Chapter 2: Vocal exploration is locally regulated during song learning

Data collection and experimental design used in this chapter was contributed by Dina Lipkind.

2. 1 Background and Rationale

Exploratory variability is essential for sensory-motor learning. We know that variability across renditions of syllables decreases over song development as the structure of the song becomes increasingly similar to the model (e.g. the acoustic error decreases). It has been shown that this exploratory variability is necessary for song learning (Brainard and Doupe, 2000; Olveczky et al., 2005). But it is not known how and at what time scales can variability be regulated. It could be, for example, that acoustic error is estimated globally and as it decreases, the exploratory variability of the entire song decreases with it (global consolidation). On the other hand, observing local consolidation (e.g. where variability decreases locally in parts of the song where the target is approached) would imply that acoustic error is estimated locally.

In order to disambiguate between these possibilities we manipulated song learning in zebra finches to experimentally control the requirements for vocal exploration in different parts of their song. We first trained birds to perform a one-syllable song, and once they mastered it we added a new syllable to the song model. If the birds can regulate the exploratory variability only globally, we predicted that the addition of a new (and unstructured syllable) to the song would result in the increment of variability gain across the entire song bout, including the already mastered syllable. On the other hand, if variability can be regulated locally we predicted that exploration would be confined only to the newly added syllable, which would also imply that acoustic error can be estimated locally.

2.2 Methods

2.2.1 Animal care

All experiments were conducted in agreement with USNIH guidelines and were reviewed and approved by the Institutional Animal Care and Use Committee of City College of New York, City University of New York.

2.2.2 Training procedure

Birds were bred in family cages. Fathers were removed when clutch mates were 7-8 days old or less, and thereafter birds were raised by their mothers and were not exposed to songs. On day 30-32 post-hatch, male birds were individually isolated in sound-attenuation chambers. All birds were tutored with operant song playbacks from day 43 to 90 days post hatch, as described in Tchernichovski et al., 1999. Eight birds used in the study presented in this chapter were trained using an altered-target training procedure as described below.

2.2.3 Song recording and analysis

We audio recorded (16 bit, 44.1kHz) each bird continuously from day 32 to day 90 post-hatch using Sound Analysis Pro 1.4 (Tchernichovski et al., 2000). Recording epochs containing songs were automatically identified and saved, and song features (amplitude, pitch, Wiener entropy, etc.) were computed as in Tchernichovski and Mitra (2004). Multitaper spectral analysis (Tchernichovski et al., 2000) was performed with time windows of 10 ms, advancing in steps of 1 ms such that song features were computed for every millisecond. Syllable boundaries were identified using a stationary threshold of sound amplitude (segmentation).

2.2.4 Altered-target training procedure

Zebra finches were trained sequentially with two song models (“source” and “target”) as in Lipkind and Tchernichovski (2011, Fig 2.2 *A*, *B*). Source and target song models for training were composed from natural syllables. Twenty-eight birds were trained with playbacks of the source song, composed of a single syllable (AAA…) from day 43 post-hatch, and songs were analyzed daily to determine if the source model was imitated. For birds that learned the source before day 63 (n=15), we switched their training to playbacks of the target song (ABAB…). Eight birds that succeeded to learn the target song were selected for analysis. In six of these birds, the novel syllable B was a harmonic stack (see Fig. 2.2 *A*) and in the remaining two birds, a broadband, highly modulated syllable.

2.2.5 Cluster analysis

After the segmentation of song bouts mean acoustic features of the segments were computed. In the acoustic space these segments tend to fall into clusters as shown in the Figure 2.1.

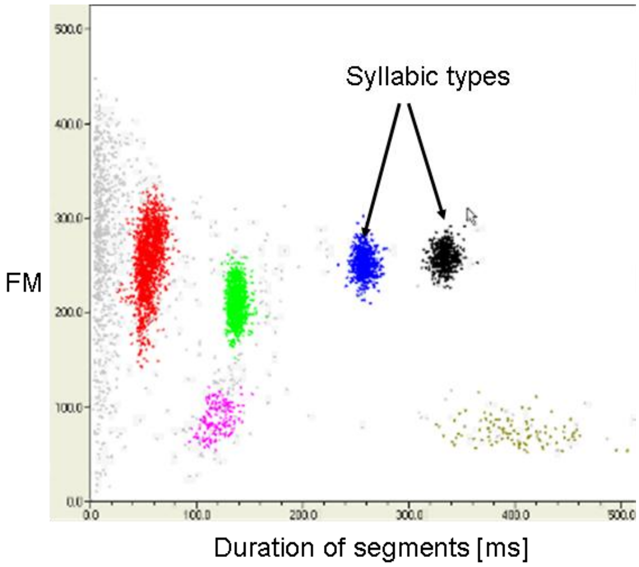

Figure 2.1 Hierarchical nearest neighbor cluster analysis methods was used to identify syllabic types. Song segments typically form clusters in the acoustic feature space. Only two features (FM and duration) are shown in the scatter plot but the clustering algorithm operates on multiple features. Each point represents mean song feature values of a segment. The colored points belong to identified clusters; gray points are not clustered (residuals).

We used a hierarchical nearest neighbor clustering algorithm, using Sound Analysis Pro 1.4 (Tchernichovski et al., 2000) for the cluster analysis. In the figure above identified clusters are colored. Although only two song features are shown (FM and duration) cluster analysis is performed across multiple features.

2.3 Results

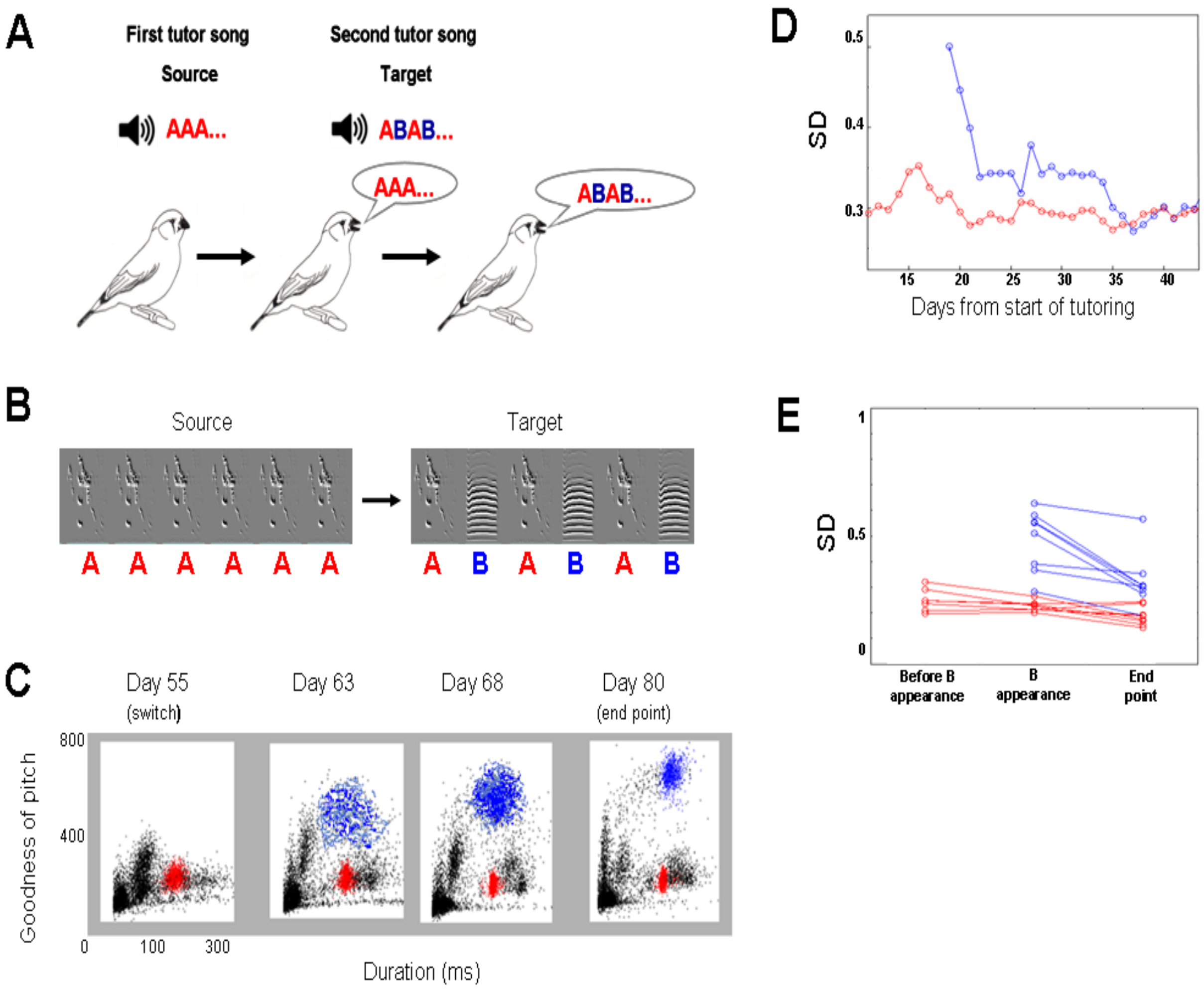

Figure 2.2 Vocal exploration is confined to newly added syllables. ***A,*** The AAAA→ABAB altered target training procedure. ***B,*** Spectral derivatives (sonograms) showing the source and target song models. ***C,*** Scatter plots of syllable features (goodness of pitch versus duration). The red cluster corresponds to syllable A and the blue cluster to syllable B. ***D,*** Variability (SD_{syll}) of syllables A and B across development. This example from one bird shows variability pooled over syllable features. ***E,*** Same as in ***D***, but across birds (n=8), 3 days prior, 3 days after the appearance of cluster B, and last 3 days before the end-point. Note, the variability of A cluster does not change after the appearance of B (p>0.98), while variability of B drops significantly (p<0.003, single-tailed *t-test*).

2.3.1 Inclusion of a new syllable to the song bout

We first manipulated song learning so that only one part of the song would require vocal exploration. We used an altered-target training approach (see 2.2.4 and Lipkind and Tchernichovski, 2011), training juvenile zebra finches (day 43 post-hatch) first with a source song model (AAAA) consisting of a single syllable and then altering the training to a target song model (ABAB), which included an additional new syllable B (Fig. 2.2 *A*,*B* and Methods). Most birds succeeded in inserting the new syllable into their song bouts and started to produce the target song.

2.3.2 Recording and analysis of the syllables

We recorded and analyzed the entire vocal output of each bird during the transition (AAAA→ABAB), automatically segmenting the songs into syllables. The structure of each syllable was summarized by four features: duration, mean Wiener entropy, mean frequency and goodness of pitch (Tchernichovski et al., 2000). We then performed cluster analysis of syllable features in order to identify the A and B syllable types (see Methods).

Figure 2.2*C* presents an example of one bird, showing scatter plots of two syllable features (duration versus mean goodness of pitch) in different stages of song learning. By the time we altered the tutoring, the cluster that corresponded to syllable A (red) was already small and dense. In contrast, the cluster that corresponded to the newly learned syllable B (blue) was initially much larger and highly scattered. The variability across renditions (SD_{syll}) of syllable B (blue cluster in Fig. 2.2*C*) then gradually decreased until it became similar to the variability of syllable A (Fig 2.2*D*). Interestingly, there was no apparent increase in the variability of syllable A when syllable B appeared (Fig 2.2*C*), indicating that when the bird sang ABAB it rapidly altered between performing a highly stereotyped and a highly variable syllable.

2.3.3 Exploratory variability is confined to the newly included syllable across birds

To test across birds (*n*=8) if variability of syllable A was affected by the appearance of new, highly variable, syllable B, we calculated the variability (SD_{syll}) of both clusters during three time periods: just prior to the appearance of B (days -3 to -1), just following the appearance of B (days 1-3) and at the end-point (days 90-93 post-hatch). As shown in Fig. 2.2 *E*, the variability across renditions (SD_{syll}) of syllable A (red traces) did not increase during the three-day period after the highly variable cluster B (blue traces) emerged (p=0.98, paired *t-test*) and did not differ from SD_{syll} at the end of development (p>0.98, paired *t-test*). At the same time, SD_{syll} for syllable B decreased strongly (p<0.003, paired *t-test*). This result indicates that exploratory variability was confined to the newly added syllable across birds.

We also compared variability during morning singing to afternoon variability. As will be shown in Chapter 4, morning variability of sub-syllabic structure tends to be higher than afternoon variability, suggesting that most exploration takes place during morning singing, which could be related to the “morning effect” (Derégnaucourt et al, 2005). Figure 2*A* shows examples from three birds. As in Fig. 2.2*D*, variability decreases across development in syllables A and B. Note that in some cases of A or B syllables the morning variability (traces marked with “*”) is consistently higher than afternoon variability (traces marked circles). This however was not the case across all birds in syllable B, while in syllable A the median variability was significantly higher in the morning than in the afternoon (Fig. 2.3*B*).

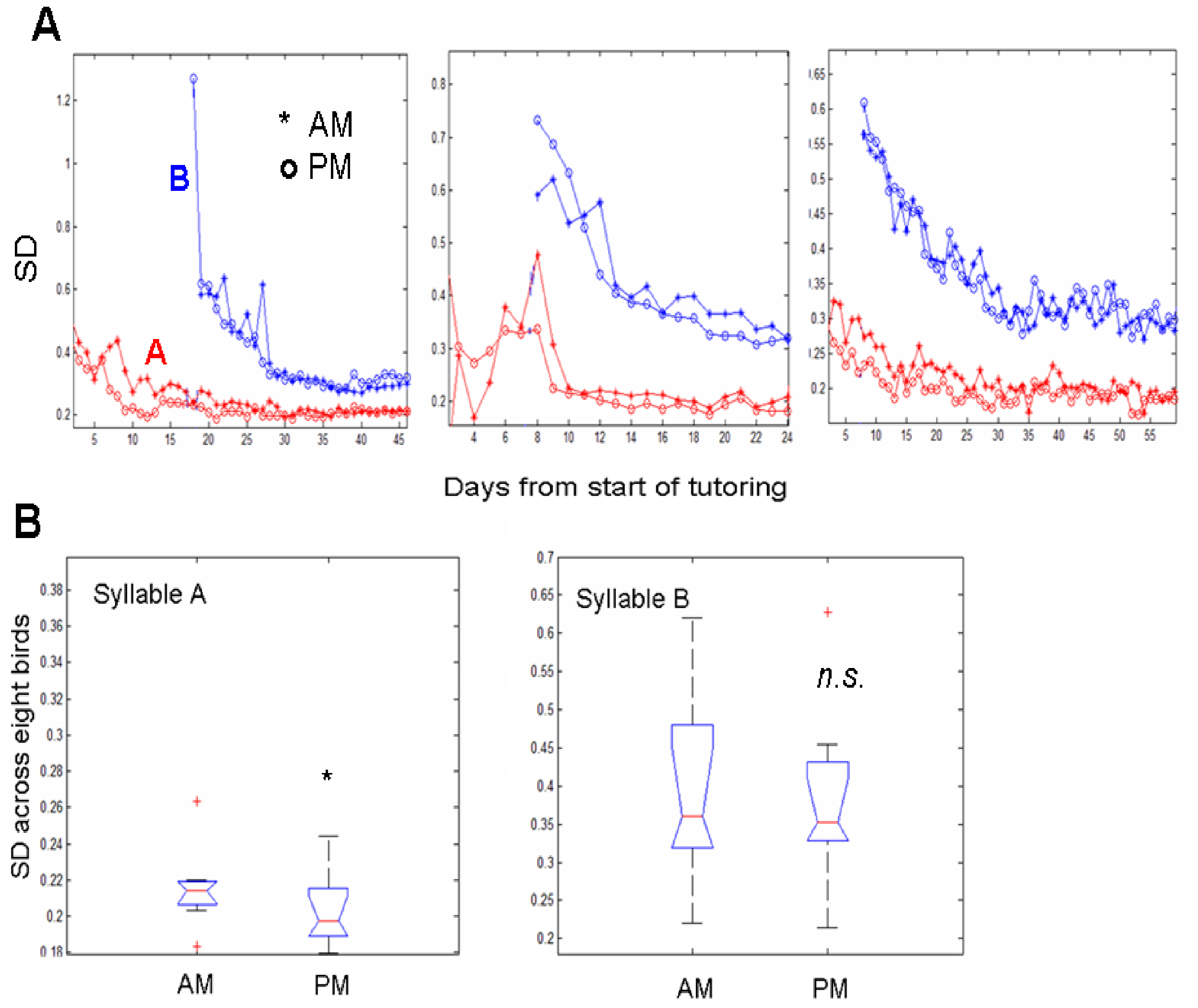

Figure 2.3 In some birds morning variability of syllables is higher than afternoon variability. Morning variability (SD across features) is shown with "*" symbol in ***A*** and afternoon variability with circles. Variability of syllable A is shown with red trances and variability of be with blue traces. Three examples of birds trained with sequential training procedure are shown. We did not find significant difference between morning variability and afternoon variability across all eight birds in syllable B. While syllable A is less variable than syllable B, its morning variability was higher than afternoon variability, across all eight birds, as shown in ***B*** (p<0.017, single-tailed *t-test*). We did not detect significant difference between morning and afternoon in syllable B. However, in the cases where there was a difference it was always in the direction of higher morning variability (as shown in

When new syllables are incorporated into a song bout, the exploratory variability required to learn the new model (target) is confined to them alone. This finding supports the hypothesis that, when faced with the conflict between consolidation and exploration during learning of continuous actions, the bird segments the singing action and adjusts vocal exploration level locally. We can conclude that learning trajectories of syllable A and syllable B are separate as they consolidate independently one from another. That is, the song can be learned in a piecemeal fashion.

What is the source of exploratory variability? As noted in Chapter 1, the temporary silencing of AFP ("learning pathway") by injecting TTX into LMAN promptly reduce variability between syllabic renditions to that of a fully stereotyped adult song (Olveczky et al, 2005). Similar results (elimination of variability between syllabic renditions) were obtained even in syllables that were highly variable (Andalman and Fee, 2009). It would be interesting in this context to perform the sequential training procedure in birds where LMAN could be inactivated. Given the above mentioned results the prediction would be that both, syllable A and the new syllable B would become stereotyped. A similar experiment could be carried out using directed vs. undirected singing. As noted in Chapter 1, during directed singing the syllables become much less variable (Hessler and Doupe 1999 a,b) and it has been shown that this difference in variability can be used to study the role of AFP in learning (Kao, 2005).

We took advantage of the "morning effect" to see if there is a difference in the diurnal variability oscillations associated with the effect, between syllables A and B. We predicted that if, as we will suggest in Chapter 4, most exploration takes place in the morning, the oscillations should be greater for the newly added syllable B. We observed a small, but significant difference between morning and afternoon variability in syllable A ($p<0.017$, single-tailed *t-test*) However, we did not observe any significant difference between morning and afternoon variability in syllable B It should be noted that in all birds where the difference in morning vs. afternoon variability was significant (the “morning “effect” was there) it was always the morning variability that was higher. However, as we shall see in Chapter 4, this is a misleading finding since variability at the syllable level is often a compound outcome of the means and variances of intra-syllabic events.

It is possible that while an overwhelming portion of variability could be explained by the active injection of noise to the motor pathway by the AFP (Olveczky et al, 2005) there is a second source of variability caused by differential sensitivity of the motor pathway to the AFP input. According to this view the morning variability could be higher because the motor pathway becomes more responsive (sensitive) to the noise injected by the AFP. This explanation would also be consistent with the view that morning song is more plastic than the afternoon song (Derégnaucourt et al, 2005).