Chapter 3: Regulation of vocal exploration during continuous actions

Most of the data used in this chapter was contributed by Dina Lipkind.

3.1 Background and Rationale

In Chapter 2 we show that exploratory variability can be adaptive such that a bird can confine the exploration only to the syllable that needs to improve most. The syllable that is already mastered does not become more variable even when the new variable syllable is added. We concluded that this can allow for a piecemeal learning of the song.

But what are the smallest units and the natural time scales of vocal exploration? One possibility is that these are the song syllables, which are somewhat discrete units of song production. Our null hypothesis is that vocal exploration is globally applied across the entire syllable. We will test if similar piecemeal learning strategy can take place in smaller units during the continuous action of the syllable. To answer this question we examined song learning at the sub-syllabic level, using song development data from six birds trained with the altered-target training procedure (as described in Chapter 2) and two birds trained with a single song model.

While quantifying the variability across renditions of whole syllables was relatively straight forward, getting such measures of variability of structure within a syllable was difficult because intra-syllabic structure (notes) are not easy to segment, and are not developmentally stable.

We therefore had to develop methods for robustly identifying significant time events during the continuous syllable, and for measuring the exploratory variability of those *intra-syllabic events*. For this reason Chapter 3 is dedicated to presenting these new methods and their application.

3.2 Methods

3.2.1 Training regimes and animals

The development of inta-syllabic events was measured in eight birds: Six of those birds were trained with alternate training procedure (see Methods in Chapter 2 for description of the procedure and Fig. 2.2 *B* for the song model), and in each of those birds we analyzed the first syllable (A) they were trained with. Syllable A has complex structure and having six repeats of the learning of the same syllable allowed us to compare the learning of specific intra-syllabic song elements across birds. To make sure that our findings are not limited to this training regime or to the particular choice of syllable, we analyzed another syllable from two birds that were trained with a single song model (Fig. 1.1). For the analysis of intra-syllabic events we used the one complex syllable from those two birds. All the findings described in this dissertation apply to both groups, and since we have seen no apparent differences in results, we pooled the data and present those eight birds as a single group.

3.2.2 Syllables selected and analytic approach

For each bird, we selected for analysis <u>one complex syllable</u> with at least three distinct song elements (but typically 5-6 song elements) at the endpoint of song development and tracked occurrences of local minima in Wiener entropy within each syllable across development.

In previous studies (Du and Troyer, 2006), complex song syllables were often segmented to smaller units called notes. Although segmenting syllables to notes is possible in adult zebra finches, we find this task much more difficult during song development. A somewhat easier task is of detecting significant time events within a syllable. The advantage of this approach is that it does not require detection of boundaries (which are often blur in developing songs) but only center of events. Further, if, due to noise, our probability to detect such events is not particularly high, say 50%, a large enough sample of syllables should allow us to detect those events quite easily. In contrast, trying to segment a syllable to notes with 50% noise level in detecting boundaries would have resulted in highly variable segments.

One of the most robust features of the developing song is Wiener entropy, and its minima indicate a local peak in tonality (highest local concentration of spectral energy). To detect significant time events within syllables, we measured Wiener entropy minima for all occurrences of a syllable type for each day of development (roughly 600-10,000 renditions per day, per bird).

3.2.3 Smoothing procedure

For each rendition of the syllable type we smoothed the Wiener entropy time course using 30 ms Hanning window using the Matlab 7 *hann* function and then automatically detected the local minima in the smoothed curve (usually 3-6 per syllable). These local minima often consistently repeated across syllabic renditions to form clearly visible clusters (see Figure 3.3) that we call “intra syllabic events” (see Fig. 3.3). The 30 ms size of the Hanning window was determined empirically in a subset of our data to optimize the detection of distinct intra-syllabic events. Experimenting with different windows sizes, we found no or little impact of the window size on the qualitative results: the same clusters (intra-syllabic events) were identified over a range of different window sizes. Further, in the vast majority of the cases we were able to trace intra syllabic events over prolonged developmental epochs.

3.2.4 Time histograms of significant events

In order to identify intra-syllabic significant events (consistently identifiable events of Wiener entropy minima within a syllable) we aligned samples of 100 syllables of the same type and computed the histograms of time positions of Wiener entropy minima. Those histograms had sharp peaks that could be easily detected at the time-positions of highest density, which we call intra-syllabic events. The minima of the histograms were used to delineate the outer limits of each intra-syllabic event. These limits (separations between different intra-syllabic events) were set 5 milliseconds in each direction from the histogram minima.

3.2.5 Detecting song elements within a syllable

Much of our analysis was based on detecting robust song elements within syllables during song development. As mentioned above, our goal is to analyze the role of exploratory variability within a syllable. We encountered two problems arising from measuring the variability across renditions of continuous actions: alignment and segmentation. For example, to estimate the variability, one could align renditions of syllables according to the points where a particular acoustic feature crosses a threshold, as shown in Figure 3.1*A*.

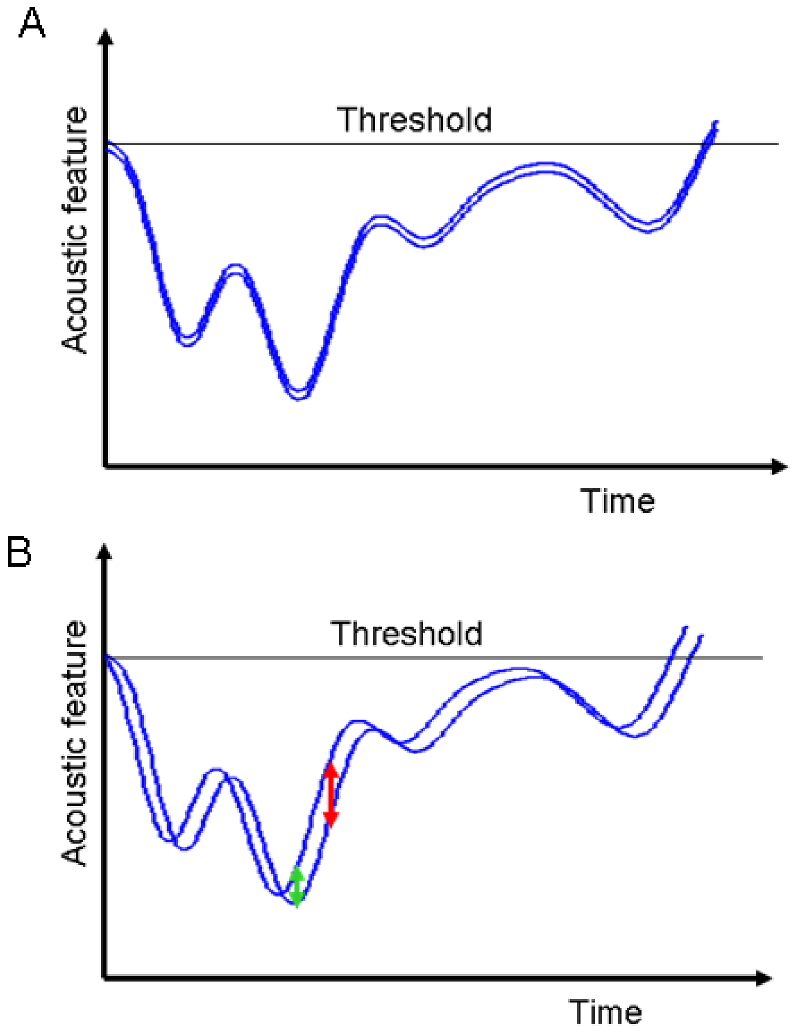

Figure 3.1 Misalignment of syllables can result in biased measurements of variability across renditions such that the slope and variability estimates will be positively correlated.   See text.

Now one could measure variability across renditions as standard deviation, for each millisecond of the syllable. With perfect alignments this simple method should produce reliable measures of variability at each time point (millisecond) within a syllable. But the problem appears when there are slight misalignments between syllabic renditions. Note that such “misalignments” can result from bird’s own temporal imprecision. The problem is illustrated in Figure 3.1*B*.

Note in Fig. 3.1*B*, that when the slope of syllabic features is low, the measured difference between the two renditions (and thus SD) is also low (green arrow) while in time-positions located at high slopes, the measured difference between renditions is relatively high (red arrows). Therefore our measure of variability between syllabic renditions will trivially depend on the location within a syllable. Another way to see the problem with this method is this: we are comparing the values of acoustic features between renditions of syllables at a particular *time* within a syllable. For example, we compare what a bird did at millisecond 20, in one syllabic rendition, to what he did at millisecond 20, in another rendition. But millisecond 20 does not necessarily correspond to a particular (consistently identifiable) acoustic event within a syllable.

In order to circumvent this problem of comparing syllabic renditions at each millisecond, we identified events within a syllable that were consistently present in each rendition.

We found, empirically, that Wiener entropy is the acoustic feature with which we could identify such consistently present events, early in development after the first appearance of the syllable (we will discuss the time of appearance of these events later). We experimented with using other acoustic features (amplitude, mean frequency and pitch) for the purpose of identifying significant time events (traceable across development) but Wiener entropy produced the best results. We therefore used Wiener entropy as a master feature to identify time-positions of the intra-syllabic events. From these time-positions we extracted the values of other acoustic features as well.

Wiener entropy is a measure of spectral noise (it estimates the width of the power spectrum). Instead of linear scale (zero entropy = order, one = disorder) we use a logarithmic scale to increase the dynamic range, and zero (log(1)) stands for disorder, i.e., white noise, and negative values for order (harmonic stacks and pure tones, Tchernichovski et al., 2000). A Wiener entropy local minimum represents a moment within the syllable where the concentration of spectral energy reaches a local peak – this corresponds to moments where harmonic stacks or pure tone notes are most clearly defined, as shown below (Fig. 3.2):

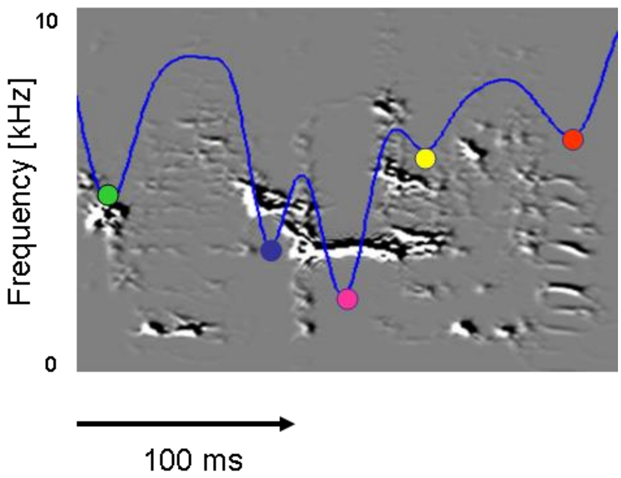

Figure 3.2 Detection of significant intra-syllabic events. Wiener entropy was used as the acoustic feature in order to identify significant intra-syllabic events. The blue trace, overlaying the spectrogram of a syllable, is the smoothened Wiener entropy. The local minima were selected as significant intra-syllabic events (colored circles) if their time positions were consistent across syllables. Note that the intra-syllabic events thus identified capture the points of high local spectral density (or harmonic sounds), with the middle event (pink circle) as the global Wiener entropy minimum.

In Fig 3.2 above note five detected events within a syllable indicated by colored circles. The blue trace represents the smoothened Wiener entropy (see Methods in this chapter for smoothing procedure). If these events consistently repeat across syllabic renditions, we call them significant intra-syllabic events. In the next sections we show how these intra-syllabic events can be tracked.

3.2.6 Tracking intra-syllabic events

In previous section we showed how we detected significant events within a syllable that could be used to compute variability across syllabic renditions, circumventing the alignment problem. For these events to be useful they have to be present consistently across syllabic renditions, e.g. we should be able to track them over the time of syllabic development.

Figure 3.3 demonstrates the identification of intra-syllabic events as consistently occurring events within a syllable.

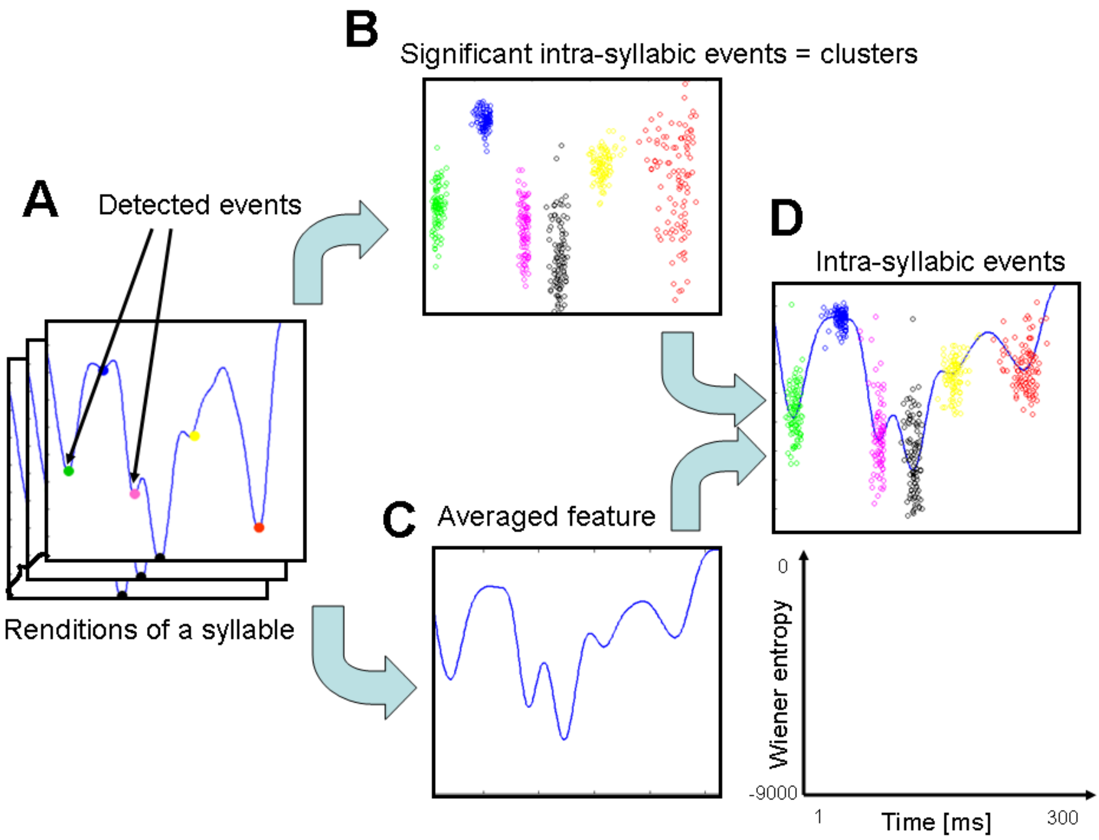

Figure 3.3 Procedure for identifying of intra-syllabic events. ***A***: smoothened Wiener entropy (blue trace) of one syllabic rendition. Six significant events (Wiener entropy minima) were identified: colored circles. ***B***: Clusters of significant events form after several syllabic renditions. These clusters we call "intra-syllabic events". ***C***: Averaged (across renditions) shape of Wiener entropy. ***D***: Averaged Wiener entropy and intra-syllabic events overlaid.

First Wiener entropy local minima were detected, as described in the previous section (Fig. 3.3 *A*). After several renditions of a syllable, these local minima were detected as clusters (Fig. 3.3 *B*). Note that the clusters are well separated from each other in time, so a time histogram could be used to identify them automatically. We called thus identified clusters *intra-syllabic events*.

Some intra-syllabic events are more "subtle" then others, meaning that they do not necessarily appear in each rendition of a syllable. For example, in Figure 3.3*A* the blue and yellow minima in Wiener entropy detected are less pronounced than other minima and are not detected in every rendition of the syllable. As a result the total number of detected minima over a day varies among different intra-syllabic events. For this reason we computed all statistics of intra-syllabic events using a fixed sample size of each intra-syllabic event (i.e., using the same running widow size as presented in section 3.3.2.). Since we record and analyze the entire song production of each bird, keeping sample size equal across events has only a negligible effect of having data epochs with slightly different boundaries for the analysis of different intra-syllabic events.

In summary, intra-syllabic events are Wiener entropy minima that clearly and consistently appeared in a recognizable timing within a syllable.

3.3 Results

3.3.1 Tracking of intra-syllabic events across syllable development

Figure 3.4 demonstrates the tracking of intra-syllabic events across syllabic development in one bird.

On the third day after the onset of training, a syllable first appeared as a district type (identifiable cluster in feature space such as in Fig. 3.4*C*). Three to seven days later, events appeared that seem to fall into clusters as described above (Fig. 3.4*B*).

Figure 3.4*A* shows the approximate location of the intra-syllabic events on the sonograms. Notice that the almost pure tone event (black color) appeared after much later than other events.

This pattern of intra-syllabic event appearance was observed in all eight birds: syllable types appeared 3.2±2.1 days (means and SD) after the onset of training while intra-syllabic events appeared 7.6±3.1 days after the training onset. This delay suggests a hierarchical process, where coarse structures (at the syllabic level) consolidate prior to the fine structures (at the sub-syllabic level).

In practice, once an *intra-syllabic event* could be recognized, we were usually able to track it continuously and automatically thereafter (using time histograms as described above). In some cases such as bifurcations we had to identify clusters by visual inspection.

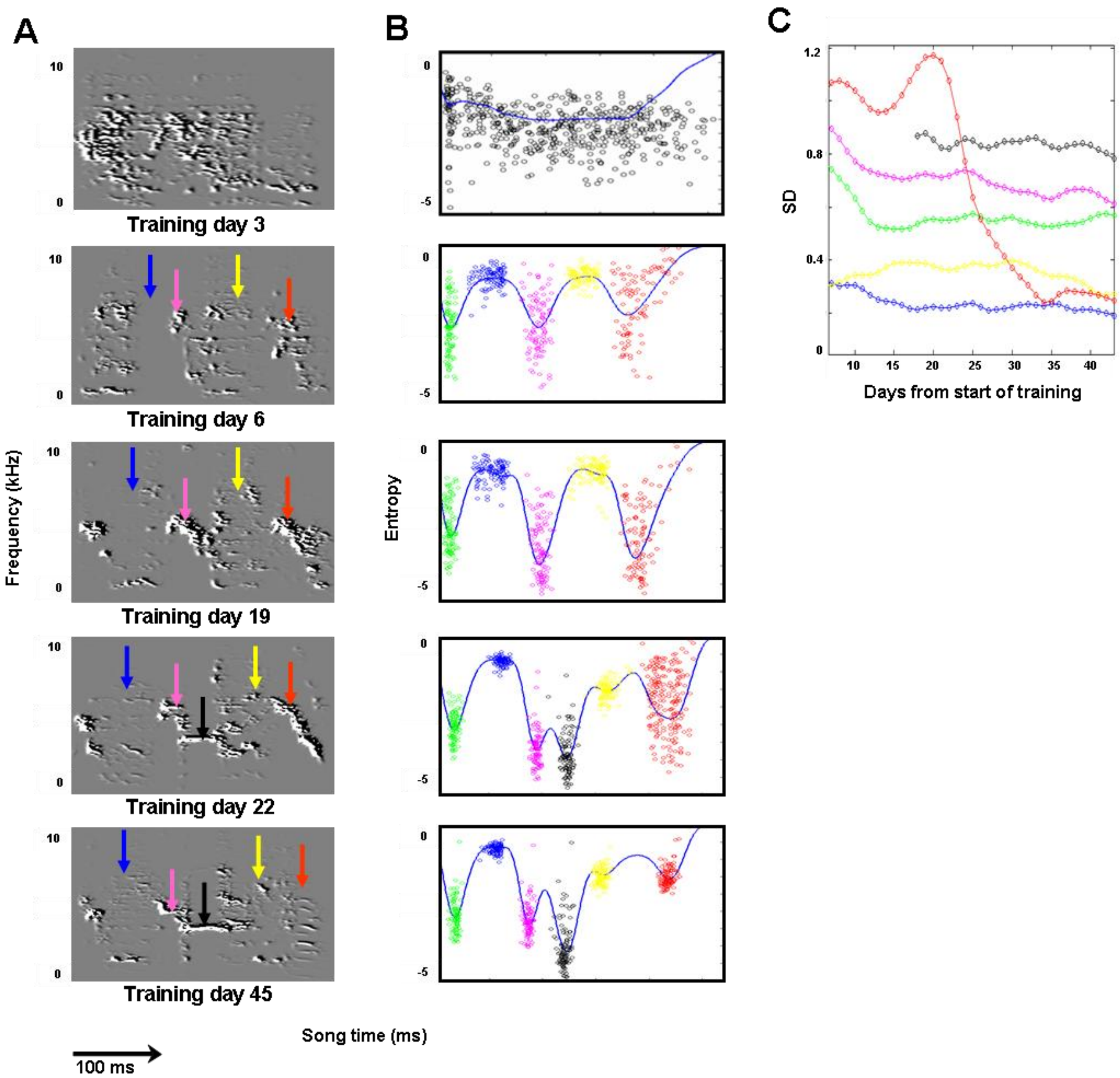

Figure 3.4 Time courses of changes in variability within a syllable. ***A,*** Sonogram of a developing syllable from its first appearance (day 3) until its consolidation (day 45). Arrows point to approximate locations of the intra-syllabic events identified. Note that on Day 3 we could not yet identify any intra-syllabic events. ***B,*** Identified intra-syllabic events can be traced across trials. The blue curve represents mean Wiener entropy traces averaged across 50 syllables. Overall, the intra-syllabic events become less variable as the clusters get smaller. (The variability of the red cluster, however, seems to increase on day 22.) Note that at day 3 we can identify syllabic types but Wiener entropy minima (black circles) do not form intra-syllabic events (clusters) until day 6. ***C,*** Developmental time courses of variability ($SD_{intrasyll}$) for all identified intra-syllabic events. Variability decreases across all intra-syllabic events but this decrease is asynchronous.

variability of each intra-syllabic event, independently of alignment.

3.3.2 Statistics of intra-syllabic events

In the previous section we show how intra-syllabic events can be identified tracked independently of the alignment. Because we analyzed complex syllables, with at least three intra-syllabic events (typically 4-5 events), this method also allows us to segment the syllables. This allows us to compute the statistics for different parts of the syllable across renditions.

The two main statistics that we used were mean and standard deviation. The former was used to determine the feature magnitude of a part of a syllable and the latter was used as a measure of exploratory variability for each part. As noted in previous sections Wiener entropy was the main acoustic feature with which intra-syllabic events were identified. Values of other acoustic features were then identified at the time-positions of these intra-syllabic events.

Mean values and standard deviations across 50 renditions were computed for all features of intra-syllabic events. We will refer to the variability of intra-syllabic events as $SD_{intra\text{-}syll}$. In order to compute $SD_{intra\text{-}syll}$ we first removed the trend (the shift of the mean) which results from the natural progression of learning. This procedure produced a few hundreds of mean and variability estimates per day. To obtain daily estimates of variability for each day we pooled the standard deviations, computed as described above, for whole-syllable features. To avoid the strong oscillations in song structure during mornings (Derégnaucourt et al., 2005), we pooled $SD_{intra\text{-}syll}$ values from the latest third of daily renditions only (which corresponded approximately to evening singing) to obtain stable daily estimates. We will discuss these daily oscillations in Chapter 4.

Figure 3.4*C* is showing time courses of daily values of variability ($SD_{intra\text{-}syll}$) for six intra-syllabic events. Note that the variability of all six events is decreasing but this decrement is not synchronous: the variability of the red intra-syllabic event, for example, increased approximately from day 15 to day 25 after the start of training and decreased thereafter (this is also evident from the size of red cluster in Fig.3.4*B*), while the variability of most other intra-syllabic events in Fig. 3.4*C* decreased before that period.

We wondered if such asynchronous decrements of variability could indicate the presence of adaptive exploration (regulation of exploratory variability) within a syllable. Because we did not have the same advantage of alternate training procedure to manipulate learning of intra-syllabic events (as we did with whole syllables, see Chapter 2) we had to look at how close intra-syllabic events were to the target (the song model). We turn to this issue in Chapter 4.