.60 \times 10^{-3}$). C: gap length vs. backward probability ($R_{CV}^2 = 19.98\%$, $P = 1.31 \times 10^{-3}$). D: gap length vs. transition frequency: $R_{CV}^2 = 16.79\%$, $P = 2.93 \times 10^{-2}$). The inset bar in B shows the typical variation in gap distributions (the median IQR over all gaps). E–G: distributions corresponding to B–D.
RESULTS

Transition probability and gap length are negatively correlated. To investigate the relationship between sequence statistics and timing, we compiled a large corpus of hand-annotated Bengalese finch song data from 31 birds that contained 310 distinct syllable types and 573 distinct transitions (see Materials and Methods). We measured gap length on a log scale (base 10) and plotted median gap length vs. forward probability (Fig. 2B). The forward probability for the transition AB is defined as the number of AB transitions divided by the total number of transitions starting with A. The forward probability also has been termed the “divergence probability” (Bouchard and Brainard 2013; Wohlgemuth et al. 2010). Gap length and forward probability are negatively correlated (N = 573, P = 3.60 × 10⁻³⁶), with the linear regression accounting for 23.41% of the variation in log median gap length. The slope of the regression was −0.471, corresponding to a 41.87% decrease in median gap length per 50% change in probability. We quantified the width of the distribution for a typical gap length using the IQR of the log base 10 gap lengths. Across transitions, the median IQR was 0.0869 (vertical scale bar in Fig. 2B). Thus the width of a typical gap distribution is small relative to the change in median gap length across transitions, with the slope of the regression being 5.42 times as large as the median IQR. This means that the gap length distributions for high- and low-probability transitions have very little overlap. These findings confirm previous results showing that forward probability and gap length are negatively correlated (Matheson and Sakata 2015).

Forward transition probability is only one of several measures that can describe the statistics of variable sequencing. Forward probability is equal to the conditional probability of a transition given that a divergence syllable has just been sung. However, factors affecting probabilities and timing may include long-term plasticity and metaplasticity mechanisms such as postsynaptic weight normalization (Bouchard et al. 2015). Therefore we examined two additional sequence statistics and their relation to gap length: for transition AB, backward probability is defined as the number of AB transitions divided by the total number of transitions ending at syllable B. The backward transition probability has also been termed the “convergence probability” (Bouchard and Brainard 2013; Wohlgemuth et al. 2010). Transition frequency is defined as the number of AB transitions divided by the total number of transitions. Both statistics are also negatively correlated with gap length (backward probability: Fig. 2C, R²_CV = 19.98%, P = 1.31 × 10⁻³⁶; transition frequency: Fig. 2D, R²_CV = 16.79%, P = 2.93 × 10⁻²⁶).

To understand the relative strength of forward probability, backward probability, and transition frequency as predictors of gap length, we performed a multiple linear regression (Table 1, top row). Regressions were regularized using the Lasso technique, optimized for cross validated R² (see Materials and Methods). Forward probability had the greatest linear contribution to predicting gap length, followed by backward probability and then transition frequency.

We were concerned that the large number of transitions with probability equal to 1 (Fig. 2, E and F) might have an undue impact on the linear regression. Therefore we segregated transitions into four groups based on whether forward probability and backward probability were equal to 1: transitions with forward and backward transition probabilities both <1 are termed “nondeterministic” (N = 357); transitions with forward and backward transition probabilities both equal to 1 but forward probability <1 are termed “forward deterministic” (N = 74); and transitions with backward probability equal to 1 but forward probability <1 are termed “backward deterministic” (N = 79). Performing a multiple linear regression of forward probability, backward probability, and transition frequency vs. gap duration for nondeterministic transitions resulted in a marked reduction in the explained variance and a large change in the factor weightings (Table 1, bottom row). The difference between the regression results for the full and restricted data sets indicates that short gap lengths for transitions with probability equal to 1 (Tachibana et al. 2015) are not due to the simple extrapolation of a linear relationship between probability and gap length.

Considered as groups, we found that forward and backward deterministic gaps were significantly shorter than nondeterministic gaps (P = 6.63 × 10⁻⁷ for forward vs. nondeterministic; P = 5.31 × 10⁻⁹ for backward vs. nondeterministic; 2-population t-test; Fig. 3A). Moreover, deterministic gaps were significantly shorter than either forward or backward deterministic gaps (P = 1.24 × 10⁻⁸ for forward deterministic vs. deterministic; P = 9.36 × 10⁻⁹ for backward deterministic vs. deterministic). Forward and backward deterministic gaps were not significantly different in length (P = 0.587).

For forward and backward deterministic transitions, we also investigated the relationship between gap length and probability for the “other” direction; that is, we plotted forward probability vs. gap length for backward deterministic transitions, and backward probability vs. gap length for forward deterministic transitions (Fig. 3, B and C). We found that having probability equal to 1 in either direction largely eliminated the relationship between gap length and transition probability in the other direction (N = 79, R²_CV = −0.038%, P = 0.213 for backward deterministic transitions; N = 74, R²_CV = −0.036%, P = 0.417 for forward deterministic transitions). Thus the mechanism generating the negative correlation between transition probability and gap length is confined to nondeterministic transitions.

Given our intent to investigate the relationships between sequence statistics and timing, our remaining analyses will be

<table>
<thead>
<tr>
<th>N</th>
<th>Forward Probability</th>
<th>Backward Probability</th>
<th>Transition Frequency</th>
<th>Total R², %</th>
</tr>
</thead>
<tbody>
<tr>
<td>All data</td>
<td>573</td>
<td>−0.300 ± 0.008</td>
<td>−0.226 ± 0.008</td>
<td>−0.154 ± 0.004</td>
</tr>
<tr>
<td>Fwd. and bwd. prob. &lt;1</td>
<td>341</td>
<td>−0.172 ± 0.025</td>
<td>0.126 ± 0.028</td>
<td>−0.355 ± 0.022</td>
</tr>
</tbody>
</table>

Values are means ± SE from N = 100 bootstrapped resamples. Predictor variables were transition frequency, forward probability, and backward transition probability. All data, all transitions. Fwd. and bwd. prob. <1, transitions with forward and backward probability <1.
restricted to the 357 nondeterministic transition types that connect divergence points to convergence points (46,754 total transitions). We will also focus on the 95 out of 163 divergence points in which all outgoing transitions terminate at convergence points and on the 90 out of 155 convergence points in which all incoming transitions originate at divergence points.

Our analyses are consistent with previous results showing that forward transition probability is negatively correlated with gap duration (Matheson and Sakata 2015) and that deterministic transitions have short gaps (Tachibana et al. 2015). We additionally have shown that the short gaps associated with deterministic transitions are not due to a simple extrapolation to a probability of 1 of the linear relationship between gap duration and probability (derived from fits to the data for probabilities <1).

Simple competition models lead to overlapping distributions of winning times. To investigate possible circuit mechanisms that might result in a negative correlation between transition probability or frequency and gap length, we turned to computational models. The main mechanistic hypothesis put forward to explain probabilistic sequencing in finches is a branched chain (Jin 2009), in which the termination of a divergent syllable initiates a competition between distinct neural populations, each of which can trigger the production of a subsequent syllable. To investigate these ideas, we simulated several competition models based on classical models of decision making in two-alternative forced-choice tasks (Bogacz et al. 2006; Ratcliff and Smith 2004).

We started by considering a “race” between two populations, each of which follows an independent buildup of activity governed by a random walk process with drift. The first population to reach a threshold value of activity wins the competition and determines the next syllable in the sequence. It has been suggested that stronger connections between syllable representations will lead to a process that reaches threshold with shorter latency and hence wins the race more often, leading to a negative correlation between latency and transition probability (Glaze and Troyer 2013; Matheson and Sakata 2015).

To explore the relationship between forward transition probability and latency, we fixed the drift rate for one population (Fig. 4, A–D, blue) and varied the drift rate for the other population (Fig. 4, A–D, red) so that the winning probability covered at least 95% of the range between 0 and 1. Noise and drift parameters were set to match the typical median and IQR of the data when the two drift rates were equal, leading to winning probabilities of 50% (see MATERIALS AND METHODS).

Figure 4A shows the relationship between median time to threshold (T-thresh) and the drift rate for the red population (Drift-red). The shaded area shows the IQR of the timing distributions. Here, the time to threshold is measured without considering the competition. The inset shows the probability of winning for both populations vs. the drift rate for the red population. As expected, the faster process wins the competition more often. Combining the two effects yields the expected negative relationship between forward probability and median latency to threshold for the red population (Fig. 4B).

Figure 4, A and B, embodies the main intuition behind previous explanations for the negative correlations between forward probability and gap length.

However, this explanation fails to consider the effect of competition on timing. Figure 4C shows the drift rate of the red population vs. the observed winning times (T-win) for the two processes. Note that even though the drift rate is only changing for the red population, the winning times for both populations roughly follow whichever population has the fastest drift rate. The explanation for these results comes from noting that the winning-time distribution excludes threshold crossings that are too slow to win the race. Thus the transition with fastest drift rate largely dictates the winning time for both populations.

Plotting the winning times vs. probability for both transitions (Fig. 4D) reveals a shallow dependence of winning time on probability for the red population and a positive relationship between winning time and probability for the blue population. This positive relationship stems from the fact that the blue
population loses the competition when the red drift rate is fast. In this case, the observed winning times from the blue populations are those that are fast enough to outcompete the relatively fast red population.

Taken together, the negative relationship between the red winning time and probability is slightly stronger than the positive correlation for the blue population, and a linear regression fit to all the data has a negative slope. However, the relationship between winning time and probability is much weaker than the correlation between threshold crossing time and probability.

For a more complete comparison between the model and our behavioral data, we varied drift rates for both populations, covering a regular grid of values (Fig. 4E). A plot of probability vs. winning time for the entire grid is shown in Fig. 4F. The orange line and shaded area show the median and IQR of winning times for simulations where the ratio of the drift rates is changing but the overall geometric mean of the drift rates is constant (orange diagonal in Fig. 4E). The black line shows the linear regression fit of all the data.

To compare the simulation results with our song data, we found all branch points in our song data that had two transitions. For these 180 transitions, we found that the slope of the regression line from our simulations was 0.08345, yielding a slope-to-IQR ratio of 1.42. The median IQR of these transitions was 0.0855, yielding a slope-to-IQR ratio of 1.35. These values are similar to the values from our full data set (Fig. 2B) and indicate that typical distributions of gap lengths for low- and high-probability transitions have little overlap. The regression line is shown in green in Fig. 4, F–H, and is surrounded by the median IQR. In contrast, the slope of the regression line from our simulations was −0.0499, and the median IQR was 0.08345, yielding a slope-to-IQR ratio of −0.598.

We explored two other decision models, both of which consider competitive effects between populations before reaching decision threshold. The classical drift diffusion model is based on a single underlying state variable representing the relative level of activation in the two populations. The dynamics follow a random walk, and the two possible transitions are triggered when the variable reaches either +1 or −1. Nonsymmetric outcomes can be implemented by adding a consistent drift toward one or the other boundary. Figure 4G shows probability vs. winning time for simulations at varying drift rates. Linear regression reveals little relationship between timing and probability in this model (slope = 0.000832).

We also considered a mutual inhibition model in which increased activity in one population caused proportional inhibitory drive to the other population (see MATERIALS AND METHODS). We performed simulations using a grid of excitatory weights (as in Fig. 4E) for each of several levels of inhibition. Figure 4H shows results when the inhibitory weight was set to twice the excitatory drive for simulations at the center of the grid. At this relatively high level of inhibition, winning time is largely determined by the relative magnitude of excitatory drive to the two populations, yielding a one-dimensional relationship between probability and winning time. Adding this strong inhibition increased the inverse relationship between timing and probability, but the slope and slope-to-IQR ratio were still much smaller than in the data (slope = −0.113; slope-to-IQR ratio = −1.42).

In our simulations of competitive decision models, the inverse relationship between winning time and probability was either not present or was substantially weaker than seen in our behavioral data. In all models considered, winning-time distributions had substantial overlap between the two outcomes, whereas the slope of the regression between probability and
gap duration in our data was much steeper than the spread in gap durations for a typical transition. These negative results do not rule out competition between neural populations as a mechanism for probabilistic sequencing, but they suggest that such models do not naturally account for the strong negative correlation between gap duration and probability observed in Bengalese finch song data.

There is strong, shared variation at convergent syllables. The difficulty we had in matching the models to song data led us to think more carefully about the sequence of events that occurs during the silent gap between syllables. In particular, it is an oversimplification to equate the entire gap with the duration of a competitive process for syllable selection. Most basically, one can divide the gap between syllables $A$ and $Y$ into three distinct epochs: 1) the period after the termination of syllable $A$ but before the competition begins, 2) the period of the competition, and 3) the period between the end of the competition and the onset of syllable $Y$ (Fig. 5). The postcompetition period can be further broken into two: a period of time in which the subsequent syllable has been determined but neural activity is still influenced by the previous syllable (epoch 3a) and a period before the onset of the selected syllable that is relatively invariant across transitions (epoch 3b). Variation in the median duration of epoch 1 across divergence points should show up in our data as correlations in the gap lengths of syllables that share a divergence point ($AX$ and $AY$ in this example). Similarly, variation in the median duration of epoch 3b across convergence points should show up as correlations in the gap lengths of syllables that share a convergence point ($AY$ and $BY$ in this example).

To look for the correlations corresponding to variations in epoch 1, we focused on divergent syllables and found a significant positive correlation between the median gap length of a given transition and the average of the median gap lengths for all other transitions sharing that divergent syllable (Fig. 6A; $N = 241, R^2_{CV} = 9.50\%, P = 1.12 \times 10^{-7}$). Similarly, we examined the correlations corresponding to variations in epoch 3b by correlating median gap duration for a given transition with the average median gap length for transitions sharing that convergent syllable. Here we found a very strong correlation, with the other gap lengths at a convergent syllable able to predict nearly 60% of the variance in gap length (Fig. 6B; $N = 249, R^2_{CV} = 58.1\%, P = 5.74 \times 10^{-50}$).

This is the central finding of this paper, and it has two important implications. First, the shared variation at convergent syllables suggests that the largest contribution to gap length stems from the latency between the time at which the transition to a given syllable has been finalized and the onset of that syllable (epoch 3b in Fig. 5). Second, the bulk of the variation in gap durations cannot be directly attributed to competitive processes that take place at individual branch points. This is because the competition that results in a transition from $A$ to $Y$ is distinct from the competition that leads to a different transition to $Y$, such as $B$ to $Y$ (Fig. 5).

Forward transition probability is not correlated at convergent syllables. One possible explanation for the negative correlation between probability and gap duration is the existence of a third factor that affects both variables but in opposite directions. As an example of such a third factor, it is possible that syllables have varying degrees of "excitability" with highly excitable syllables having shorter gap durations and greater probability of winning all competitions that have that syllable as a possible outcome. Under this hypothesis, one expects to find shared variation in forward probability as well as in timing, since all transitions converging on an excitable syllable should have more success in winning their respective competitions.

To look for shared variation, we calculated the "within-group" variance for transitions sharing a convergent syllable (as in an ANOVA) and expressed this within-group variance as a fraction of the total variance. If forward probability or gap timing take on similar values at a given convergent syllable, we expect within-group variances for that variable to be small. To quantify the within-group variance expected if there were no

---

Fig. 5. Three epochs during the gap between syllables: 1) before competition, 2) period of competition (box), and 3) period after competition before syllable onset. Epoch 3 can be further subdivided by whether neural activity is (epoch 3a) or is not (epoch 3b) influenced by the previous competition.

Fig. 6. Gap length vs. mean of other gaps at branch points. A: divergent syllables ($N = 241, R^2_{CV} = 9.50\%, P = 1.12 \times 10^{-7}$). B: convergent syllables ($N = 249, R^2_{CV} = 58.1\%, P = 5.74 \times 10^{-50}$).
and gap length can be attributed to the negative covariance between the mean forward probability and mean gap length at convergent syllables: $r = -0.383$, $r_{\text{btw}} = -0.259$, $r_{\text{within}} = -0.123$, and $r_{\text{btw}}/r$ = 67.7% (249 transition types at 90 convergence points).

We expect some degree of correlation between group means for any two correlated variables, even if the grouping is independent of the correlation. To test whether the negative correlation between probability and gap duration depends on shared covariation at convergent syllables, we randomly shuffled transitions, decomposed the correlation $r = -0.383$ into $r_{\text{btw}} + r_{\text{within}}$, and repeated the process 1,000 times. The measured between-group correlation accounted for a greater proportion of the negative correlation than every shuffled value, indicating that the correlation between forward probability and gap duration has a significant dependence on the grouping of transitions at convergent syllables (Fig. 8; $P < 0.001$, actual data $r_{\text{btw}}/r = 67.7$%; shuffled data $r_{\text{btw}}/r = 36.18 \pm 8.39$%, mean $\pm$ SD).

In sum, our data show that variations in gap duration are correlated at convergent syllables and that a significant fraction of the negative correlation between probability and gap duration is shared at convergent syllables. In contrast, forward probability does not display shared variation at convergent syllables. The hypothesis that is most consistent with these analyses is that transition probabilities are established first, followed by a reduction in latency for those syllables that happen to be visited with high probability (see DISCUSSION). According to this hypothesis, one expects average gap duration for all transitions terminating at a given convergent syllable to be negatively correlated with the frequency of that syllable. Indeed, frequency and average gap duration are negatively correlated across convergent syllables ($N = 90$; $R^2_{\text{CV}} = 12.83\%$, $P = 1.72 \times 10^{-5}$).

**DISCUSSION**

We investigated the relationship between sequencing and timing in Bengalese finch song to gain deeper insight into the mechanisms of probabilistic sequence generation. The starting...
point for our investigation was the negative correlation between forward probability and the length of the gap of silence between the two syllables defining that transition (Fig. 2B). It had been suggested that this relationship could result from a branched-chain model in which the termination of a divergent syllable initiates a buildup in activity within several neural populations, with syllable transitions triggered by the first population to reach a threshold for action. Under this mechanism, stronger input to a given population is expected to result in a shorter latency to activation and a higher probability of winning the race, leading to a negative correlation between forward probability and transition time (Glaze and Troyer 2013; Matheson and Sakata 2015).

To investigate this idea, we performed numerical simulations of several models of competitive branching. In general, these models result in a surprisingly weak relationship between probability and timing. This is because the only transition times that are observed are those that are short enough to win against all competitors. Therefore a population with a relatively long latency to reach activation threshold in the absence of competition will have a lower probability of winning, but the distribution of winning times for that population will have a large overlap with the winning times for high-probability options.

Overlapped timing distributions are seen in reaction time data for correct and error trials in human two-alternative forced-choice experiments (Ratcliff and Smith 2004). However, the difference in gap durations between low- and high-probability outcomes in our song data is usually more than three times the typical IQR. Thus, although branched chaining may be the mechanism leading to probabilistic transitions, it does not provide a natural explanation for the strong link between forward probability and gap timing in Bengalese finch song.

Our simulation results forced us to think more carefully about the sequence of neural processes during the gaps between syllables (Fig. 5) and to more closely examine the correlation structure in our data. Our central finding is that roughly 60% of the variance in gap length can be predicted by knowing the gap lengths for other transitions converging on a common syllable (Fig. 6B). Because transitions that share a common target syllable occur at distinct branch points in the song sequence, variations in transition timing cannot be attributed to syllable selection mechanisms acting independently at distinct divergence points.

The magnitude of the shared variance at convergence points also suggests that variations in gap timing are dominated by the latency between the end of the competition period and the onset of vocalization of the syllable that ends the gap. It follows that any competitive process determining sequence transitions is likely to be completed relatively early during the gap period between syllables. An early period of competition followed by a relatively longer period of latency to syllable initiation is reminiscent of variations in patterns of inspiration during singing (Andalman et al. 2011). Both zebra and Bengalese finches take small inspiratory minibreaths during most syllable transitions (Cooper et al. 2012; Goller and Cooper 2004). Andalman et al. (2011) showed in zebra finches that these minibreaths are characterized by a sharp initial increase in inspiratory air sac pressure, followed by a slow decline that is terminated by a sharp transition toward expiratory pressure at the end of the minibreath. For a given bird, the rapid rise in inspiratory pressure is stereotyped across transitions, with variations in duration mostly due to changes in the rate of the slow decline. The parallels between these results and ours suggest that the competition process at branch points may overlap with the period of the initial rise in inspiratory pressure during inspiration. Our data also suggest that the largest fraction of the gap is occupied by processes related to the syllable terminating that transition. Consistent with this, when juvenile zebra finches learn to sing a transition from A to B where syllables A and B are copied from different tutors, the gap between A and B corresponds most closely to the song of the tutor singing syllable B (Williams and Staples 1992).

The shared variation in gap length at convergent syllables only concerns timing. It does not explain the negative correlation of gap length with forward probability. Therefore we investigated whether the grouping of variables at convergent syllables applied to forward probability and gap length and whether the correlation between gap length and forward probability depended on grouping at convergent syllables. In contrast to gap length, we found that the variation in average forward probability at convergent syllables is consistent with a random assignment of forward probabilities (Fig. 7B). However, we found that the correlation between forward probability and gap duration does group at convergent syllables, showing much stronger correlation between the mean probability and mean gap length at convergent syllables than expected from chance (Fig. 8).

There are four basic causal explanations for the negative correlation between probability and gap length: 1) higher probability leads to shorter gaps, 2) a short gap leads to higher probability, 3) a third factor leads to shorter gaps and higher probability, or 4) more complex causal links such as bidirectional interactions between gap duration and probability lead to negative correlations between these variables. We have shown that both gap length and the correlation coefficient between gap length and probability group at convergent syllables, whereas transition probabilities do not. These results favor causal explanation 1: transition probabilities are established first, subsequently leading to alterations in gap durations that are shared by transitions that converge on the same syllable. The three other explanations posit causal influences that either flow from gap duration to probability or are shared by both gap length and probability. In these cases, the grouping of gap duration at convergent syllables would result in a grouping of probability that is inconsistent with our data.

A simple neural mechanism linking probability and gap duration in a manner consistent with our data is for syllable latencies to be reduced with repeated singing, with greater latency reductions for syllables that are visited with higher probability. Consistent with this notion, we find that syllable frequency and the mean gap duration at convergent syllables are negatively correlated. Previous studies show that gap durations decrease significantly after the initial period of song crystallization in zebra finches (Brainard and Doupe 2001) and the magnitude of the decrease in gap duration is correlated with the increase in transition probability (Glaze and Troyer 2013). Bengalese finches also show correlated changes in transition probability and gap durations over the first year of life (Mathe-son and Sakata 2015). Presumably, frequency-dependent changes require the bird to sing many song renditions to affect syllable latency. In addition to tracking changes over develop-
ment, Matheson and Sakata (2015) found that introducing a female bird resulted in short-term changes in both forward probability and gap duration in a male’s song. However, in contrast to developmental changes, the direction of these acute changes in probability were not systematically related to the changes in gap duration.

Although we have focused on the shared variation in gap length at convergent syllables, we found other correlations that explain a smaller fraction of gap length variance. For example, the shared variation at divergent syllables suggests that gap length is affected by variations in latency between syllable offset and the start of the competition process determining the next syllable transition. Also, a portion of the negative correlation between forward probability and gap duration remains even after accounting for the correlation between average forward probability and average gap duration at convergent syllables. This suggests that the competition process might indeed resolve more quickly for more probable transitions. Finally, it should be noted that although the correlation in timing at convergent syllables is very strong (60% of variance explained), the relationship with probability or frequency parameters is modest (<25% of variance explained). This suggests that a significant fraction of the variation in postcompetition syllable latency is due to factors unrelated to syllable frequency or transition probability.

Most of our analyses focus on the inverse relationship between timing and sequencing for transitions connecting divergence points to convergence points. We also examined transitions in which either the forward or backward probability was equal to 1, i.e., transitions that precede or follow stereotyped sequences of syllables. These transitions do not show a significant negative correlation between timing and probability (Fig. 3), suggesting that mechanisms affecting syllable transitions around stereotyped sequences may be distinct from the mechanisms that govern transitions in the highly branched portions of the song. There is evidence that such stereotyped “chunks” of Bengalese finch song are treated as functional units in both the motor and perceptual domains (Seki et al. 2008; Suge and Okanoya 2010; Takahasi et al. 2010). If the stereotyped chunks of Bengalese finch song are analogous to the stereotyped motifs in zebra finch song, it is possible that the mechanisms leading to the negative correlation between gap duration and probability analyzed here may be most analogous to mechanisms controlling motif initiation in zebra finches (Glaze and Troyer 2006, 2012; Horita et al. 2008; Rajan and Doupe, 2013).

Despite the fact that syllable branching occurs during periods of vocal silence, we have exploited the rich branching structure of Bengalese finch song to constrain the possible mechanisms governing motor sequencing. In particular, our models of vocal silence, we have exploited the rich branching (Glaze and Troyer 2006, 2012; Horita et al. 2008; Rajan and Doupe, 2013).

REFERENCES


