Supplementary Results

End-of-motif transitions

In the 128 stimulation trials that we performed with Uva → HVC_I connections (Fig. 5B), 8 showed a transition to another syllable at the end of the motif rather than a stop at the end of the motif, within the 350 ms simulation. In the 128 stimulation trials that we performed without Uva → HVC_I connections (Fig. 5C), 26 showed such a transition. The reason for this difference is unknown. In 30 unstimulated trials with Uva → HVC_I connections, only 1 showed such a transition, while in 30 unstimulated trials without Uva → HVC_I connections, only 2 showed such a transition.

Timing dependence of stimulation effects

As Fig. S3 shows, we saw evidence for timing dependence of stimulation effects, and we observed different timing effects depending on whether we included Uva → HVC_I connections or not.

In Fig. S3, the corrected stimulation time is the approximate time of the stimulation's effect on HVC relative to the onset of the current syllable. More specifically, it is the time of stimulation, plus the mean delay to HVC in the model (9.9 ms for RA, 8.1 ms for DM, and 5.6 ms for Uva), minus the mean onset time of the most recent HVC_{RA} syllable. Because we only stimulated at 25, 50, 75, 100, 125, 150, 175, and 200 ms relative to the start of the simulation, all

corrected stimulation times for the feedback pathway fell within one of three 6 ms bins: 0 to 6 ms, 25 to 31 ms, or 50 to 56 ms.

Fig. S3 shows the grouped effects for all trials in which we stimulated the feedback pathway. In the model with Uva $\rightarrow$ HVC_I connections (Fig. S3A), we observed no distortions when the corrected stimulation time was in the 50 to 56 ms bin, significantly fewer than the number in the other two time bins ($p < 0.05$, Liddell's exact test). We also observed significantly fewer transitions in the 0 to 6 ms time bin than in the other two bins ($p < 0.05$, Liddell's exact test). Corrected stimulation time is defined in Methods.

In the model without Uva $\rightarrow$ HVC_I connections (Fig. S3B), there were no significant differences among the numbers of observed distortions at the three stimulation times ($p > 0.1$, Liddell's exact test). However, stimulation of the feedback pathway at a corrected time of 50 to 56 ms evoked significantly more transitions than did stimulation at 0 to 6 ms ($p = 4.3\times10^{-2}$, Liddell's exact test) or 25 to 31 ms ($p = 3.8\times10^{-6}$, Liddell's exact test), and stimulation at 0 to 6 ms evoked significantly more transitions than stimulation at 25 to 31 ms ($p = 6.7\times10^{-3}$, Liddell's exact test; the latter did not evoke any transitions). The reason for these differences is unknown; further analysis of this is beyond the scope of the present paper.

Compensation for feedback delay

The brief pause in HVC activity at transitions in our model (4.6 ± 1.0 ms; n = 10) can probably be accounted for by the difference between the time it takes the excitation from Uva to

bring the HVC_{RA} initiation neurons of the next syllable network to spiking threshold and the time it takes the feedforward inhibition from Uva to suppress the previous syllable network. The time from the beginning of the depolarization to the first spike peak in the HVC_{RA} initiation neurons is 5.9 ± 0.7 ms (n = 10). In the version of our model without Uva → HVC_I connections, the pause is absent: the first spike of an HVC network occurs 1.0 ± 0.5 ms *before* the last spike of the previous HVC network (n = 10).

HVC cooling

The discrepancy between syllable and gap expansion during HVC cooling may be small. For example, suppose that the syllable duration is 150 ms, the gap duration is 30 ms, the feedback delay is 24 ms, and $t_{delay1} = t_{delay2}$. If HVC activity is slowed by 50%, and if the gap is controlled by the beginning of the syllable network, then the new gap duration will be 45 ms (a 50% expansion), while the new syllable duration will be 213 ms (a 42% expansion). These considerations predict that the shorter the feedback delay, the more similar the expansion of gaps and syllables will be.

Supplementary Discussion

The syllable-specific populations of neurons in dRA, DM/PAm, and Uva could potentially be spatially organized as “syllabotopic” maps, i.e. topographic maps organized by syllable.

As we have discussed, the syllable-specific input to HVC is necessary during variable sequences, in order to maintain interhemispheric coordination. We implement the simplest version of the model, in which the sequence in HVC is controlled by syllable-specific input no matter how stereotyped the syllable sequence is. In a more complex version of this model, HVC input could be syllable unspecific when the sequence is stereotyped and syllable specific when it is variable. It is possible that the dRA $\rightarrow$ brainstem $\rightarrow$ Uva $\rightarrow$ HVC feedback loop provides only syllable-unspecific timing pulses and that during states of sequence variability, a bilaterally coordinated MMAN $\rightarrow$ HVC input or NIf $\rightarrow$ HVC input could play the central role in promoting syllable transitions.

The model assumes that both syllables and intersyllable gaps are controlled by parts of the HVC_{RA} networks. In principle, gaps could be controlled by HVC_{RA} neurons anywhere along the HVC_{RA} network. However, given the evidence that syllables are units of song (see Introduction), it is reasonable to assume that gaps are controlled by the beginnings and/or ends of HVC_{RA} networks, near HVC syllable transition points. In the present model, gaps are controlled by the ends of HVC_{RA} networks (Fig. 2D)

Even without the Uva $\rightarrow$ HVC_I connections, the model contains mechanisms for suppressing persistent activity at the ends of syllables and motifs. Each syllable network inhibits the previous network, preventing persistent activity at the ends of syllables; and stimulation of the beginnings of all syllable networks by Uva at the end of a motif causes brief activity in all of them, which causes inhibition of all of them via the HVC_I neurons and terminates their activity.

The dramatic rewiring that is the basis for the stuttering in Fig. S1 is unlikely to occur in real birds, but this simulation is suggestive of a more likely possibility, in which HVC syllable network 3 sends a weaker projection to $dRA_{DM/PAm}$ subpopulation 2 as well as the projection to $dRA_{DM/PAm}$ subpopulation 3. Variable excitation of $dRA_{DM/PAm}$ neurons, perhaps due to synaptic input from lateral LMAN (Johnson et al. 1995), would permit subpopulation 2 to be activated with some probability, resulting in a stutter.

Our model is not designed to simulate the distortions observed by Ashmore et al. (2005), since it does not yet include respiration, sound production, or projections either from DM/PAm to RAm and nXIIts or from PAm and RAm to spinal motor neurons innervating respiratory muscles.

As Hahnloser et al (2002) suggest, the assumption that that repeated syllables are associated with repeated bursts in HVC_{RA} neurons is consistent with the fact that repetitions of a syllable type are associated with very similar ensemble burst patterns in RA (Leonardo and Fee 2005).

Experimentally, selective stimulation of a single subpopulation is likely to be very difficult, unless such subpopulations are arranged topographically. The model results suggest that only by *both* using a long stimulus *and* stimulating a single subpopulation can one evoke a large fraction of syllable transitions rather than truncations and song stops. Additionally, it is possible that our model is only partly correct—that some of its predictions are accurate, while

others are the result of perturbing the model beyond its range of applicability. The model will have served its purpose if it prompts researchers to look at the system in a new way and to devise and perform new experiments. Future experiments will enable the refinement of the model and the generation of new experimental predictions which are both more accurate and more precise.

Since MMAN may play part of the role we have proposed for Uva (see Introduction and Fig. 1C), it is possible that stimulating MMAN or its afferent thalamic nucleus, DMP, may evoke the syllable transitions that our model predicts for stimulation of the brainstem feedback pathway. For MMAN, too, our model results suggest that syllable transitions are much more likely to be evoked with longer stimulation of specific subpopulations.

Supplementary Figure Legends

FIG. S1. Stuttering of the last syllable, generated when RA subpopulation 2 is more strongly excited than RA subpopulation 3 by neurons near the end of HVC syllable network 3. In this simple example, all of the synapses from neurons near the end of HVC syllable network 3 excite RA subpopulation 2.

FIG. S2. *A:* Repeated neural "motifs" can be generated by continuous stimulation of the beginning of the first HVC_{RA} syllable network. We stimulate the first neuron of the first syllable network in HVC (*HVC_{RA} Neuron 1*; voltage trace shown above raster plot) with a current lasting from t = 0 to t = 500 ms, which could represent continuous excitatory synaptic drive to the

initiation region of the first syllable network ("intention" to sing). HVC_I neurons inhibit this neuron as activity propagates along the three syllable networks (*"Motif 1"*). When this activity ceases, the constant current causes the first neuron to reach threshold again and initiate a new neural motif (*"Motif 2"*). The neural patterns of the two motifs are the same. In one of 10 simulations, the model also exhibited a stutter of syllable 1.

FIG. S3. Timing of stimulation in the feedback pathway (same data as for Fig. 5). *Corrected stimulation time:* approximate time of the stimulation's effect on HVC relative to the onset of the current syllable (see text). Percentage of simulations showing HVC_{RA} distortions (*distort*), truncation and song stop (*truncate & stop*), or syllable transition (*transition*) for corrected stimulation times in three 6 ms bins. *A.* Model with Uva $\rightarrow$ HVC_I connections. *B.* Model without Uva $\rightarrow$ HVC_I connections.

FIG. S4. Effect of HVC cooling. *A:* Activity of all HVC_{RA} neurons in a normal simulation of the type shown in Fig. 3. The last cluster is the last to exhibit spiking. *B:* Activity of all HVC_{RA} neurons in a simulation in which HVC activity has been slowed. Now the third-to-last cluster of each syllable network is the last to exhibit spiking, as a result of the suppressive effect.

Supplementary References

Leonardo A, Fee MS. Ensemble coding of vocal control in birdsong. *J Neurosci.* 25: 652-61, 2005.