

## **Chapter 5: Song learning at the articulatory level: the development of respiratory pressure patterns**

Data used in this chapter were collected by Primoz Ravbar, Franz Goller and Brent Cooper at the Vicario lab & Goller lab. EMG data were collected by Franz Goller at the Goller lab. Pratik Shukla also contributed to data collection and provided major help with chronic recording maintenance. Xiao Jianqiang improved the recording techniques used in this chapter.

### ***5.1 Background and Rationale***

In previous chapters we explored the development of a syllable by analyzing its acoustic structure. We showed that birds can regulate variability across syllabic renditions such that parts of the song that need to change more (because they are further from the target) receive higher exploratory variability. The techniques were developed to track various parts of a syllable from their developmental origin onwards. We showed that the scope at which vocal exploration takes place could be quite narrow. We know (Kao et al., 2004) that perturbation of syllabic structure by the “learning pathway” – AFP – can be very precise, both in time and in the specificity of acoustic features affected.

But while we know that birds can manipulate song at short time intervals it remains less clear how the ability to do so develops. How does the *skill* to perform, modify and explore song elements develop? (We define *skill* as the ability to manipulate the song at short time scales.) It would be worthwhile to know, for example, which sets of

muscles are involved in song production and how does the activity of those sets of muscle develop with song learning. Unfortunately analyzing the acoustic structure alone can tell us little about muscle activity.

This was the main motivation for identifying a physiological variable that can be followed developmentally to study basic processes that underline the emergence of motor skill. This variable is *respiratory pressure* and can be directly related to activity in a particular set of muscles - those controlling respiration (Suthers 1999; Goller 2002). In this chapter we will discuss the methods used to record and analyze the respiratory pressure patterns and see how these methods can be used to measure development of motor skill. But first let us examine the background of song production.

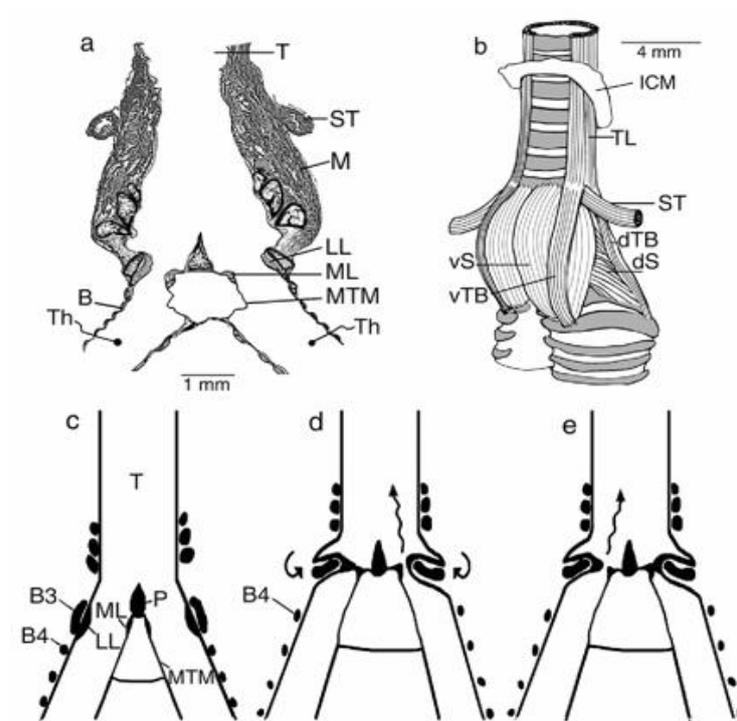
### **5.1.1 Major components required for song production**

The song is produced by coordinated activity of the vocal apparatus, which include three major components:

1. *The Syrinx*
2. *The respiratory system*
3. *The beak and the upper vocal tract*

These components have to be well orchestrated during singing. Now we will present the functional role of each component.

## The Syrinx



**Figure 5.1 Schematics of the syrinx.** See text.

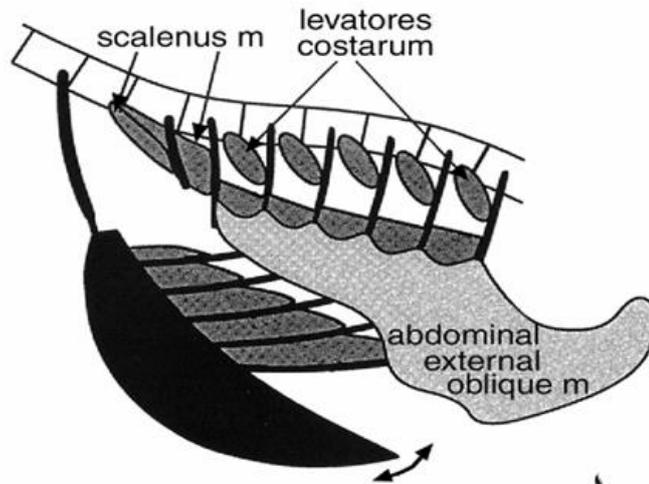
*Taken from Suthers et al., 1999*

Syrinx, the vocal organ of songbirds, is located in the interclavicular air sac between the two bronchi and the trachea (in zebra finches). It is a complex organ composed of modified cartilage and six muscles at each side. Those muscles are responsible for adductions (closing) and abductions (opening) of labia in each bronchus. The vibrations of the labia result in sound. Labia on both sides of the syrinx are fully abducted during silent respiration (Fig. 5.1C). During singing labia adduct (which increases the resistance in the bronchus and creates the right conditions for sound producing oscillations (Fig. 5.1B). Also, when labia are completely adducted they can act as vents, closing the ipsilateral bronchi (Fig. 5.1E). Extensive EMG recordings during singing revealed functions of different sets of muscles in the syrinx (Goller and

Suthers, 1996). The *dorsal muscles* (dS and dTB, Fig B) are involved in closing (adduction) the bronchus, decreasing the air flow through the syrinx (Suthers et al, 1999). When these muscles contract, they adduct the labia by moving the cartilage of the bronchus. *Ventral muscles* (vS, vTB), on the other hand, control the frequency of sound by manipulating the tension of labia (Goller and Suthers, 1996).

Interestingly, the two sides of the syrinx can (in some species) produce sound quite independently from each other (Suthers, 1997). This “lateralization of song control” increases production capabilities of some species of songbirds: the two sides of the syrinx are often specialized for different frequency ranges. In domestic canary, for example, the right side produces high frequency parts of “FM sweeps” (these are syllables or parts of syllables where mean frequency is decreasing as to produce a “sweep” on a sonogram), and the left side produces the low frequency parts (Suthers et al, 2004). In other species, like in brown headed cowbirds, songs are produced by alternation between left and right side of the syrinx (Suthers and Roderick, 1997). In the zebra finch, however, most song syllables are produced by co-operation between the two sides of the syrinx, except from high pitch notes which are produced by the right side only (Goller and Cooper, 2004). But this lateralization of song production might also pose additional challenges of coordination; now the motor system has to coordinate not only respiratory muscle activity with the syrinx but also the two halves of the syrinx with each other.

## The respiratory system

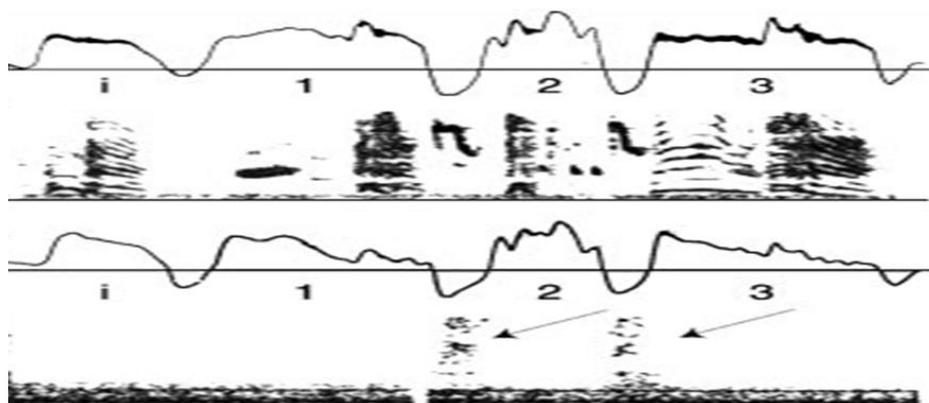


**Figure 5. 2** Abdominal muscles are involved in production of expiratory pressure. The production of inspiratory pressure is controlled by *scalenus m*, *levatores costarum* muscles, while expiratory pressure is produced by abdominal muscles. Taken from Suthers et al., 1999

The respiratory pressure is produced by a set of about a dozen muscles acting on a complex of interconnected air sacs. The activity of these muscles has to be precisely coordinated with the syringeal muscle activity. In contrast to syrinx, the lateralization of song control does not apply to the respiratory pressure (Goller and Suthers, 1999). As such, the pressure amplitude is a global articulatory state, common to both bronchi. In birds, both inspiration and expiration are active processes (Konishi, 1965). The expiration pressure is produced by the activity of abdominal muscles, while the inspiratory pressure is controlled by a set of thoracic muscles (*scalenus m*, *levatores costarum*, see Fig, 5.2 above) (Suthers et al., 1999).

How does the respiratory pressure relate to the vocal output? When the air sac pressure is measured during singing we can observe a number of interesting properties. Unlike in human speech the respiratory pressure amplitude follows closely

the acoustic structure of the song. The pressure signal is composed of repeated pulses of expiration which we will call “EPs” (expiratory pulses). EPs are correlated with the acoustic signal (Fig 5. 4). On a coarse scale the pressure correlates with syllable boundaries, but on a finer scale, it correlates with intra-syllabic structure. Note that pressure shape reflects, to a large extent, the input *to* the syrinx and not a combined effect of respiratory muscles and syringeal resistance. That is, EP shape is not significantly determined by changes in syrinx resistance (Cooper and Goller, 2004). This is shown in Figure 5. 3. Fine structure of the EPs is not affected by fixing of the syrinx such that it stays always open. Under this condition the resistance of the syrinx does not change and can therefore not contribute to the finer structure of the EPs observed. Notice that “singing” is silent under those conditions, while the pressure signal is not affected (Fig. 5. 3). This is an indication that EP shape is determined by the respiratory system alone. This fact is critical for the main aim of this chapter, namely to relate the development of the skill (ability to manipulate song at short time-scales) to the activity of muscles.

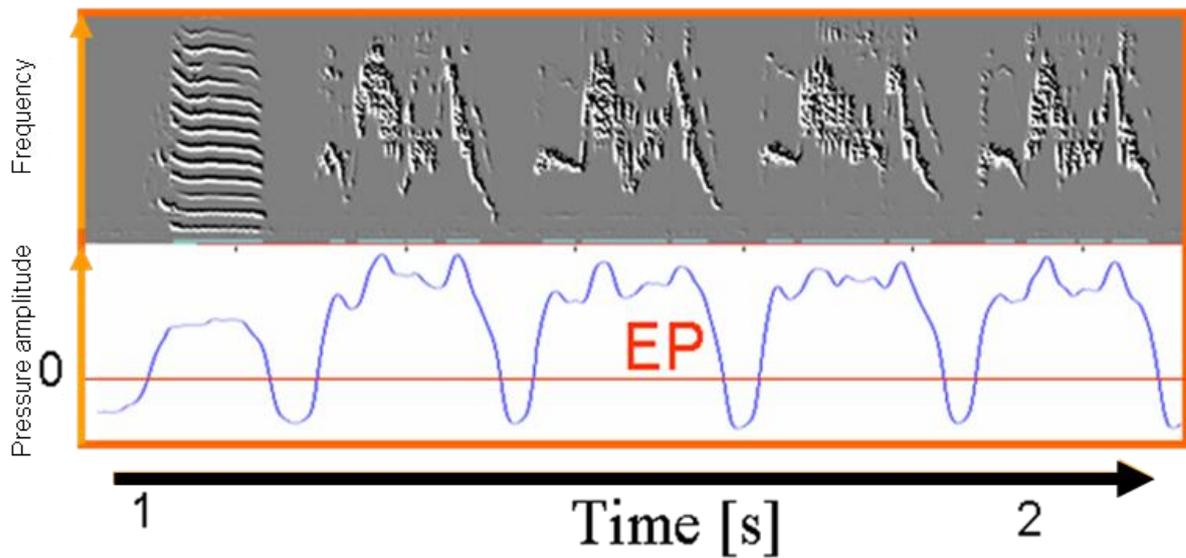


**Figure 5. 3** The fine temporal structure of respiratory pressure does not result from the activity of the syrinx. The top trace is showing respiratory pressure pattern during vocalization (sonogram bellow). After muting procedure where the air resistance was set constant the pressure pattern retained all of the fine temporal structure (lower pressure trace; the sonogram after muting bellow). *Taken from Cooper and Goller, 2004*

This highly structured pattern of respiratory activity must in turn be tightly coordinated with syringeal activity, ensuring that the syrinx receives just the right amount of pressure for producing the correct sub-syllabic acoustic structure.

Another interesting feature of respiratory pressure is that while most animals exclusively vocalize during positive respiratory pressure phase (e.g. during expiration), birds also (albeit not often) vocalize song syllables during inspiration. Acoustic notes produced during inspiration are observed more often in early, non-crystallized songs (Goller and Daley, 2001). The song syllables are separated by short, usually silent periods during which the bird is inhaling. These short inhalations are called “minibreaths” (Calder, 1970) and are much deeper than inhalations produced during quiet respiration.

These properties make pressure especially important for the temporal organization of the song. If a singing bird is interrupted by a strobe light, the song usually (but not always) gets truncated usually *between*, but sometimes also *within* a syllable (Cynx, 1990). These “continuation to completion” effect is even stronger at to the level of expiratory pulses (EPs) observed on air sac pressure (Franz and Goller, 2002). Thus, when the song is interrupted by the strobe light, it breaks down *only* between the cycles of inspiration and expiration. In a sense the respiratory pressure determines the temporal structure of song and the motor units of its production it could be referred the *articulatory* component of singing.



**Figure 5. 4** Pressure amplitude follows closely the acoustic output of song. Note that pressure follows the sub-syllabic structure of acoustic syllables (repeated units in the spectrogram above).

### The beak and the upper vocal tract

Another motor system, and a major component of song production, that has to be integrated with the syrinx and respiratory systems consists of the beak and the upper tract. It was shown that beak movements during the song are stereotyped (Goller et al., 2004). In zebra finches the beak gap is also well correlated with the peak frequency of different notes in a syllable, namely large beak gape usually results in removal of high frequencies from harmonic stacks. This may indicate that opening the beak shortens the upper vocal tract, thereby changing its filtering properties. Beak gap is also closely correlated with the amplitude of respiratory pressure. This coordination

between respiratory system and mandibular motor control is not well understood. A lot more is known about coordination between respiratory system and the syrinx.

### **5.1.2 How are respiratory muscles and syrinx muscles coordinated?**

This is a subject of current research and we will summarize just one possible mechanism by which the nervous system is coordinating respiratory activity with the activity of syringeal muscles.

The forebrain nucleus RA (*robustus archistriatalis*) sends projections to pre-motor nuclei of both parts of the vocal apparatus: syringeal and respiratory. On the syringeal part of the vocal apparatus, neurons from ventral RA ipsilaterally project onto motor neurons in the hypoglossal nucleus (nXIIIts). These motor neurons drive the activity of syringeal muscles, both ventral and dorsal. On the respiratory part of the vocal apparatus, dorsal RA projects onto RAm (nucleus retroambigualis) and RVL (ventrolateral medulla, controlling quiet respiration). RAm, in turn projects to both nXIIIts (bilaterally) and onto respiratory muscles (Wild et al, 2000). More recent findings (Kubke et al., 2009) suggest that RAm consists of two distinct neuronal populations both projecting onto nXIIIts, as well as populations projecting onto motor neurons of respiratory muscles. These various populations of RAm neurons are likely involved in coordination of the syringeal muscle activity with the activity of respiratory muscles, exciting and inhibiting the nXIIIts during different parts of respiratory cycle. It is thought that RAm may play a role in gating of the nXIIIts activity so that appropriate *phase delays* are established between syringeal and

expiratory muscles. These delays allow for enough expiratory pressure to be created (by expiratory muscles) before syringeal vocalization starts (Suthers and Margoliash, 2002; Kubke et al., 2009).

Now when we presented all the major components involved in song production let us go back to our main aim in quantifying the development of skill. But first let us present the two conditions under which we will measure skill development.

### **5.1.3 Acoustic isolates vs. tutored birds**

In order to quantify the skill development we compared birds that were tutored with a song model to those that were raised in acoustic isolation. The latter birds can hear themselves vocalize but have no memorized tutor song to copy. The rationale for this choice of experimental groups is that we would expect more skill development in birds who attempt to imitate the song model, than in acoustic isolates. That is, we would expect tutored birds to develop in the direction of being able to manipulate their song at shorter time scales. Interestingly, however, song undergoes some developmental changes even in these birds, although their song has atypical features and is not completely stereotyped (Marler, 1985).

## ***5.2 Methods and definitions***

### 5.2.1 Recording of respiratory pressure during singing

In all the birds, we used a surgically implanted scilastic tube (a cannula) to record respiratory pressure continuously for periods from a few days to a few weeks, together with the acoustic signal. Under these conditions, birds sing spontaneously.

### 5.2.2 Analysis of respiratory pressure patterns

From these data, we identified pulses of expiration on the respiratory pressure signal. These pulses usually correspond to syllables in the acoustics (Franz and Goller, 2002). There are also inspiratory pulses (negative pressure), usually associated with silences separating the syllables. Therefore when we study the activity of respiratory muscles, resulting in expiratory pressure pulses (EPs), we can infer that this activity is in fact involved in syllabic development. And, not surprisingly perhaps, as the syllables develop so do their corresponding EPs. Thus the basic units of our pressure pattern analysis are the EPs.

We then quantified the developmental changes in EP waveforms (raw data) by using the Fast Fourier Transformation (FFT), to compute a power spectrum for each EP analyzed. This method will be described in detail in this chapter, but its essential use is to quantify the addition of increasingly fine temporal structure to the pre-existing coarse structure of the EPs, over the period of song development. This process of addition of fine structure to the coarse structure we call *pressure refinement*, which is our measure of motor skill. Most importantly, our method allows us to compare pressure refinement (motor skill acquisition) between juvenile and adult birds as well as between tutored and untutored birds

We found that, although the song of untutored changes, it does not show the same refinement as in tutored controls. Thus, refinement (addition of fine temporal structure to EPs) does not occur merely as a function of vocal practice, but appears to require guidance by an acquired “template”.

### **5.2.3 Data collection**

Once the singing activity was stable we started recording the respiratory pressure from a plastic cannula inserted into the thoracic air sac. Surgery was performed under gas anesthesia (isoflurane), the skin was punched just under the rib cage and a cannula was fixed into position by surgical suture and tissue adhesive. This procedure is relatively non-invasive and the bird resumes singing 1-2 days after the surgery. The cannula was then led out to a pressure transducer incorporated in a backpack attached to the back of a bird. Pressure data was digitized together with the sound using NIDAQ card and the Sound Analysis Pro software (Tchernichovski et al., 2000) for training and recording. The pressure signal was acquired at the sampling rate of 44.1 kHz and amplified. Filters were set to record signals from DC – 3 kHz. We monitored the pressure signal during singing and quiet respiration. If the signal to noise ratio (SNR) decreased drastically we repeated the cannula implant surgery. Only the data with consistently high SNR were analyzed. This procedure allowed us to record respiratory pressure in individual birds for continuous periods lasting 4-14 days. Pressure data was analyzed using Matlab 7. Raw signal was averaged to millisecond resolution. Data with no singing activity or low signal-to-noise ratio were discarded.

## **5.2.4 Segmentation**

The pressure signal was segmented by a stationary threshold set at zero pressure (the average of silent respiration). This way EPs (expiratory pulses) crossing the threshold are isolated. We used mean amplitude and duration to segregate EPs and cluster them into different types. Only EPs belonging to the same type (cluster) were used for analysis.

## **5.2.5 Alignment and averaging**

EPs belonging to the same type were aligned by the beginning (the first point at which they crossed the stationary threshold). These EP segments were then averaged across renditions to produce the averaged shapes of EPs presented in Figure 5.11 of the Results. We tried alternative alignment methods but did not observe a significant difference in averaged shapes of EPs.

## **5.2.6 Fast Fourier Transformation (FFT)**

Any complex signal can be decomposed into a number of sine waves of different frequencies. After EPs of the same type were aligned we high-pass filtered them, using differential filter. The input signal was padded with zeros on both ends as necessary so that power spectra of the EPs could be computed on the filtered data (Matlab 7 *fft* function). Averages of 100 spectra of sequential EP renditions were taken in order to remove non-repeating frequencies. These averaged spectra were then

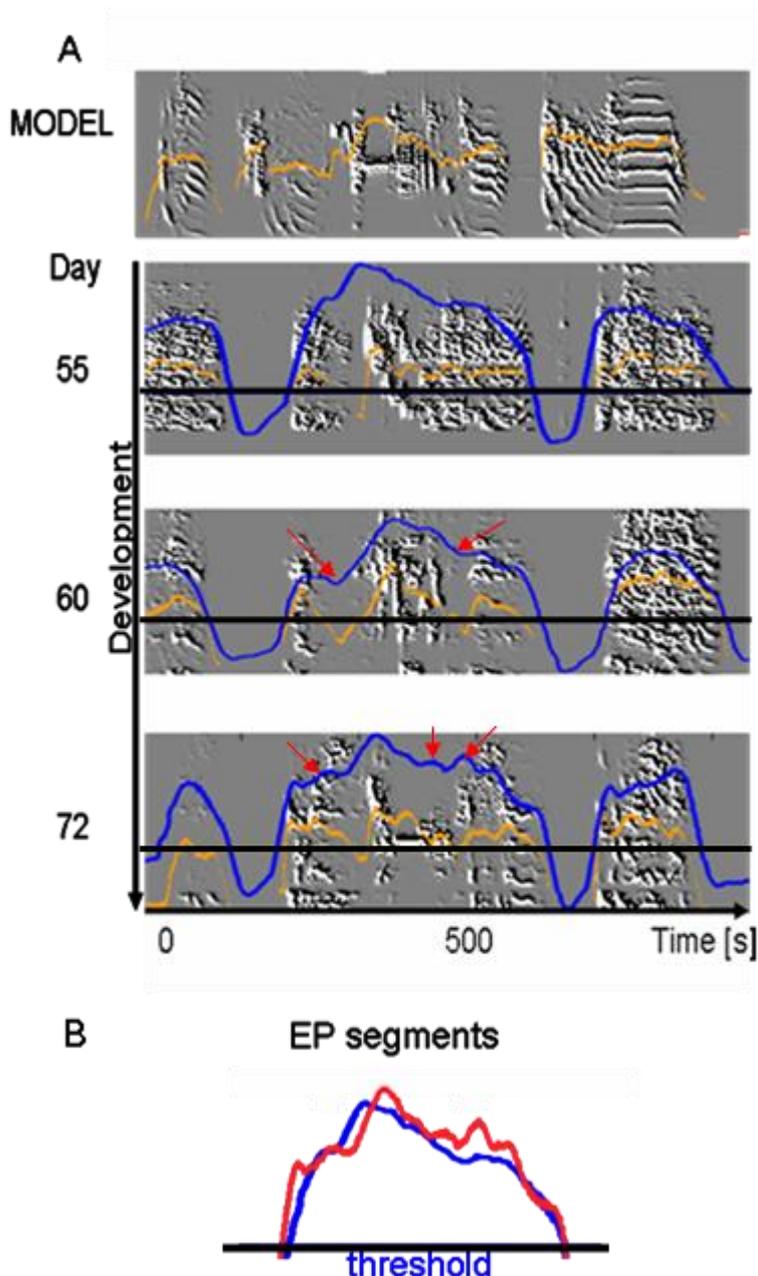
used to calculate their mean frequencies (our measure of motor skill): of the center of gravity of Fourier coefficients (power) was taken as an estimate of mean frequency.

## **5.3 Results**

### **5.3.1 Development of EP structure during song learning**

The developmental changes in a representative zebra finch song (in a bird tutored with a song model) that lead to a stable, stereotyped “crystallized” song are illustrated in Figure 5.5. This figure shows an introductory note and two different syllables that make up a song bout. As can be seen from the sonograms, the syllables become progressively more similar to the model as the song develops from Day 55 of age to Day 72 (training started on Day 43). These changes are accompanied by changes in the pressure signal (blue trace), as well as the acoustic amplitude signal (orange trace). The horizontal black line indicates zero respiratory pressure (room pressure or averaged pressure of quiet respiration). This line is used as a threshold to segment the pressure trace into positive pressure pulses (pressure above the zero-line) and negative pressure pulses (below the threshold line). We refer to the positive expiratory pressure pulses as EPs. Nearly all vocalization is produced during the EPs although a subclass of syllables can be produced on inspiration (Goller et al., 2001). Fine temporal modulations can be seen not only on the acoustic amplitude but also on the respiratory pressure, suggesting the contribution of respiratory as well as syringeal control to song production (see Discussion of this chapter).

Importantly, the modulation of the respiratory pressure signal increases over the period of development from Day 55 to Day 72, as is evident from the pressure signal (blue trace) in Figure 5.5A. The changes are seen in the superimposed EP segments in 5.5B. The red arrows in 5.5A indicate the points where new valleys and peaks have appeared.



**Figure 5. 5** Song development. *A*, sonograms (vertical axis is frequency; horizontal axis time) show song development from day 55 to day 72 after hatching. The blue trace shows the respiratory pressure and the orange trace represents sound amplitude. The horizontal black lines identify zero pressure also used as a threshold to segment expiratory pulses (EPs). The duration of these song bouts is 1.0 sec. Red arrows point to examples of added fine structure. *B*, EP segments from Day 55 (blue) and Day 72 (red) are superimposed for comparison.

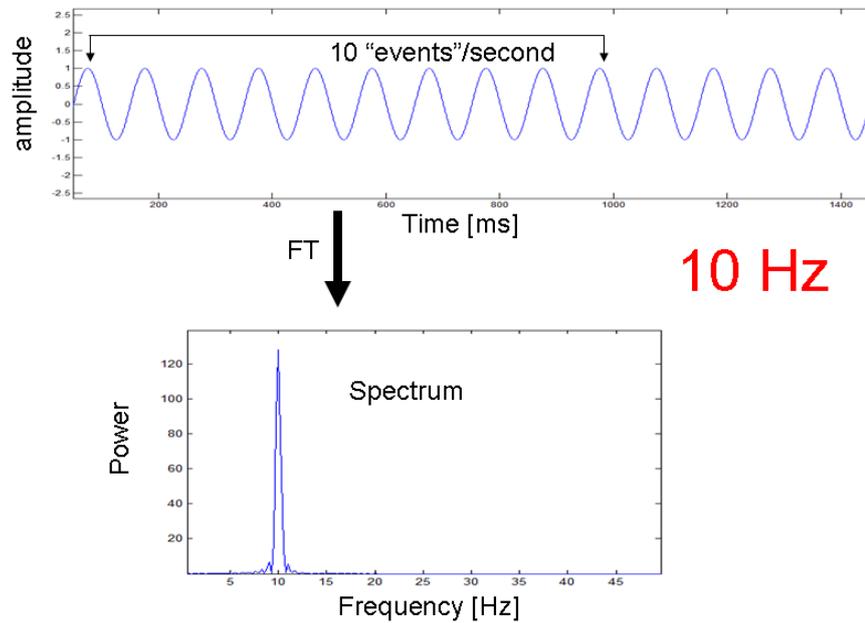
Changes in EP waveform over the recording period were seen in all birds, but differed dramatically between tutored and acoustically isolated birds. Figure 5.11 compares early and late examples of vocalizations from three tutored and three untutored birds (Fig. 5.11A and 5.11B respectively). The averaged EP shapes (across 100 renditions) are shown for each bird (blue and red traces represent early and late recordings, respectively), next to their corresponding sonograms. For each bird, the upper sonogram (in blue frame) is a sample taken from the beginning of a pressure recording period and the lower sonogram (in red frame) is an example obtained from the end of the recording period. For each bird we also present the EP spectra (produced by FFT analysis of EPs), which we use to quantify the amount of temporally fine structure present in the EPs. It is now time to describe this method as is done in the following section.

### **5.3.2 Application of Fourier Transformation to quantify the presence of fine temporal structure in the EPs**

As noted earlier, when EP structure develops over the period of song learning, temporally finer structure is added to the pre-existing coarse structure. We refer to this process as pressure refinement or motor skill acquisition. The addition of fine structure can be quantified by first decomposing the EPs by Fourier Transformation. The spectra thus obtained can be collapsed across renditions. The averaging of EP spectra is done in order to determine the repetitive frequencies e.g. the fine structure

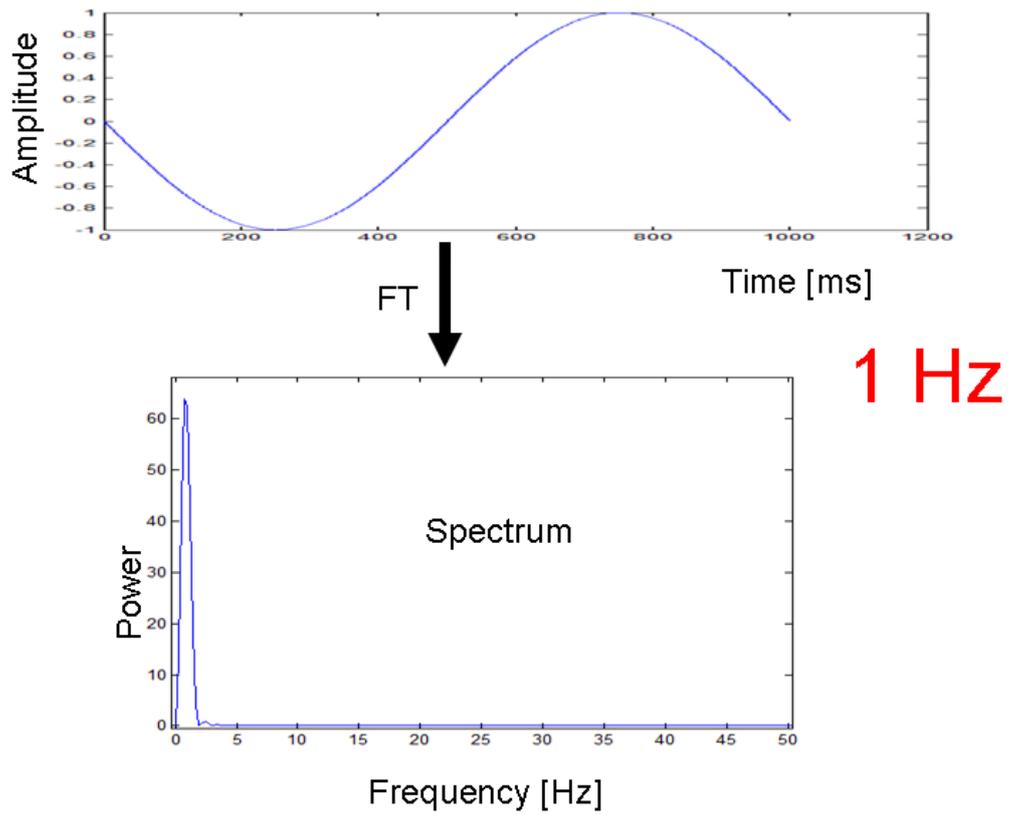
of EPs that is consistently repeated across renditions (see Methods and definitions for details).

We can demonstrate this method by a simple simulation. Any signal can be decomposed in its frequency components by Fourier Transformation (FT). Take a simple sine wave with a frequency of 10 Hz and decompose it by FT as shown in Fig. 5.6 below:



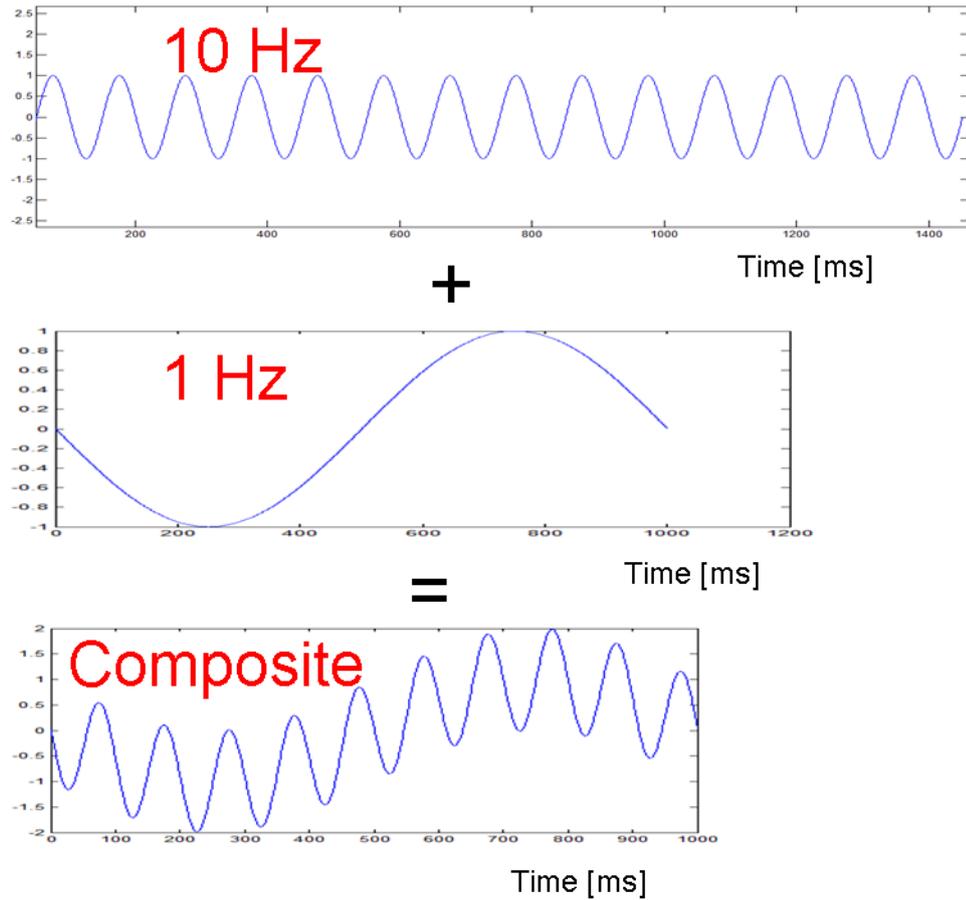
**Figure 5. 6** Fourier transformation of a 10 Hz signal. See text.

The obtained spectrum has one peak at 10 Hz. Now let us look at another sine wave of 1 Hz frequency and decompose it (Fig. 5.7):



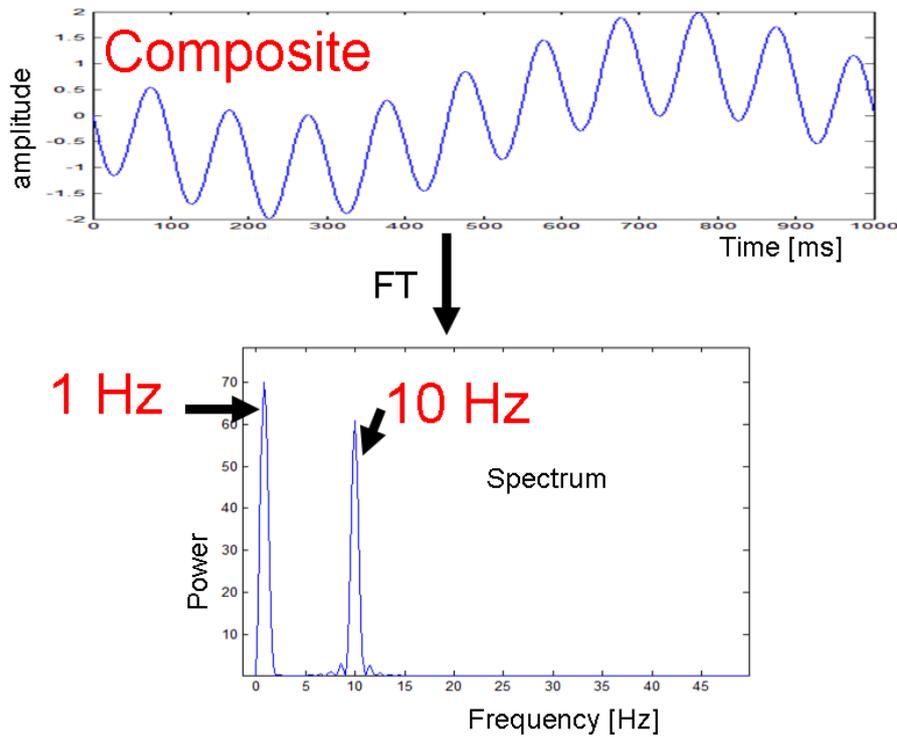
**Figure 5. 7** Fourier transformation of a 1 Hz signal. See text.

This signal has a single peak at 1 Hz. Next let us *add* the 10 Hz signal to the 1 Hz signal to obtain the *composite* signal (Fig. 5.8):



**Figure 5. 8** Adding signals with different frequencies results in a composite signal. See text.

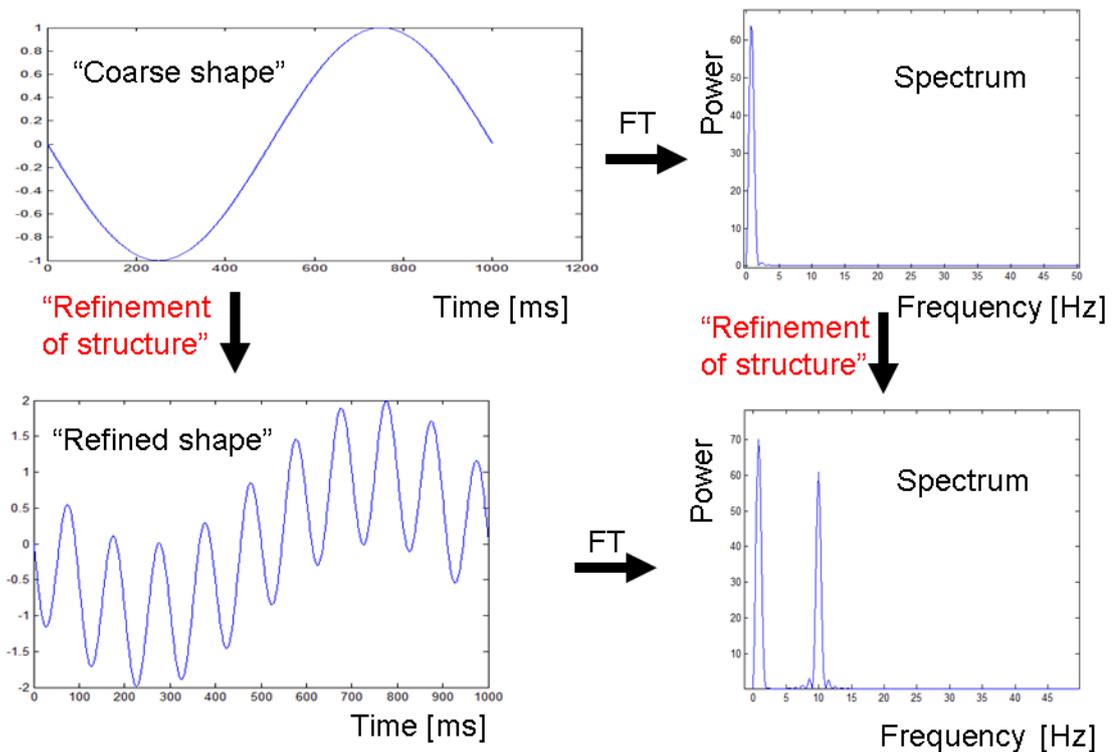
This composite signal can then be decomposed by FT to obtain its spectrum as shown below (Fig 5.9):



**Figure 5. 9** Fourier transformation of a composite signal. See text.

The spectrum of the composite signal has two peaks, one for each component: 1 Hz and 10 Hz peaks.

We can represent the coarse structure of EPs with the low frequency signal (low level of structure modulation) and the “refinement” of EP structure with the *addition* of a high frequency signal to the pre-existing coarse structure. Therefore the refinement of structure should result in a new, higher frequency peak in the spectrum as summarized in Fig 5.10 below:



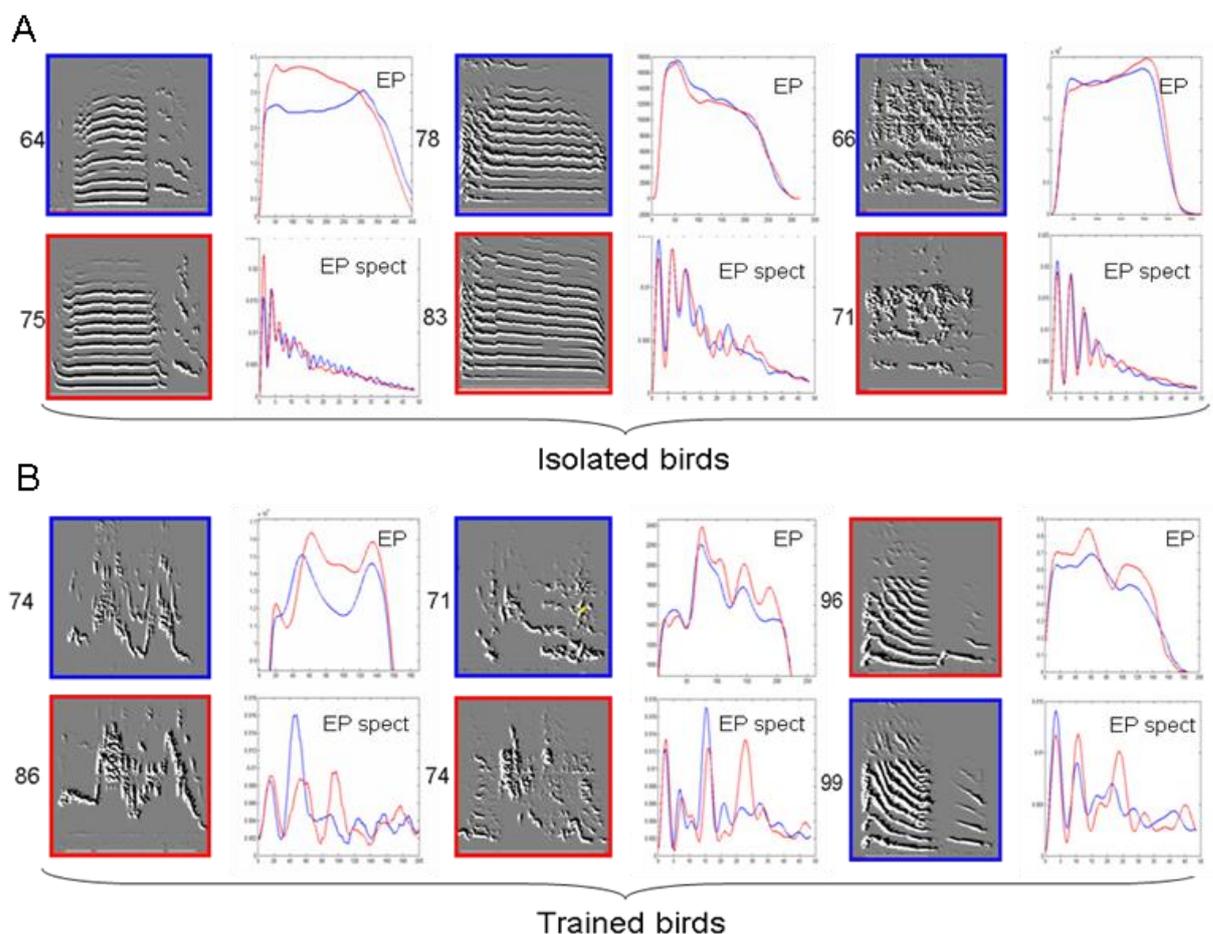
**Figure 5. 10** Adding a higher frequency “fine” structure to the low frequency “coarse” structure results in a shift of spectral power towards higher frequency. See text.

Thus we predicted that with refinement of EP structure of trained birds the spectral power would shift toward higher frequency. On the other hand, spectra of isolated birds, where no or little structural refinement happen should show no such shift of spectral power. (To eliminate the effect of noise that could potentially get added to the EP measurements we calculated means across spectra so that only the repeated frequencies “survive”.)

### 5.3.3 Structural refinement happens predominantly in tutored birds

In Figure 5.11 compare the spectra of isolated birds to the spectra of tutored birds.

Note that even at the end of the recording period (red spectra) untutored birds have less power concentrated in lower frequency ranges, which indicates that their EPs do not have as much fine temporal structure as the EPs of tutored birds. Moreover, isolated birds do not show *improvement* of motor skill, as measured by the shift of power of EP spectra towards higher frequency ranges, while trained birds do show a shift, indicative of pressure refinement.

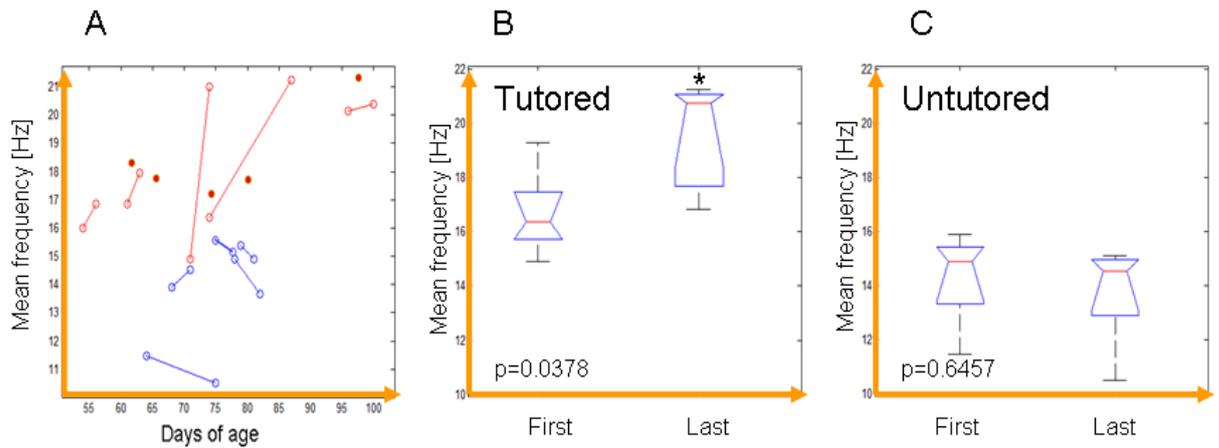


**Figure 5. 11** Comparing early to late EPs and their spectra in trained birds and isolates.

**A**, Three examples of isolated birds are shown. Blue traces indicate early samples, red traces late samples. The corresponding sonograms are shown (blue frames indicate early samples, red frames late samples). **B**, In trained birds spectral power shifts from lower frequencies in early recordings (blue curves) toward higher frequency ranges in late recordings (red curves).

To quantify these differences further we performed this spectral analysis across five tutored birds and compared them to 5 un-tutored birds. Figure 5.12 shows the mean frequency of spectra for tutored and untutored birds. Empty circles indicate beginnings and ends of recording periods for individual birds. Each mean frequency is calculated as an average of mean frequencies of all types of EPs analyzed per individual bird (typically 1-2 types of EPs). Red circles and links indicate 5 tutored birds, while blue circles and links represent acoustically isolated birds. It is possible that the difference of EP structure refinement that we measure could be explained with the particular song model we used to tutor the birds with. So we also analyzed end-point samples of EPs from birds tutored with different song models (filled red circles). We found that the mean frequency of all 10 tutored birds was higher than the mean frequency of all isolated birds (Fig. 5. 12A).

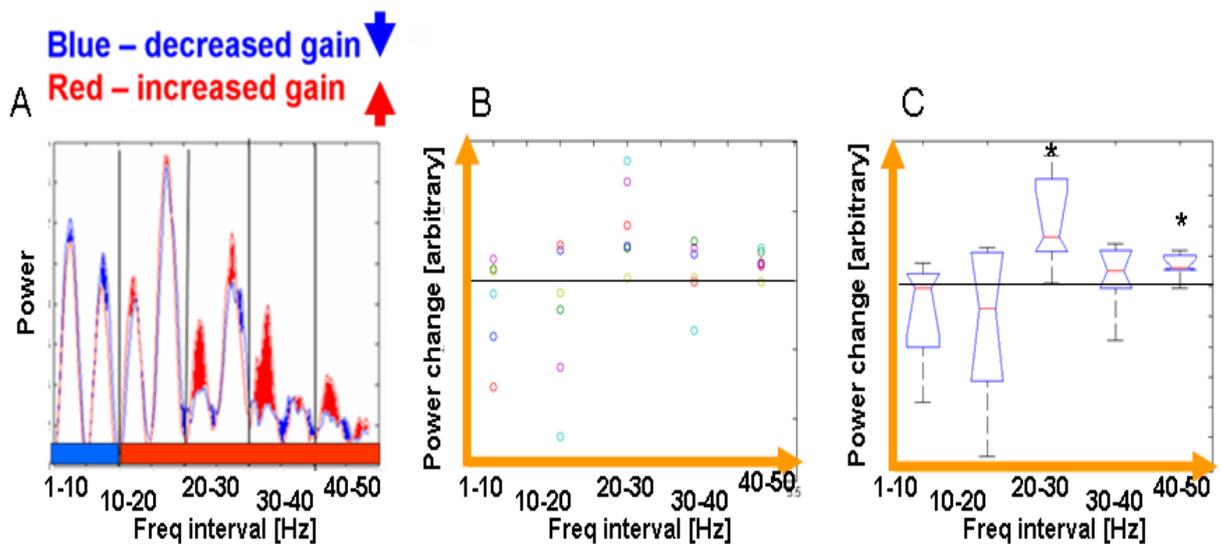
While isolated birds may show less fine structure of their EPs than tutored birds, they might still be able to improve this measurement during a period of development. However, we found that across birds this mean frequency shift, indicative of refinement process, was significant only in tutored birds (ANOVA1;  $p=0.0378$ ) but no significant change (ANOVA1;  $p=0.6457$ ) was observed in isolated birds (Fig. 5. 12C). Thus we detected no improvement of EP refinement in untutored birds.



**Figure 5.12** The mean frequency of EP spectra increases in trained birds and is overall higher than in acoustically isolated birds. **A**, Circles indicate mean frequencies at the beginnings and ends of recordings (see methods). Red: trained birds; blue: isolates. Full red circles: samples of birds trained with different models. Mean frequency shifts across trained birds and is overall higher **B**, than in isolates **C**.

We wanted to characterize in what frequency ranges addition of fine structure happens, which corresponds to the *time scales* at which a bird refines his song. Figure 5.13A compares spectra from the beginning and end of recording in one tutored bird, and this example shows how power is added to certain higher frequency ranges (red fill), while power is lost in lower ranges (blue fill). Figure 5.13B shows the changes in power spectra in 5 tutored birds (different colors of circles represent the 5 subjects). Data points above the zero line indicate an increase of power for the corresponding frequency range. The points below the zero line indicate decreases of power. Figure 5.13C shows a box-plot of the data from Fig. 5.13B. Power significantly increased in high frequency ranges of 20-30 Hz and 40-50 Hz (*t-test*; *p-value* = 0.0244 and 0.0143 respectively), with the maximal increase in 20-30 Hz range. The addition of fine structure in these frequency ranges is also significantly higher than in 1-10 and 10-20 Hz ranges.

In terms of time-scales this result means that the greatest addition of fine structure to EPs spans from ~30-50 ms in length. For example, this would correspond to the broader three peaks of the middle EP in Figure 5.5A (note the deepening of the valleys – red arrows, separating these 3 parts of the EP, from Day 55 to Day 60 in Figure 5.5). The addition of fine structure was also significant (*t-test*; *p-value* = 0.0143) at the shorter time-scales, from 20-25 ms (40-50Hz in Fig. 5.8C). This finer structure could correspond to the peaks and valleys such as those added to the EP on Day 72 in Figure 5.5A.

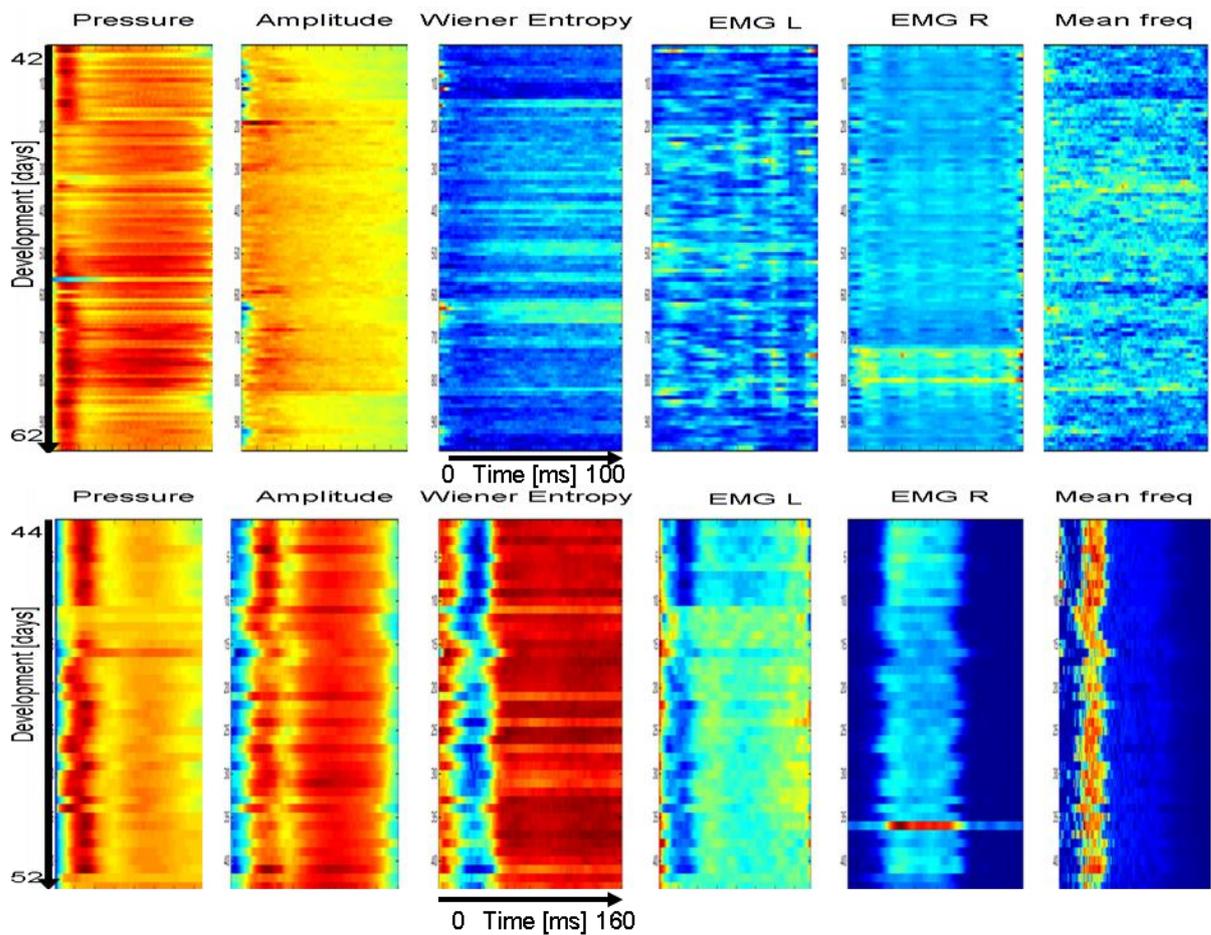


**Figure 5.13** Identifying the frequency intervals where power changes occur in tutored birds. **A**, Example of power changes in one bird (red: increase; blue: decrease). **B**, Power changes in 5 EPs of trained birds in 10 Hz intervals. Units of change are arbitrary. **C**, Box plots of power changes across 5 EPs of trained birds. Significant changes of power are indicated by “\*”. The highest increase of power was in 20-30 Hz interval, corresponding to time scale of 30 – 50 ms. We also see significant increase of power in 40-50 Hz interval indicating the addition of temporal fine structure as short as 20-30 ms.

### 5.3.4 Changes in respiratory activity are closely associated with changes in syringeal muscle activity

As already noted, previous studies have shown that fine temporal structure of respiratory pressure is not affected by the resistance of the syrinx (Cooper and Goller, 2004). This fact is critical for the interpretation of our results because we wanted to associate the structural refinement with the activity of a particular set of muscles (respiratory muscles). We also wanted to know whether the refinement of respiratory pressure would precede the refinement of activity other components of the song production system (see Discussion).

To find the answer we recorded the activity of the syringeal muscles by EMG, over a period of song development, in three birds. We cross-correlated the developmental time courses of parts of EPs with the same time courses in the EMG signal where developmental changes were observed and found no lag (median across three birds:  $r=0.47$ ,  $lag=0$ ). Figure 5.14 shows the aligned EPs (first column), acoustic features and the EMG in two birds. Note that especially where the correlations between EPs and EMG are strong (lower row) there is no apparent lag between the two correlated signals. Although this is a small sample of birds, the correlation between changes in EP shape and changes in EMG signal from the syrinx is indicating that structural refinement is probably coordinated between the two components of song production.



**Figure 5.14** EMG activity is correlated with respiratory pressure with no lag. Respiratory pressure and EMG was recorded together with sound (in the laboratory of Franz Goller, University of Utah). EMG was recorded from both left and right muscles of the syrinx (EMG L and EMG R). Three birds were analyzed (two are shown). In cases where respiratory pressure and EMG signals were strongly correlated (lower row) the lag of cross-correlation was zero (compare pressure to EMG L in the lower row).

## 5.4 Conclusion and discussion

In this chapter we compared the syllabic development between tutored and untutored birds. Here are three of our main findings:

1) Tutored birds add more fine temporal structure to their EPs (expiratory pulses) than do untutored birds, over the course of song development. This confirms the previous finding (Méndez et al., 2010).

2) There is no indication that untutored birds add any fine temporal structure to their EPs during song development, although changes in their EP shapes do occur. These developmental changes do not seem to include addition of new peaks and valleys or other increments of modulation (such as peaks getting higher, valleys getting deeper), that could be detected by our method.

3) In the samples of EMG data from syringeal muscles where developmental changes of EMG shapes were closely associated with changes in the EP shapes, we did not find a lag of either signal. While due to small sample size (three birds) this result is not fully conclusive it could suggest that the addition of fine structure has to happen on both, respiratory and syringeal level of song production system simultaneously.

In human speech respiration has to be coordinated with articulation (Conrad and Schönle, 1979) and the control of the breathing pattern is a critical art that has to be learned by singers (Sundberg Johan, 1993). The right amount of air has to pass through the articulatory and vocal structures in order to produce the desired sounds. On the other hand, the gas exchange in the lungs has to be uninterrupted during speech or singing. However, in human vocal behavior there is no need for fine temporal coordination of respiration. Instead, most of the articulation is performed by

articulatory structures controlled by different muscle systems. As the result human respiratory pattern lacks time complexity found in some songbirds.

In zebra finches, on the other hand, the respiratory activity during singing is itself involved in articulation and, as a result of this, possesses a great temporal complexity that corresponds to the acoustic complexity of the syllables (Wild et al, 1998). This is perhaps surprising because, in principle, the fine articulation of song structure could be carried out by the syrinx alone (this would be analogical to human articulation), which would pose a lesser challenge for precise temporal coordination between syrinx and respiratory muscles. If this was the case, then it might suffice for respiratory activity to determine the rhythm or the syntax of the song, while the fine intra-syllabic structure would be learned on the level of the syrinx. But in zebra finches temporal structure of respiration closely follows the complex acoustic structure of syllables. This would imply that the learning of fine temporal structure of respiration must thus co-occur with the learning of other motor gestures (syringeal activity and beak movements). We show that tutored birds add fine temporal structure to the coarse respiration pattern during the song development.

While the major advantage of our method is in that we can indirectly measure the activity of (respiratory) muscles rather than acoustic structure (which results from complex interaction between muscles, membranes in the syrinx and respiratory muscles), there is also a possible disadvantage. It lays in the fact that respiratory signal is smoother than some acoustic features (see Fig. 5.5) so we may be missing even finer time-scale segmentation of syllables than what we can see in respiratory pressure. To detect such fine time-scale events (and relate them to central activity)

one could record chronically from nuclei such as RA or downstream from it. Our attempts to produce such data have not been successful.

#### **5.4.1 Bias towards structure refinement in tutored birds**

That the song of untutored birds is significantly different from “normal”, tutored birds has been known for some time (Marler and Sherman, 1985). The difference in the respiratory pressure pattern between tutored and untutored birds has also been established in a recent study (Méndez et al., 2010) by measuring the degree of EP modulation at the end-point of song development together with the syringeal muscle activation. But, because the study is focusing only at the end-point of development, it cannot establish whether the untutored birds also follow the trend of refinement of motor gestures over development, but to the lesser extent than tutored birds.

Our results suggest that while there is a measurable systematic bias in birds exposed to the song model toward addition of fine temporal structure to the respiratory pressure patterns, this bias was not detected in untutored birds. Thus the difference between the two groups is not merely in the degree of motor gesture refinement (which has been shown before) but in the *direction* of development (qualitative difference). The lack of motor refinement in untutored birds could be caused by the broad time scale at which the vocal exploration is carried out (and at which the song can be modified). Because there is no model to imitate, the “goal” of song development may not be locally defined at all. In terms of hierarchical learning, which we discussed earlier, isolates may be stuck at the lowest level, where only

broad acoustic changes can be made. Indeed, our observations show that developmental changes of EP structure do occur in untutored birds (the song does develop) but these changes remain broad. In other words, the untutored birds may not be able to produce developmental changes at the short time scale.

However, as mentioned earlier, even the untutored birds sometimes produce songs with complex but variable temporal structure (e.g. their songs are inconsistent). One advantage of our method is that we detect only *consistently produced* fine temporal structure of the song (by first decomposing EP shapes by FT and then *averaging* the EP spectra, see Methods and definitions). Therefore any untutored birds with highly temporally complex but variable songs would still be detected as possessing low level of temporally fine structure.

While it is in principle possible that untutored birds can “catch up” with the tutored birds later in development, this is unlikely for two reasons: first, as discussed above, untutored birds do not show any bias in the direction of motor refinement, and the second reason is, that the acquisition of motor skill is thought to be limited by the “critical period”, when substantial developmental changes can still be made (Konishi, 1965). The oldest tutored bird that we recorded pressure from (the recording session took place between 94 to 99 days post-hatching) also shows the lowest amount of pressure structure refinement (Figure 5.12A). This would suggest that the end of the “critical period”, when refinement of motor gestures can still occur, was already quite close for this bird. Therefore, it does not seem that untutored individuals would ever “catch up” with the tutored ones.

The inability of untutored birds to “catch up” with the tutored ones is, however, probably limited to individuals but not communities. A recent study by Feher et al. (Feher et al., 2008) looked at the syllables of zebra finches that were raised in isolation from normal (“wildtype”) songs but not from each other, and demonstrated that (acoustically) parts of syllables (“vocal states”) tend to get shorter over several generations of “tutors” and “pupils”. In other words, there is a systematic bias of the whole population to culturally evolve the song with fine temporal structure.

One problem with identifying just the length of vocal states from acoustic signal alone is that this method will not likely detect all developmental changes. A potential advantage of our method is that we can detect any changes in modulation of EP structure, even if the lengths of vocal states do not change. Therefore any trend in untutored birds towards “normal” song should be detectable by our method.

Veit et al. (Veit et al, 2011) have recently shown that at least one aspect of respiratory structure in very young birds is established before the acoustic structure (namely the length of gaps between syllables is adjusted on respiratory pressure first). This study therefore indicates that in young birds, where AFP has the dominant control of song production, the learning is focused on the establishment of respiratory patterns. There is also anatomical evidence (Vicario, 1991) that there is some degree of segregation in nucleus RA between parts controlling respiratory system and parts controlling the syringeal muscles, making it possible that refinement of articulation would be learned separately from refinement of syringeal activity.

However, this body of evidence does not necessarily imply that on the level of the syrinx the acquisition of fine motor skill (such as learning of intra-syllabic events) lags behind the respiratory system. Even in the absence of acoustic evidence the refinement of activity of syringeal muscles may still be present. As discussed earlier in this chapter, the acoustic output is the final product of integration of respiratory, syringeal and beak activity. The imprecision of this integration may result in apparently delayed measurements of refinement on the acoustic level. Indeed, as our results from EMG recording of syringeal muscle activity seem to suggest, there is no apparent lack of synchronicity between developmental changes in the pressure pattern and changes in syringeal EMG activity. Therefore we suggest that learning of fine temporal structure of syllables occurs on all levels of song production at the same time.

