

behavioral measurements (Werker et al., 1981; Tsao et al., 2004), auditory Event Related Potential (ERP) recordings (Mills et al., 2004; Rivera-Gaxiola et al., 2005; Wunderlich et al., 2006) and functional Magnetic Resonance Imaging (fMRI) (Dehaene-Lambertz et al., 2002). In songbirds too, song development is associated with changes in auditory perception (Braaten et al., 2006) and auditory responses: single neurons in the song system and auditory nuclei develop stronger responses to their tutor's song, as well as to their own song, over a sensitive period for vocal learning (Volman, 1993; Doupe, 1997; Doupe and Solis, 1997; Nick and Konishi, 2005; Amin et al., 2007). In addition, motor vocal activity is essential for development of song preferences in males, as muted male zebra finches (*Taeniopygia guttata*) are not able to develop behavioral preferences for conspecific song, in contrast to normal birds (Pytte and Suthers, 1999). Does the shaping of auditory responses to vocal sounds depend on early song learning experience? If so, what sort of early auditory, vocal or social experience, is required for shaping response selectivity?

In European starlings, the development of selective auditory responses depends on social experience. Social deprivation during development affects auditory responses in adulthood (Cousillas et al., 2006): multiunit recordings across multiple sites in the principal forebrain auditory nucleus Field L indicated that, as in mammals, social deprivation results in a larger auditory area (increase in number of sites that show auditory responses) and reduced specificity of responses (units tend to respond to multiple types of stimuli). Further, in wild caught birds, but not in the socially deprived birds, time courses of responses show sharp and highly synchronized onset responses. According to Cousillas et al, larger auditory areas and poorer selectivity are an effect of social deprivation, as socially deprived birds that could hear songs show similar abnormalities.

Here we revisit this question using alternate methods and a different species of songbird (zebra finch). Instead of using multi-unit electrophysiology we combined fMRI and ERP approaches to test how early auditory experience shapes responses in the entire auditory forebrain of the adult zebra finch. This approach allowed us to quantify the overall pattern of brain activation – in brain space and in time - to songs and syllables; such large-scale brain activation may be correlated with perception (Murray et al., 2002). We quantified the consistency in auditory responses to conspecific songs, calls, and syllables, comparing birds that had different developmental experiences: from birds who had rich social and acoustic experience during development, through birds that had no social experience and only minimal song learning experience, and birds that had no social or song learning experience during the sensitive period for song learning. We were interested in the possibility of coupling between the sensory-motor song learning period and the shaping of stimulus-specific auditory response patterns to conspecific sounds, because sensory plasticity might guide the development of motor skills, and vice versa. As noted, the zebra finch female provides us with an interesting contrast: as opposed to males, the females do not sing. However, females select their mates by judging the quality of their songs: what sort of early experience is required for females to develop their taste in songs (O'Loghlen and Rothstein, 2003)?

Methods

All experiments were approved by the Institutional Animal Care and Use Committee of Weill Cornell Medical College and the City College of City University New York.

Subjects

Auditory responses were measured (using fMRI or ERP) in adult birds (age range: 4-24 months old).

Isolated males—fMRI n = 5 birds, ERP n = 9 birds. Males hatched in a mixed gender colony room. Between day 4-7 post hatch the father was removed and the cage was taken, together with the nest and the mother, to a colony room occupied by females and chicks only. This procedure prevents early exposure to songs, as zebra finches have elevated hearing thresholds prior to day 10 post hatch (Amin et al., 2007). Chicks were raised by their (non-singing) mother until day 29 post hatch. From day 30 on males were kept individually in sound-attenuating chambers until auditory responses were measured in adulthood, between ages 4-11 months post hatch.

Isolated females—Females (fMRI n = 5, AEP n = 5) were raised, isolated and tested in the same way as the isolated males described above.

Box trained males—Males (fMRI n = 5, AEP n = 9) were isolated and tested as described above for the isolated males, except that each bird was trained with one of three song models, from day 43 to day 90 post hatch. Training procedure has been described elsewhere (Tchernichovski et al., 1999). Briefly: we placed two keys in each box, and the birds learn to peck them to trigger a brief (about 1.5 sec) song playback. We limited playbacks to 20 per day. On day 90 post hatch, we removed the keys and stopped the training. Birds were kept isolated in the sound chambers without training until auditory responses were measured in adulthood, between ages 4-11 months post hatch.

Colony males—Males (fMRI n = 5, AEP n = 5) were randomly selected from our zebra finch colony. Birds were kept in family cages of one reproductive pair and their offspring (of up to 10 birds per cage). The age range of colony males was higher than age ranges of other groups studied here: 9-24 month old. However, we did not observe any apparent age-related effect on auditory responses in this, or any of the groups, where age varied between 4-11 months. Colony males were not isolated prior to auditory testing to avoid stressing them.

Auditory stimuli

fMRI stimuli—In the fMRI experiment auditory stimuli were delivered using a pair of stereo headphones with the magnets removed. (The constant magnetic field of the MRI scanner is strong enough to replace the field of the headphone magnets.) The sound pressure level of the auditory stimuli at the head position was about 100 dB; the background noise during the EPI sequence was constant at about 80 dB. Because of the complexity of natural stimuli (songs, calls, etc), we could not fully normalize auditory parameters (including durations and amplitude variance), and we used the following adjustment to reduce non-stimulus specific variance given these constraints: stimuli were normalized with respect to peak amplitude (keeping amplitude range the same for all sounds). The auditory stimuli were the same across groups unless stated otherwise: A 2 kHz pure tone (TONE); a conspecific song (CON); a second conspecific song (CON2) played to females only to compensate for missing stimulus categories; bird's own song (BOS, uniquely defined for each male); the tutor song (TUT, varied across males); two calls (SYLL1, SYLL2). We designed our experiment starting from the naturally given constraint that TUT and BOS stimuli are uniquely defined for each individual bird. For TONE, we restricted the experiment to one sound, too, as one representative for the class of all synthetic sounds. In the same way, to keep our overall experimental design consistent, we chose representatives for CON and the syllables, which were not further varied. All stimuli were applied to all birds with the exception of BOS and TUT which were not applied to females, and CON2 which was only applied to females (in order to balance the number of stimuli, replacing BOS). Stimuli were delivered in 16 blocks each consisting of a 32 s “on” and a 32 s “off” part. The durations of the stimuli were highly variable: CON 730 ms, CON2 746 ms, BOS

1407 ms (standard deviation: 325 ms), TUT 1095 ms (standard deviation: 358 ms), SYLL1 299 ms, SYLL2 174 ms. We set tone duration to 1000 ms, and then stimuli TONE, CON, CON2, and BOS were played out every 2 s. The two calls were played out every second to partially compensate for their shorter duration (at the expense of more repetitions). This scheme ensured that the beginning of each MRI signal sample phase coincided with the beginning of the stimulus. The overall scan time per experiment was 1024 s (256 sequential data points).

ERP stimuli—Auditory stimuli were delivered in free-field using a small speaker located approximately 20 cm above and perpendicular to the bird's head. The sound pressure level of the auditory stimuli at the head position was adjusted to 75 dB SPL. Stimuli were naturally produced female calls and male song syllables: SYLL 1 194 ms, SYLL 2 175 ms, SYLL 3 89 ms, CALL 1 252 ms, CALL 2 184 ms, CALL 3 80 ms. SYLL 1 and SYLL 2 were the same syllables presented in the fMRI experiment (SYLL 1, SYLL 2). One recording session consisted of 100 consecutive presentations of a single stimulus with 3.5 seconds between stimulus onsets.

fMRI procedure

We followed the protocol used in Voss et al. (Voss et al., 2007). Shortly, each bird was sedated with 50 μ l Diazepam (Abbott Labs.) intra muscular (1.66 mg/ml Diazepam in normal saline solution) 10 min prior to MRI scanning and immobilized in a restraining device made of two soft plastic tubes. The head tube was fixed in a custom-made coupled solenoid-type radiofrequency coil. The restrained bird was then placed into a custom-made sound-attenuating box that fit within the scanning bore (dimensions: inside box 25 \times 18 \times 12 cm – height; outside box 40 \times 30 \times 29 cm - height). The walls of the box consisted of multiple layers of acoustical foam and the outside form was wrapped in several layers of acoustical rubber. Under those conditions, average stimulus intensity was about 100 dB SPL in a background scanner noise of about 80 dB.

MRI parameters—BOLD sensitive images were acquired on a GE Excite 3.0 T MRI scanner with 50 mT/m gradients using a four-shot 2D gradient EPI sequence with TE/TR = 25/1000 ms, yielding an effective repeat time of 4 s. Eight sagittal slices of 1 mm thickness along the sagittal plane, 4 cm FOV, and a matrix size of 128 \times 128 were acquired. Each voxel measured 1 mm (sagittal slice thickness) \times 0.3 mm \times 0.3 mm. Slices were prescribed in right-left direction, covering the forebrain. Additionally, in-plane anatomical images and field maps were acquired.

fMRI post-processing—BOLD sensitive EPI images were corrected for distortions using a reversed phase gradient method and field correction maps (Voss et al., 2006). The images were then despiked and motion corrected using AFNI (Cox, 1996) and further processed using in-house software written in MATLAB: Data were spatio-temporally smoothed and significance for a positive BOLD effect was defined by correlating the signal intensity with the “on-off” block stimulus indicator function. In this process, the motion correction parameters were taken as nuisance parameters and were regressed out. Correlation-coefficient-based statistical parametric maps (SPMs) obtained in this manner were co-registered (Periaswamy and Farid, 2003) to an EPI brain template. Activation in areas with an EPI intensity baseline less than 20% of the maximum slice intensity and activation in the eyes was discarded. In the SPMs, bird-averaged activations with significance $p < 0.005$ are shown. P-values were corrected for spatial multiple testing based on Gaussian random field theory (Worsley et al., 1996), after converting correlation coefficients into z-values. The area of activation was measured for every bird as the number of voxels significantly activated (voxelwise $p < 0.005$) in a region of interest including the posterior forebrain,

summed over the two medial slices 4 and 5. Areas of activation were localized by comparison with the location of field L in the three-dimensional zebra finch MRI atlas kindly provided by C. Poirier (Poirier et al., 2008). The brain and cerebellum were manually outlined and overlaid to our template anatomical MRI images, and a good match was observed between these data sets (area of interest we are looking at is larger than the observed activations). Note that after filtering eye movement artifacts, area and intensity of BOLD activation were computed for each bird as described above. Those per-bird measures were used as the input for the ANOVA, as well as for obtaining average activation images across birds. Our measures of significance did not rely on averaged data, which were used only for visual displays.

ERP Procedure

Electrode implantation—Epidural electrode arrays consisted of copper alloy pin electrodes (0.5 mm) mounted on a plastic substrate (we used conventional machine interconnect strips) adapted from Espino et al. (2003). Three electrode pairs (6 electrodes) were implanted in each hemisphere at different positions along the coronal plane (anterior pair, posterior pair, middle pair, see Figure 1d in main text). One single pin was implanted above the cerebellum as a reference electrode. Electrode pins were coated with epoxy so that copper alloy was exposed only at the tip. Pin length was adjusted to conform to the skull curvature. During implantation the bird was anesthetized with a Xylazine-Ketamine mix (16.25 mg/ml and 8.12 mg/ml respectively in normal saline solution) i.m. and mounted in a stereotaxic device. Using the bifurcation of the sagittal sinus as a reference (0.0 mm anterior, 0.0 mm lateral) the skull was marked at the following coordinates on the left and right hemispheres: 0.0 mm, 2.0 mm (posterior medial); 0.5 mm, 4.1 mm (posterior lateral); 2.7 mm, 1.8 mm: (medial medial); 3.2 mm, 3.9 mm (medial lateral); 5.0 mm, 1.1 mm (anterior medial); 5.5 mm, 3.2 mm (anterior lateral); one reference electrode was placed over the cerebellum at -2.0 mm, 0.0 mm. Both layers of skull were perforated at these coordinates using a 27 g surgical needle. The bottom pins of the electrode array were inserted into the holes until they touched the dura, and the plastic substrate rested atop the skull. Dental cement (Tylok Plus, Fisher Scientific) was used to secure all electrodes and formed a permanent electrode cap on each bird's skull.

Post-operatively birds were injected with Yohimbine (0.49 mg/ml in normal saline) i.m. to quicken the return to wakeful state. Birds were monitored for several days after the electrode implantation and all birds used in these experiments had normal behavior, singing, perching, eating, and hopping around the cage.

ERP recording procedure—To reduce movement artifacts during recording, birds were placed in a partial restraint and then placed in the electromagnetically shielded and sound-attenuating chamber and electrodes were connected to amplifier leads. Stimuli were presented in blocks of 100 with 3.5 seconds between stimulus onsets (100 presentations of one stimulus type). Responses did not appear to habituate over the course of one session. Responses were amplified using a $\times 10,000$ battery-powered Bioelectric amplifier (SA Instrumentation, San Diego, CA) and the analog signals were filtered (0.1 – 1000 Hz) prior to sampling at 5 kHz (Data Acquisition Card - PCI-DAS1602/16, Measurement Computing, Middleboro, MA). For each bird, stimulus blocks were presented consecutively, with 5 minutes of silence before each new stimulus block, and were limited to 5 or fewer blocks per day.

ERP post-processing—Data was epoched in segments of 2.4 s, median subtracted, notch-filtered (2nd order Butterworth filter, 58-62 Hz), artifact-rejected (1.5 standard deviations from noise floor), and down-sampled to 500 Hz. Stimulus onset was recorded as

an additional analog channel, which was used to align recordings to stimulus onset. Event related potentials (ERP) were then obtained by averaging across trials for each bird and each stimulus.

Features of evoked response

ERP time course—To determine the overall time course of the evoked activity we computed ERP *power* as the square sum across electrodes for each sample. Averaged over groups and syllables ERP power showed multiple response components in time (Supplementary Fig. 2 shows the root mean square in μV): onset response at 0-40 ms; a main peak between 40-120 ms; offset responses between 120-350 ms; and a weak prolonged response from 350-1000 ms. ANOVA tests for significant response patterns in these four time windows were corrected for multiple comparisons using Bonferroni correction.

ERP spatial distribution—To quantify the strength and predominant distribution of the evoked activity we performed principal component analysis. The first (spatial) principal component of the evoked potentials computed using all time samples, birds, and syllables indicated a rostral-caudal response gradient (explaining 74% of the variance in the ERP) with left-right symmetry. This motivated us to measure ERP amplitude as the potential difference between rostral electrodes and caudal electrodes. This is dominated by the activity over the caudal areas, with rostral electrodes serving primarily as reference electrodes.

Event-related de-synchronization—To quantify potential effects on oscillatory activity we computed trial-averaged spectrograms of the EEG in the mean across electrodes. This is the conventional approach used in EEG studies to assess the strength of oscillatory activity

Statistical ERP methods

Bootstrapping on F-statistic—In order to judge differences in ERP response patterns of isolated males and box-trained birds we computed the F-statistic for random selections of birds irrespective of their group. Specifically, the 18 box-trained and isolated male birds were randomly assigned to one of two groups of 9 birds each. The F-statistic for each of these random groups was computed and their difference in F-value noted. This was repeated 1000 times to estimate the likelihood that the observed difference occurred by chance. More extreme values than the observed difference occurred at a rate of $p < 0.05$.

For consistency, the same method was applied to the fMRI data F-statistics for the same groups (isolated males and box-trained birds), also resulting in a significant p-value ($p < 0.05$). The same method was applied to ERP and fMRI data of isolated males and isolated females to judge the difference in response patterns of these two groups.

Measurement of song imitation—Judged by visual inspections of sound spectrograms all box trained birds copied some of their tutored song. To quantify imitation accuracy we performed similarity measurements using Sound Analysis Pro version 1.4 (Tchernichovski et al 2001). We used the default settings, and used the % similarity measure.

Results

To examine the role of developmental experience in shaping adult auditory responses we controlled the auditory and social experience of the birds. Without exposure to song, isolated male zebra finches develop an abnormal, loosely structured song. Several seconds of operant song playback per day over development is sufficient to induce song learning and shape normal song structure in male birds that are otherwise acoustically and socially isolated

(Tchernichovski et al., 1999). Training with song playbacks allows us to disentangle social factors from song learning. We compared responses of isolated birds that did not learn a song, isolated birds that did learn a song from song playback, and colony birds, who were raised in a socially and acoustically rich environment. We examined auditory responses across groups using two imaging techniques: functional magnetic resonance imaging fMRI, for spatial resolution, and auditory ERPs, for temporal resolution. Both fMRI and ERPs can reveal large scale brain responses to sensory stimuli and thus may serve as probes of perceptual processing (Murray et al., 2002). We use these methods as independent assessments of auditory responses.

The effect of social isolation and early song exposure on BOLD responses to vocal sounds

We looked at activation area of the Blood Oxygen Level-Dependent (BOLD) response to determine location and extent of response to pre-recorded natural stimuli. Responses to auditory stimuli can be recorded using fMRI in songbirds both in the awake sedated (Voss et al., 2007) as well as anesthetized states (Van Meir et al., 2003; Van Meir et al., 2005; Boumans et al., 2007). We scanned lightly sedated birds through a custom-made sound-attenuating chamber. Playback (peak amplitude 100 dB) included unfamiliar conspecific songs (CON), birds own song (BOS), tutor song (TUT), repeated syllables (SYLL) and a pure tone. In all birds we observed strong BOLD activation levels in the posterior region of the forebrain, centered at auditory areas NCM and Field L (Fig. 1a-c). Taking the low BOLD signal-to-noise ratio and the relatively small distortions induced by echo planar imaging (EPI) into account, for statistical group analysis we took a conservative approach and pooled BOLD responses over the entire medial posterior forebrain. Occasionally we also observed BOLD responses in more rostral areas of the forebrain and in the midbrain (Figs. 2 a,c,e). Due to their relatively low statistical significance in the group statistical parametric maps, in stark contrast to the posterior responses, group comparisons and attempts to exactly localize these regions were not performed. We first examined BOLD responses in male birds raised in a colony (*colony males*). Colony males showed strong BOLD activation to songs and weaker activation to repeated syllables (Fig. 2a). Responses differed significantly across stimuli (Fig. 2b), with the strongest activation to songs and weaker activation to calls and tones (1-way ANOVA, $p < 0.01$, $F = 4.5$, $n = 5$).

We wondered if song tutoring by playbacks in isolation is sufficient to induce response patterns as observed in the colony males. *Box trained* birds were raised by the mother alone (who does not sing) from 7-29 days post hatch, isolated from song until training. From day 30-100 (during song development) birds were kept individually in sound-attenuated chambers and were trained with thirty seconds of song playback every day, starting from day 43 post hatch. This minimal exposure induced significant song imitation in all of the trained birds (mean similarity = $61 \pm 8\%$, Sound Analysis Pro 1.4, Tchernichovski et al 2001). In these box-trained birds heterogeneity of responses was still apparent (Fig. 2c). In agreement with previous results (Cousillas et al., 2006), we found that areas of activation were larger in the socially deprived box-trained birds compared to colony birds (Means and SEMs: 89 ± 14 voxels and 34 ± 9 voxels respectively). However, across stimuli, differences in responses were still significant and similar to the pattern observed in colony males, with strongest responses to songs (Fig. 2d, 1-way ANOVA $p < 0.05$, $F = 2.7$, $n = 5$).

We then examined BOLD responses of male birds that were raised in sound attenuated chambers like the box-trained birds, but were not trained with playbacks (*isolated males*). Isolated males developed abnormal isolate songs and showed activation to all sounds (Fig. 2e), but responses were highly variable across birds and on average equally strong to socially relevant (songs) and less relevant stimuli (e.g. tone). In isolated male birds too, areas of activation were large compared to colony birds (Mean and SEM isolate males: 108

± 6 voxels) but in contrast to the box-trained birds, auditory responses did not differ significantly across stimuli (Fig. 2f, 1-way ANOVA $p = 0.9$, $F = 0.2$, $n = 5$). In other words, the tendency (at the group level) to respond differently to different stimuli, which we observed in both colony birds and in box-trained birds, was no longer detectable in the socially and acoustically isolated males who were not exposed to song playbacks (see also BOLD responses of individual birds in supplementary figure 4).

Functional imaging results in males suggest that early exposure to song is necessary to shape auditory responses of the adult male brain. In females too, early exposure to songs affects song preference and neural responses in adulthood (Riebel, 2000; Lauay et al., 2004; Terpstra et al., 2006; Hauber et al., 2007). We wondered if early exposure to song is required for the development of stimulus-specific responses in adult female zebra finches, who do not sing. We examined BOLD auditory responses in 5 isolate females. Note that isolated males still hear auditory feedback of their own songs; isolated females do not hear any song and therefore experience even greater auditory deprivation than the isolated males. Despite this, and in contrast to adult isolated males, adult isolated females showed significant differences in responses across stimuli, much like the adult males that have learned a song (Fig. 2g & 2h, $p < 0.05$; $F = 4.9$, $n = 5$, Mean and SEM area of activation isolated females: 72 ± 14 voxels).

We summarize the findings across groups by using the F-statistic (for each group) as a yardstick of how variably each group responds to different stimuli. Namely, the F-statistic is a measure of the group tendency to respond more strongly to some sounds, and less strongly to other sounds (Fig. 2i). We refer to this measure as a *stimulus-specific response pattern*. As shown, response patterns are strong in the colony males, weaker in the box-trained males and absent in the isolated males (see also, Supplementary Fig. 3). In the isolated females response patterns are comparable to those found in the colony-trained males. To test directly if response patterns are significantly different across groups we performed bootstrap tests, computing the F statistic for random selections of birds irrespective of their group (box-trained, isolated male, isolated females). We found that response patterns in both the box-trained males and in the isolated females were significantly different from those obtained in the isolated males ($p < 0.05$). Note, however, that CON and CON 2 elicited responses that differed substantially even though these are the same class of stimulus, suggesting that idiosyncratic differences across stimuli might contribute to the outcome. In other words, a significant *stimulus-specific response pattern* does not necessarily imply *stimulus-category-specific response pattern*.

Event related potentials as an independent measure of stimulus-specific responses

As in all experiments on live animals, results obtained with fMRI should only be interpreted on the basis of the specific conditions that apply to the animal during the experiment. In general, the unfamiliar, dark, and noisy fMRI recording environment is probably stressful to the animal and the necessary restraint and sedation or anesthetization are likely to affect auditory responses (Cardin and Schmidt, 2003). Furthermore, the BOLD response cannot be directly interpreted as a neuronal response. Therefore, we used auditory Event Related Potentials to obtain an independent test of our findings, as well as to diminish some of the experimental burdens on the animal mentioned above: ERPs are a direct measure of large scale changes in neuronal activity that reflect a temporal component of the brain response, and they can be recorded relatively quickly and without sedation in a quiet environment. Exploring evoked responses across all birds we found that stimulus-sensitive response patterns were stronger at short latencies (see Supplementary Fig. 1), during the onset of the evoked responses (corresponding to the first 40 ms of the sound). Although we see some evidence for stimulus specificity during the later phase of the ERPs (Supplementary Fig. 5) those differences were not statistically significant. We therefore limited our investigation to

different calls and syllables during stimulus onset, and did not include entire songs in the stimuli. A posteriori investigation of the differences in onset responses showed no straightforward correlation with differences in amplitude or song features across stimuli (consistent with previous results) (Maul et al., 2007). Further, because of low signal to noise ratio, ERPs were averaged across 100 back-to-back repetitions of each syllable, and therefore, even the short latency patterns might mirror sensory integration over several repetitions of the entire syllable (e.g., it might indicate the latency of identifying the syllable after a few presentations).

It is difficult to judge if and how responses measured by fMRI might be related to the ERP measurements. Instead, we use the ERP as an independent measure of auditory response selectivity: for each bird we measured evoked potentials to playbacks of conspecific male song syllables (SYLL 1, SYLL 2, SYLL 3) and female calls (CALL 1, CALL 2, CALL 3), which varied in duration and acoustic complexity. Responses were recorded from 12 epidural electrodes (Fig. 1d). To assess the spatial distribution of the response variability we performed principal component analysis (PCA). The spatial distribution of the first principal component is shown in Figure 1e. This component captures 74% of the signal variance across time, stimuli, and birds. As shown, the main potential difference is in the rostro-caudal direction with a more uniform distribution of power in the lateral direction. Therefore, we measured potential differences between the eight caudal and rostral electrodes while averaging along the lateral direction to improve signal quality. This primarily captures the activity over caudal areas (coinciding with NCM/Field L) with the most distant (rostral) electrodes as reference. Across groups, song syllables ERP time courses were more complex compared to female calls (Supplementary figure 5-6) with apparent offset responses to song syllables only. Overall, more complex stimuli (e.g., SYLL1 and CALL1) induce more complex time courses of ERPs in groups that showed stimulus specific responses (Supplementary figure 5).

Figure 3 summarizes the ERP results across groups obtained for the onset response. Evoked potential of onset responses (average of 10-40 ms post-stimulus) to two example stimuli show a strong rostro-caudal potential to a male syllable (SYLL 1) but weaker rostral-caudal response to a female call (CALL 1) (Fig. 3a). Similar to the fMRI results, colony males showed significant response heterogeneity across stimuli (Fig. 3b, 1-way ANOVA, $p < 0.01$, $F = 12.9$, $n = 5$). As with fMRI, minimal exposure to song playbacks in box trained males was sufficient to elicit stimulus-specific responses (Fig 3c & 3d, $p < 0.05$, $F = 7.8$, $n = 9$); whereas isolated male birds without song exposure did not show significant stimulus-specific differences in the rostral-caudal potential (Fig 3e & 3f, $p = 0.7$, $F = 1$, $n = 9$). In contrast to the isolated males, isolated females showed a significant response pattern (Fig 3g & h, $F = 5.1$, $n = 5$, $p < 0.05$; this and preceding p-values were adjusted for multiple comparisons using the Bonferroni method). Figure 3i summarizes stimulus specific response patterns across groups using the F statistic as a yardstick. Bootstrap tests show significant differences in response pattern between boxed-trained and isolated males, and also between isolated females and isolated males ($p < 0.05$).

In sum, converging evidence from fMRI and evoked potentials suggests that the development of stimulus-specific response patterns in the zebra finch male requires only a minimal level of song experience: less than thirty seconds of song playback per day in complete social isolation (box-trained birds) was sufficient to shape both ERP and BOLD responses across male birds. In contrast to males, for females auditory experience with songs may not be required to shape stimulus specific response patterns, at least to the level of shaping observed in the trained males.

Event-related desynchronization is abnormal in isolate males but not in isolate females

Stimulus specific response patterns were statistically significant only at the onset of the evoked potentials (10-50 ms after stimulus onset). However, the evoked response continues 100-200 ms after stimulus onset (Fig. 1f) and a more subtle response may be seen hundreds of milliseconds after stimulus presentation (Espino et al., 2003). In the average over syllables we observed a prolonged post-stimulus decrease in power (0.2 – 0.8 Sec.) in the 4-16 Hz frequency band as compared to the pre-stimulus activity. Figure 4a shows trial averaged spectrograms of the ERP in the mean across electrodes. The initial peak reflects the broadband evoked response to the stimulus (0-200 ms). Following stimulus presentation, oscillatory power decreases in the 4-16 Hz frequency band (300-800 ms post-stimulus). A similar effect has been documented in humans: attending to a stimulus and engaging in a perceptual or motor task is thought to reduce large scale synchrony which is reflected in decreased oscillatory electrical potentials in the 10 Hz frequency band (alpha activity) (Pollen and Trachtenberg, 1972). In human ERP literature a decrease of power in this frequency range is commonly referred to as event-related de-synchronization (Pfurtscheller and Lopes da Silva, 1999):

Although we have no conclusive evidence that the decreased power in the 4-16 Hz frequency band is sensitive to different stimuli (no significant stimulus-specific response patterns), when averaged over stimuli the desynchrony effect was markedly weaker in male isolates as compared to all other groups (Fig. 4b). This group difference was confirmed by a 2-way ANOVA: group, $p < 0.01$; stimuli, $p = 0.7$; interaction, $p = 0.9$.

We conclude that post-stimulus event related desynchrony in male zebra finches is dependent on early experience, and minimal exposure to song playback is necessary for its development. However, in females, no such experience is necessary. The deficit observed in the isolated male birds could suggest that the circuitry that handles longer term auditory processing is less active in these birds. Inferring from human and previous animal data, reduced desynchronization perhaps suggests that isolated male birds were less engaged in listening to these stimuli.

The post-stimulus desynchrony result suggests that in addition to a lack of stimulus-specific response patterns (BOLD activation and rostral-caudal potential), isolated males also show a more generic deficit in responses, or lack of attentiveness, to all stimuli, as compared to males that have experience learning song and to isolated females.

Discussion

In humans and in songbirds early experience during the sensitive period for vocal learning can shape auditory responses to vocalizations (Werker et al., 1981; Volman, 1993; Doupe, 1997; Doupe and Solis, 1997; Dehaene-Lambertz et al., 2002; Mills et al., 2004; Tsao et al., 2004; Nick and Konishi, 2005; Rivera-Gaxiola et al., 2005; Braaten et al., 2006; Wunderlich et al., 2006; Amin et al., 2007). The current study reveals some interesting aspects of this process: we found that 30 sec/day of exposure to playback of a single song (with no variability), which induces song imitation, is sufficient to shape auditory responses to a variety of unfamiliar stimuli. The response pattern of box-trained birds is more similar to that of colony raised birds than to that of untrained isolated male birds. The evoked potential responses revealed that responses were sensitive to different syllables and calls, whereas the fMRI showed differences in representative examples of responses to songs, calls and tones. Males that learned song during a sensitive period of development showed stimulus-specific response patterns in adulthood, whereas males that did not learn song, did not show such patterns. We observed an effect of social deprivation similar to that reported by Cousillas et al (Cousillas et al., 2006). Nevertheless we observed a strong, gender specific effect of song

exposure on auditory responses in zebra finches, despite social deprivation. Interestingly the female zebra finch, unlike the female starling, does not learn song, which may account for some of the differences between the present findings and those of Cousillas et al. (Cousillas et al., 2006). Non-singing females showed response patterns, with stronger BOLD responses to songs, in adulthood despite not hearing songs during development. This result indicates that, in contrast to males, some song selectivity develops in females even without prior exposure to songs. Taken together, these results suggest a link between the sensory-motor song learning period and the shaping of stimulus-specific auditory response patterns to conspecific sounds. We do not know if this link is related to vocal practice. For example, brief exposure to songs causes strong changes in (presumably sensory) activity in the song system during the first night after exposure – before vocal changes have occurred (Shank and Margoliash, 2009). Future studies in muted males might answer questions about the role of vocal practice in shaping auditory responses in males. In the female, some degree of hardwired auditory selectivity could compensate for the absence of vocal practice to the shaping of song selectivity. An unlikely alternative possibility is that in females only, auditory responses are shaped by exposure to female vocalization prior to day 30. In either case, evidence suggests that although auditory responses patterns are similar in male and female zebra finches, the timing, and possibly the mechanism, of shaping those responses is very different across genders.

Our study has several limitations. Using the 3T fMRI machine in awake birds, we can only obtain coarse spatial resolution compared to the zebra finch brain size. Therefore we could not determine if different auditory areas of the medial posterior forebrain (e.g., sub-regions of Field L and surrounding NCM) show different responses between brains of different groups. Recent studies in zebra finches suggest that stronger magnets (e.g., 7T) provide higher spatial resolution allowing more detailed analysis of contributions of specific brain nuclei to the BOLD response to various sounds (Poirier, 2009)

We found group differences in both BOLD intensity and in the size of brain activation, in some isolated males activation included large brain areas surrounding the auditory forebrain. This result compares to the larger auditory regions observed in socially deprived European starlings (Cousillas et al., 2006) using multiunit electrophysiology. We cannot exclude that the larger area of activation might be, at least in part, due to a spread of the hemodynamic responses. In either case, the interpretation of stimulus-specific response patterns in box trained and isolated female birds remains the same.

It is important to note that claims made here about female response patterns being “hardwired” apply only to the coarse shaping of auditory responses toward broad vocal categories, whereas the fine tuning of song selectivity must be experience-dependent. Female songbirds show strong behavioral discrimination across songs, and those are likely to be acquired early in life. For example, female show behavioral preference towards songs of older males and toward “local dialect” songs (Riebel, 2000; O’Loughlen and Rothstein, 2003; Lauay et al., 2004; Terpstra et al., 2006; Hauber et al., 2007). However, the effect of early experience on female’s song preference might develop on the foundation of generic hardwired tuning toward conspecific songs. Investigating the development of such preferences would probably gain from the higher resolution fMRI animal scanners which are now becoming available.

In the male, the stimulus-specific response patterns we observed might be shaped by the saliency of the song template (sensory hypothesis), or by vocal practice (sensory-motor learning hypothesis). A third possibility, supported by the similarity in response patterns across isolated females and trained males, is that the stimulus specific response patterns observed in isolated females is the default (perhaps hardwired) state. According to this

hypothesis, in the male there is a delay in the maturation of response patterns so as to facilitate perceptual plasticity during vocal learning. Perhaps for isolated male birds who do not learn song during development the “delayed” maturation of response patterns becomes permanent, resulting in lack of auditory response patterns in adulthood.

Some features of the adult male zebra finch responses might mirror maturation whereas other features might relate more specifically to the song playbacks. Although this study does not provide enough data to attempt a distinction between the two, both fMRI and EEG approaches allow repeated measurements over development and should make it possible in future studies to look continuously at how response patterns emerge during development. Further, it should be possible to examine side by side the development of song motor skills (when song becomes more and more structured) and the emergence of structured stimulus specific response patterns. Developmental sensory and motor information may help to answer the question: which comes first, sensory or motor crystallization?

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Josh Wallman for helpful comments and discussion. This work was supported by NIH (O.T.), IBIS (S.A.H. and H.U.V), NIH (S.A.H.) and an RCMI grant to CCNY.

References

- Amin N, Doupe A, Theunissen FE. Development of selectivity for natural sounds in the songbird auditory forebrain. *J Neurophysiol.* 2007; 97:3517–3531. [PubMed: 17360830]
- Boumans T, Theunissen FE, Poirier C, Van Der Linden A. Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. *Eur J Neurosci.* 2007; 26:2613–2626. [PubMed: 17970728]
- Braaten RF, Petzoldt M, Colbath A. Song perception during the sensitive period of song learning in zebra finches (*Taeniopygia guttata*). *J Comp Psychol.* 2006; 120:79–88. [PubMed: 16719585]
- Cardin JA, Schmidt MF. Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *J Neurophysiol.* 2003; 90:2884–2899. [PubMed: 12878713]
- Catchpole, CK.; Slater, PBJ. *Bird Song: Biological Themes and Variations.* Cambridge University Press; 2008.
- Cousillas H, George I, Mathelier M, Richard JP, Henry L, Hausberger M. Social experience influences the development of a central auditory area. *Naturwissenschaften.* 2006; 93:588–596. [PubMed: 16932952]
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomed Res.* 1996; 29:162–173.
- Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L. Functional neuroimaging of speech perception in infants. *Science.* 2002; 298:2013–2015. [PubMed: 12471265]
- Doupe AJ. Song- and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J Neurosci.* 1997; 17:1147–1167. [PubMed: 8994068]
- Doupe AJ, Solis MM. Song- and order-selective neurons develop in the songbird anterior forebrain during vocal learning. *J Neurobiol.* 1997; 33:694–709. [PubMed: 9369467]
- Espino GG, Lewis C, Rosenfield DB, Helekar SA. Modulation of theta/alpha frequency profiles of slow auditory-evoked responses in the songbird zebra finch. *Neuroscience.* 2003; 122:521–529. [PubMed: 14614916]
- Grace JA, Amin N, Singh NC, Theunissen FE. Selectivity for conspecific song in the zebra finch auditory forebrain. *J Neurophysiol.* 2003; 89:472–487. [PubMed: 12522195]

- Hauber ME, Cassey P, Woolley SM, Theunissen FE. Neurophysiological response selectivity for conspecific songs over synthetic sounds in the auditory forebrain of non-singing female songbirds. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. 2007; 193:765–774. [PubMed: 17503053]
- Jusczyk, PW. The discovery of spoken language. MIT Press; 1997.
- Lauay C, Gerlach NM, Adkins-Regan E, Devoogd TJ. Female zebra finches require early song exposure to prefer high quality song as adults. *Anim Behav*. 2004; 68:1249–1255.
- Marler P, Tamura M. Culturally Transmitted Patterns of Vocal Behavior in Sparrows. *Science*. 1964; 146:1483–1486. [PubMed: 14208581]
- Maul, KK.; Voss, HU.; Parra, LC.; Salgado-Commisariat, D.; Tchernichovski, O.; Helekar, SA. 2007 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience; 2007. How developmental song learning affects auditory responses in the songbird; p. Online
- Mills DL, Prat C, Zangl R, Stager CL, Neville HJ, Werker JF. Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *J Cogn Neurosci*. 2004; 16:1452–1464. [PubMed: 15509390]
- Murray MM, Wylie GR, Higgins BA, Javitt DC, Schroeder CE, Foxe JJ. The spatiotemporal dynamics of illusory contour processing: combined high-density electrical mapping, source analysis, and functional magnetic resonance imaging. *J Neurosci*. 2002; 22:5055–5073. [PubMed: 12077201]
- Nick TA, Konishi M. Neural auditory selectivity develops in parallel with song. *J Neurobiol*. 2005; 62:469–481. [PubMed: 15616963]
- O'Loughlen AL, Rothstein SI. Female preference for the songs of older males and the maintenance of dialects in brown-headed cowbirds (*Molothrus ater*). *Behavioral Ecology and Sociobiology*. 2003; 53:102–109.
- Periaswamy S, Farid H. Elastic registration in the presence of intensity variations. *IEEE Trans Med Imaging*. 2003; 22:865–874. [PubMed: 12906240]
- Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol*. 1999; 110:1842–1857. [PubMed: 10576479]
- Poirier C, Vellema M, Verhoye M, Van Meir V, Wild JM, Balthazart J, Van Der Linden A. A three-dimensional MRI atlas of the zebra finch brain in stereotaxic coordinates. *Neuroimage*. 2008; 41:1–6. [PubMed: 18358743]
- Pollen DA, Trachtenberg MC. Some problems of occipital alpha block in man. *Brain Res*. 1972; 41:303–314. [PubMed: 5038330]
- Pytte CL, Suthers RA. A bird's own song contributes to conspecific song perception. *Neuroreport*. 1999; 10:1773–1778. [PubMed: 10501573]
- Riebel K. Early exposure leads to repeatable preferences for male song in female zebra finches. *Proc R Soc Lond B Biol Sci*. 2000; 267:2553–2558.
- Rivera-Gaxiola M, Klarman L, Garcia-Sierra A, Kuhl PK. Neural patterns to speech and vocabulary growth in American infants. *Neuroreport*. 2005; 16:495–498. [PubMed: 15770158]
- Shank SS, Margoliash D. Sleep and sensorimotor integration during early vocal learning in a songbird. *Nature*. 2009; 458:73–77. [PubMed: 19079238]
- Tchernichovski O, Lints T, Mitra PP, Nottebohm F. Vocal imitation in zebra finches is inversely related to model abundance. *Proc Natl Acad Sci U S A*. 1999; 96:12901–12904. [PubMed: 10536020]
- Terpstra NJ, Bolhuis JJ, Riebel K, van der Burg JM, den Boer-Visser AM. Localized brain activation specific to auditory memory in a female songbird. *J Comp Neurol*. 2006; 494:784–791. [PubMed: 16374807]
- Tsao FM, Liu HM, Kuhl PK. Speech perception in infancy predicts language development in the second year of life: a longitudinal study. *Child Dev*. 2004; 75:1067–1084. [PubMed: 15260865]
- Van Meir V, Boumans T, De Groof G, Van Audekerke J, Smolders A, Scheunders P, Sijbers J, Verhoye M, Balthazart J, Van der Linden A. Spatiotemporal properties of the BOLD response in the songbirds' auditory circuit during a variety of listening tasks. *Neuroimage*. 2005; 25:1242–1255. [PubMed: 15850742]
- Van Meir, V.; Boumans, T.; De Groof, G.; Verhoye, M.; Van Audekerke, J.; Van der Linden, A. Program No. 129.8 2003 Abstract Viewer/Itinerary Planner. Washington, DC: Society for

- Neuroscience; 2003. Functional Magnetic Resonance Imaging Of The Songbird Brain When Listening To Songs.
- Volman S. Development of neural selectivity for birdsong during vocal learning. *J Neurosci.* 1993; 13:4737–4747. [PubMed: 8229196]
- Voss HU, Tabelow K, Polzehl J, Tchernichovski O, Maul KK, Salgado-Commissariat D, Ballon D, Helekar SA. Functional MRI of the zebra finch brain during song stimulation suggests a lateralized response topography. *Proc Natl Acad Sci U S A.* 2007; 104:10667–10672. [PubMed: 17563380]
- Voss HU, Watts R, Ulug AM, Ballon D. Fiber tracking in the cervical spine and inferior brain regions with reversed gradient diffusion tensor imaging. *MRI.* 2006; 24:231–239.
- Werker JF, Gilbert JH, Humphrey K, Tees RC. Developmental aspects of cross-language speech perception. *Child Dev.* 1981; 52:349–355. [PubMed: 7238150]
- Worsley KJ, Marrett S, Neelin P, Friston KJ, Evans AC. A Unified Statistical Approach for Determining Significant Signals in Images of Cerebral Activation. *Human brain mapping.* 1996; 4:58–73. [PubMed: 20408186]
- Wunderlich JL, Cone-Wesson BK, Shepherd R. Maturation of the cortical auditory evoked potential in infants and young children. *Hear Res.* 2006; 212:185–202. [PubMed: 16459037]

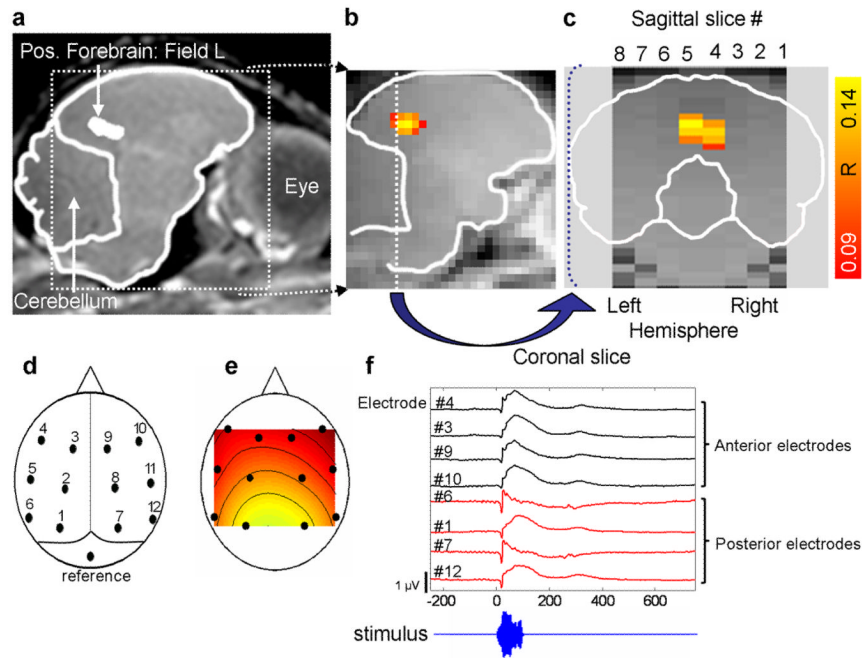


Figure 1.

Auditory responses using fMRI and ERP. **(a)** Anatomical MRI image of live, sedated zebra finch. **(b)** Functional MRI image with area of activation (voxels with activity above significance threshold) shown here for colony males averaged over birds and stimuli. In the statistical parametric maps, each colored voxel indicates stimulus-related activity (during stimulus) that is significantly above baseline activity (silence) ($p < 0.005$, multiple test corrected); color scale indicates correlation coefficient between signal intensity and the stimulus indicating function. Yellow voxel indicates greater significance of activation (higher correlation coefficient) and red indicates lower significance. **(c)** Functional image from coronal view shows medial activation (sagittal slices 4 and 5) in the posterior forebrain. FMRI analysis focused on the size of the active area in this region of interest. **(d)** Epidural electrode cap. **(e)** Spatial distribution of first principal component of electrode potentials indicates that the main variability (74%) across time, stimuli and birds in the event related potentials is in the rostro-caudal direction. **(f)** Example of event related potential (ERP) waveforms in the anterior and posterior electrodes in response to a short conspecific call.

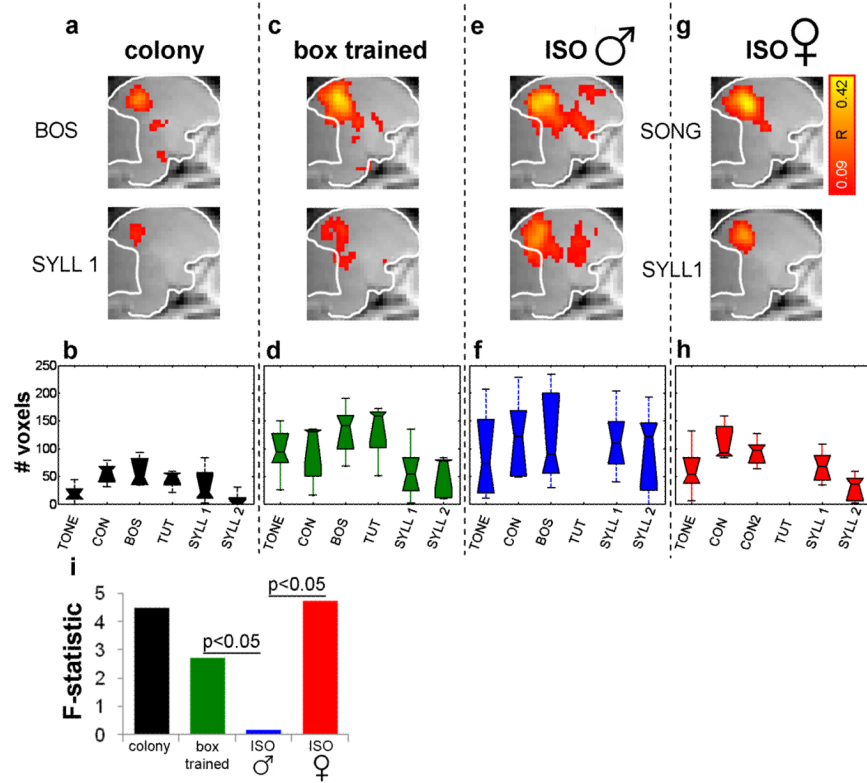


Figure 2.

Differences in BOLD activation to different stimuli. **(a)** BOLD activation maps for bird's own song (BOS) and repeated syllable (SYLL1) in colony males shows stronger activation to BOS. Each map shows average activity in left and right hemisphere 0.5 mm from midline (sagittal slices 4 & 5); color scale shows correlation coefficient (see caption of Fig. 1). **(b)** Box plot summarizing BOLD activation across stimuli in colony males. Area of activation varies significantly across stimuli with greater activation to songs (ANOVA $p < 0.01$). Stimuli: 2 kHz pure tone, conspecific song (CON), bird's own song (BOS), tutor song (TUT), song syllables (SYLL 1, SYLL 2). All conspecific songs were produced by unfamiliar, colony raised birds. **(c-d)** Same as A-B for box-trained males ($p < 0.05$). **(e-f)** Same for isolated males, except that there is no tutor song ($p = 0.9$). **(g-h)** Same for isolate females ($p < 0.05$) with an additional conspecific song (CON2) to balance cross-group comparisons. **(i)** The F statistics of ANOVAs can be used to compare response patterns across groups. Bootstrap comparisons of response patterns show significance between box-trained and isolated males, and between isolated males and isolated females ($p < 0.05$, each comparison).

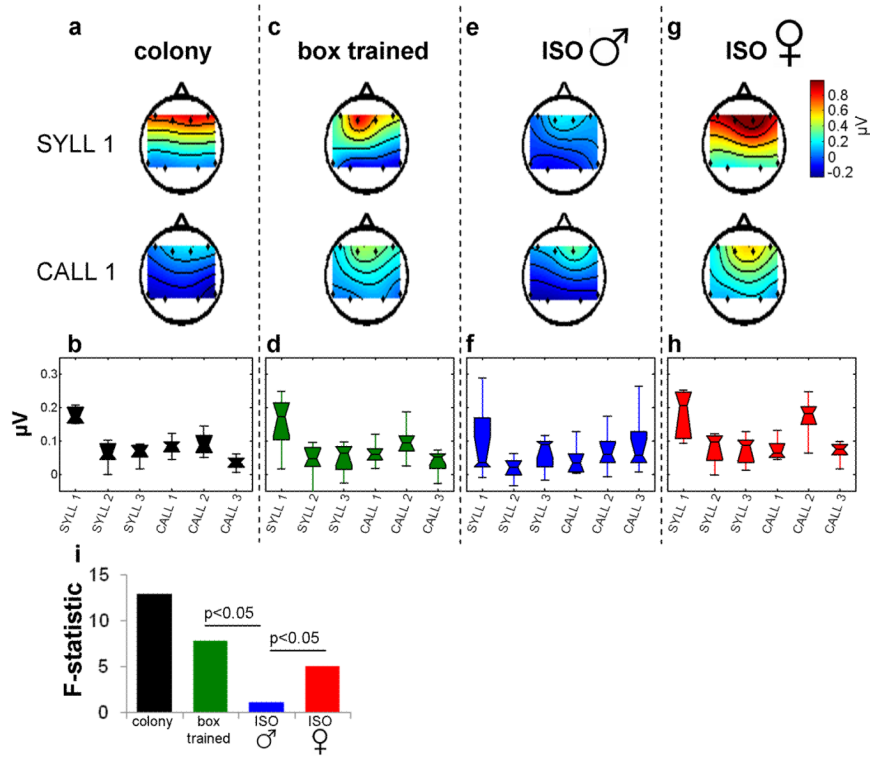


Figure 3. Differences in ERP responses to different stimuli. **(a)** In colony males rostro-caudal potentials at stimulus onset (0-40 ms) vary by stimulus in two example stimuli. Color scale indicates amplitude of evoked potential at stimulus onset (ERP values averaged over 10-40 ms post-stimulus time-window). **(b)** Box plot summarizing ERP rostro-caudal potentials across stimuli in colony males. Potentials vary significantly across stimuli (ANOVA $p < 0.01$). Stimuli: syllables (SYLL 1, SYLL 2, SYLL 3) and female calls (CALL 1, CALL 2, CALL 3). **(c-b)** Same as A-B for box-trained males ($p < 0.05$). **(e-f)** Same for isolated males ($p = 0.7$), **(g-h)** Same for female isolates ($p < 0.05$). **(i)** The F statistics comparisons of response patterns show significance between box-trained and isolated males, and between isolated males and isolated females ($p < 0.05$, each comparison).

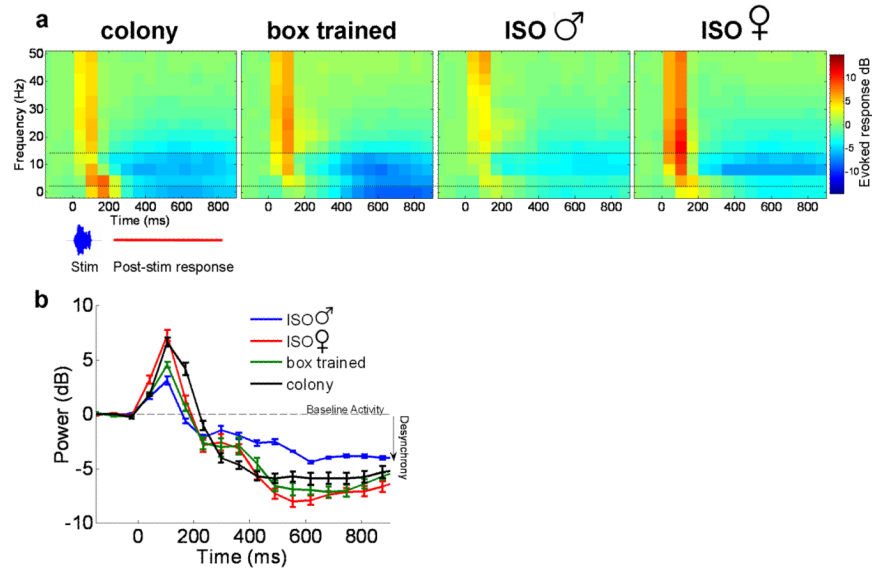


Figure 4.

Effect of song learning and gender on post-stimulus synchrony. **(a)** Trial-averaged spectrograms of the EEG across electrodes. Spectrograms, averaged across stimuli, show decreased power in the 4-16 Hz frequency band (between dotted lines) approximately 200-800 ms after stimulus onset. This effect is weak in isolated males compared to all other groups. **(b)** Mean power in the 4-16Hz band: there is significantly less decrease in power (desynchrony) in the isolate males compared to all other groups (2 way ANOVA: group, $p < 0.01$; stimuli, $p = 0.7$; interaction, $p = 0.9$). Reduced post-stimulus desynchrony in isolated males may suggest that this group is less engaged in listening to and processing biologically relevant sounds.