Dynamics of frequency and amplitude modulations in vocalizations produced by eastern towhees, *Pipilo erythrophthalmus*

Brian S. Nelson

Department of Biology, Indiana University, Bloomington, Indiana, 47405-3700

(Received 26 August 2003; revised 26 December 2003; accepted 29 December 2003)

Eastern towhees, *Pipilo erythrophthalmus* (Emberizidae, Passeriformes), appear to estimate source-SPL using spectral or temporal variables. Nevertheless, vocalizations are dynamic by nature and it remains unclear whether subjects pay attention to correlations between discrete variables or pay attention to the dynamics that these variables summarize. Sine functions are used to describe coarse (slow, <10 Hz) frequency and amplitude modulations in towhee calls and correlations between variables are identified. Towhee calls are also finely (rapidly, >400 Hz) modulated in both amplitude and frequency. Fine amplitude and frequency modulations correlate well (occur in phase) over relatively low fundamental frequencies (<3.5 kHz) and tend to have greater amplitudes and frequencies over these same frequencies. Modulations and correlations between modulations might exist due to stable dynamic interactions that occur within and between the physical forces that function to produce modulations in vocalizations. Results support the hypothesis that towhees communicate within separate sound frequency channels defined to each side of ~3.5 kHz.

I. INTRODUCTION

Vocal production and hearing mechanisms are dynamic by nature and individuals may often exchange information using a currency that is also dynamic in nature. Dynamics in vocalizations are commonly described as modulations, however, some dynamics are commonly distinguished from modulations. For example, songbird vocalizations and syllables within vocalizations are often sequenced and produced with variable repetition rates (e.g., Podos, 2001; Draganoiu et al., 2002). Repetition rates are commonly distinguished from modulations since there often may be no obvious acoustic carrier being modulated between syllables. Thus, repetition and modulation will be treated as separate entities in this paper even though auditory systems might often react similarly in response to both types of periodic signals [Fig. 1(a)].

In contrast with repetition rate, coarse (slow) amplitude modulations (AM) and coarse frequency modulations (FM; Fig. 1(b)) are ubiquitous in songbird vocalizations and have long been defined as modulations (Greenewalt, 1968; Stein, 1968). Gardner et al. (2001) demonstrated that relatively coarse changes in subbronchial air pressure and labial tension (elasticity) are often sufficient to produce many of the coarse modulations [Fig. 1(b)] that are evident in bird vocalizations. Nevertheless, few studies have attempted to quantify coarse modulation functions in recorded vocalizations and, as a consequence, little is known about the modulation functions that need to be described by mathematical models of vocal production mechanisms (e.g., Laje et al., 2002; Mindlin et al., 2003; Elemans et al., 2003). More specifically, little is known about the modulation functions that are thought to be derived from relatively slow (>~20 ms) oscillations in subbronchial air pressure or tension on the medial labia (ML) and lateral labia (LL).

Vocalizations and syllables within vocalizations commonly vary in sound frequency, duration, or amplitude [Fig. 1(c)] and correlated measurements between these variables can also often resemble modulating functions (e.g., Nelson, 2000). Functions derived across independently produced vocalizations are not normally defined as modulations but are likely to be produced by the same physical forces (e.g., subbronchial air pressure and labial elasticity). As a consequence, descriptions of the forces that function to produce intrasyllable modulations are likely to also describe the forces that function to produce intersyllable modulations. Fine (rapid) modulations have also long been recognized in bird vocalizations (Greenewalt, 1968; Stein, 1968). Unlike coarse modulations, however, fine modulations occur, by definition, over relatively short periods of time (typically <10 ms) and may occur due to rapid oscillations in sublabial air pressure or tension on the ML (as opposed to relatively slow oscillations in these same forces). Acoustical interactions between two relatively high frequency carriers (e.g., >1.5 kHz) can also produce fine amplitude modulations in vocalizations (Nowicki and Capranica, 1986a) but are not expected to generate fine frequency modulations. Fine modulations give a vocalization a “rough” acoustic quality and are not easily described using discrete acoustic variables [Fig. 1(d)]. Instead, fine modulations are usually quantified in the frequency domain using sidebands in spectra (e.g., Nowicki and Capranica, 1986b; Banta Lavenex, 1999). Sidebands do not, however, differentiate fine amplitude modulations from fine frequency modulations (Beecher, 1988) and, as a consequence, little is known about the fine modulations that occur in songbird vocalizations.
calls (N=315) that were recorded and analyzed previously (Nelson, 2000). Subjects were captured in a mist net and then placed in a vinyl-coated steel wire (2.4-mm diameter, 1.15×2.25 cm mesh) cage measuring 20×20×20 cm and elevated 1.5–2.5 m above the ground in a small clearing or fire-lane located within each subject’s territory. Recordings were obtained with one or two microphones placed parallel to the ground and 100 cm from a perch placed in the center of the cage (Bruel & Kjaer 4188 microphone with a Bruel & Kjaer 2236 sound level meter functioning as a preamplifier, or Bruel & Kjaer 4189 microphone with Bruel & Kjaer 2671 and Radio Design Labs STM-2 preamplifiers). Microphone outputs were recorded using a digital audio tape recorder (Sony TCD-D3 or HHB PortaDAT) or microcomputer (16 bits; 44.1 or 48 kHz sampling rate).

Subjects are not represented equally in this sample. However, no single individual’s calls dominated the sample and the trends that are described are not produced by the calls of any single individual. Furthermore, general trends are evident in calls produced by each individual. Subjects could rarely be recaptured, thus differences observed between individuals are not described. Correlations between variables described previously were found to be consistent across individuals (Nelson, 2000) and thus the main goal of this research is to describe modulations in vocalizations produced by individuals throughout this population.

Error in SPL calibration can influence both the correlations and the modulation functions that are described. As a result, 1-kHz, 94-dB (re: 20 μPa) reference calibration tones (Bruel & Kjaer 4231) were recorded before or after calls were recorded and calls were subsequently adjusted to appropriate absolute SPLs before analyses were performed (Nelson, 2000). In most cases, these adjustments were accurate (±0.5 dB) but in some cases slightly larger errors (±2 dB) may have occurred due to various factors. These errors may have introduced noise into measurements but do not produce the correlations nor the modulations that are described.

Time-frequency spectrograms of calls were produced using a 34.8-ms Hanning window, 46.4-ms FFT frame, and 98% overlap between each successive 46.4-ms frame. Frequency modulations were not directly quantified using spectra, however, spectrograms were inspected to ensure that there were no problems with recordings. In addition, spectra extracted from time-frequency spectrograms were used to quantify the magnitudes of sidebands in recorded calls (see below). Each vocalization was next high-pass filtered (30 dB/oct at 200 Hz) and cropped above a 65-dB rms threshold (re: 20 μPa; 0.035 Pa) to minimize background noise (see below). This cropping procedure also functioned to define the times at which calls both started and ended.

Fine modulation in the time domain was measured over the duration of each call using an autocorrelation algorithm (Fig. 2). The autocorrelation algorithm performs best when executed over at least 3 cycles and, as a consequence, frequency estimates were obtained using a sliding 1.5-ms Hanning window so that each window would include at least 3 cycles at 2 kHz. The short-term fundamental frequency within each window was first approximated as the inverse of
Coarse \((< \sim 10 \text{ Hz})\) and fine \((\sim 400 \text{ Hz})\) modulations are likely to be perceived differently and produced by distinct physical forces. As a consequence, coarse AM and FM were quantified by applying a 10-ms sliding average (boxcar) smoothing procedure to fine AM and FM waveforms. This 10-ms window preserves modulation frequencies below 10 Hz \((\sim 0.15 \text{ dB})\) but attenuates \((\sim 40 \text{ dB})\) modulation frequencies lower than 400 Hz.

Coarse FM is sinusoidal and the Levenberg–Marquardt best-fit algorithm was used to fit a sine function to each coarse FM waveform,

\[
y_0 = A \cdot \sin(f \cdot 2\pi \cdot t + \varphi),
\]

where \(y_0\) is fundamental frequency, \(A\) is amplitude (frequency excursion in Hz), \(t\) is time, \(f\) is coarse FM frequency, and \(\varphi\) is phase. Because fine modulations \((\sim 400 \text{ Hz})\) are often strong (see below), modulation frequencies \((f)\) were limited to below 10 Hz.

Coarse AM is difficult to summarize because calls have variable duration. Calls were therefore normalized to one second and coarse AM was approximated using a Gaussian function

\[
y_0 = A \cdot e^{-\left(\frac{t - t_0}{W}\right)^2},
\]

where \(y_0\) is call amplitude \((\text{Pa})\), \(A\) is the amplitude of the coarse AM function, \(t\) is time, \(t_0\) is the time of peak amplitude, and \(W\) is the width of the function.

There is no clear rationale for why coarse AM dynamics should be characterized using a single Gaussian function. Thus, onset and offset ramps were also described using best-fit sine functions [Eq. (2), where \(y_0\) is call amplitude, \(A\) is modulation amplitude, \(t\) is time, \(f\) is coarse AM frequency, and \(\varphi\) is phase]. Onset and offset ramps merge together when calls reach maximum amplitude (Fig. 4) and, as a consequence, sine functions were fit to points falling either before or after the center \((t_0)\) of a second Gaussian function that was fit to each original time-amplitude waveform. Offsets ramps were referenced to the end of each vocalization and are thus presented using a negative time scale.

First-order sidebands in towhee calls corresponding with \(f \pm m\) typically have far greater magnitude \((>15 \text{ dB})\) than second-order sidebands corresponding with \(f \pm 2m\) (see Sec. III), indicating that fine frequency and amplitude modulations in towhee calls are both coherent and sinusoidal. Preliminary observations of sidebands in time-frequency spectrograms also indicated that fine modulations typically occur with frequencies ranging from between 450 and 550 Hz. I therefore filtered each finely modulated waveform (bandpass FIR filter; 400–600 Hz; \(\sim 30 \text{ dB/oct}\)) and estimated the frequency (rate) of fine AM and FM imposed upon each call’s coarse FM carrier (fundamental) by performing a second autocorrelation analysis on each filtered waveform. The
FIG. 3. Methods provide for an accurate estimate of sound frequency and amplitude over a relatively short period of time (1.5 ms) but produce biases in estimates of fine AM and FM amplitude. (a) FM amplitude losses expected as FM index increases (FM index=FM excursion/FM frequency). (b) FM amplitude losses expected as a function of FM frequency. (c) FM amplitude losses expected as a function of fundamental frequency (coarse FM). (d) AM amplitude losses expected as a function of fundamental frequency (coarse FM). (e) AM amplitude losses expected as a function of AM index increases (AM index=AM amplitude/carryer amplitude). (f) AM amplitude losses expected as a function of FM frequency. (g) AM amplitude losses expected as a function of AM frequency. (h) AM amplitude losses expected as a function of FM frequency.

autocorrelation algorithm used for this analysis was identical to the algorithm described above with the exception that I used a longer 10-ms time frame so that each frame would encompass at least 4 cycles at 400 Hz.

Fine AM and FM should have a consistent phase relationship if both are produced by a single modulating structure or pair of coupled (integrated) structures. Correlation coefficients were therefore calculated over 8-ms time frames obtained from fine AM and FM waveforms and correlation coefficients were used to describe the phase relationship between fine AM and FM over the duration of each call. An 8-ms time frame was employed so that each frame would encompass at least 4 cycles at 500 Hz.

Calls have variable duration and measurements that vary over the duration of a call are difficult to summarize. As a result, measurements were averaged within 50-Hz sound frequency bins and are summarized as a function of fundamental frequency. Measurements were obtained using Igor Pro (Wavemetrics, Inc.).

III. RESULTS

A. Time-series measurements

Eastern towhee calls are highly variable in acoustic structure (Nelson, 2000). For example, call (a) in Fig. 4 has a relatively low bandwidth and low overall amplitude, while calls (b) and (c) have greater amplitude and bandwidth. Calls that span a large bandwidth and include fundamental frequencies above and below 3.5 kHz are especially interesting since fine amplitude and frequency modulations often have different amplitudes, frequencies, and phase relationships to each side of 3.5 kHz [Fig. 4(c); see below]. Measurements that vary over time (Fig. 4) are difficult to summarize because calls vary in overall duration (Nelson, 2000). As a consequence, measurements obtained directly from time-amplitude waveforms are not summarized.

B. Coarse AM and FM

Calls are summarized as a function of fundamental frequency (coarse FM) and Fig. 5(a) illustrates how measurements were averaged within 50-Hz bins. Residuals from best-fit operations are small (<~30 Hz) in comparison with the overall bandwidth of coarse FM [~1-2 kHz; Fig. 5(d)]. Small residual values demonstrate that coarse FM is well characterized by a sine function. Nevertheless, residuals from this analysis tend to vary predictably with fundamental frequency [e.g., Fig. 5(a)] and independent sine functions fit to sound frequencies to each side of 3.5 kHz typically produce far smaller residuals (<5 Hz vs. <30 Hz). In particular, positive residuals near ~3 and ~4 kHz and negative residuals near ~2.5 and ~3.5 kHz decrease substantially when curve fit operations are limited to fundamental frequencies to each side of 3.5 kHz.

Sine equation phase estimates obtained from curve fit operations are summarized in Fig. 5(f) and sine equation frequency estimates are summarized in Fig. 5(g). There is a significant and relatively strong negative correlation between frequency and phase measurements [Fig. 5(e); $r = -0.68, F = 921, N = 1067, p < 0.001$] suggesting that these two variables may either be controlled by the same physical forces or by forces that are coupled during phonation.

Reasonably accurate fits to coarse AM were obtained using a Gaussian function as illustrated by relatively small residuals from best-fit operations [mean±s.d.; $A = 0.44 ± 0.09, width = 0.2 ± 0.1$, and $t_0 = 0.39 ± 0.22$; Figs. 6(a) and (b)]. Nevertheless, smaller residuals were obtained when call
onset and offset ramps were fit using independent sine functions [Figs. 6(c) and (e)]. There is a significant and moderately strong negative correlation between onset phase and onset frequency [Fig. 6(d); \( r = -0.5, F = 343, N = 1067, p = 0.003 \)]. Sine equation phase estimates are not summarized for offset ramps because Gaussian functions were used to determine the center (\( t_0 \)) of each rms SPL waveform and thus the phase of each sinusoidal offset ramp. Onset durations are typically longer than offset durations and AM offset frequencies [Fig. 6(e)] tend to be higher than AM onset frequencies as a result.

Residuals from sine functions fit to both coarse AM and FM onset ramps tend to vary together over time and irregular (nonsinusoidal) changes in both fundamental frequency and rms amplitude often track each other over time. Such changes demonstrate the AM and FM are likely to be produced by the same physical forces. Nevertheless, coarse AM and coarse FM functions do not correlate well across calls because AM and FM functions tend to define different modulation frequencies over different periods of time (AM phase vs. FM phase, \( r = 0.007, n = 1067 \); AM frequency vs. FM frequency, \( r = 0.2, n = 1067 \)).

C. Fine AM and FM

Fine AM and fine FM should vary together and have a strong positive or negative phase relationship if the physical forces that function to produce these modulations are coupled (integrated) during phonation. The phase relationship between fine AM and fine FM decreases from positive to negative (from in phase to out of phase) rapidly between 3 and 4 kHz and the phase relationship is significantly higher.
at 3.25 kHz than at 3.75 kHz \(t\)-test, \(p<0.001, n=1270\); Fig. 7(a)]. Measurements are not derived directly from time-series data. Nevertheless, phase estimates may not be independent and thus statistical comparisons drawn as a function of coarse FM frequency need to be viewed with caution.

Sidebands in time-frequency spectra decrease in magnitude above \(\sim 3.5\) kHz [Figs. 7(b) and (c)] and an especially rapid decrease in sideband magnitude often occurs in the first upper sideband \((f + m)\) near \(\sim 3.5\) kHz [Fig. 7(b)]. The first lower sideband \((f - m)\) does not decrease in magnitude as rapidly as the first upper sideband [Fig. 7(e)]. These differences suggest that large changes in upper sideband \((f + m)\) magnitude may occur, in part, as a result of phase changes that occur between AM and FM signals as the fundamental frequency in each call \(f\) approaches and exceeds \(\sim 3.5\) kHz [Figs. 7(d) and (e)].

Decreases in the phase relationship between AM and FM often occur rapidly in individual calls and this decrease often corresponds with a rapid (-2 ms) decrease in both the frequency and amplitude of AM near \(\sim 3.5\) kHz [e.g., Fig. 4(c)]. An abrupt change in AM-FM phase might be expected if towhees produce fundamental frequencies to each side of \(\sim 3.5\) kHz using different sides of their bipartite syrinx or if modulations correlate positively only when calls are increasing in amplitude (see below).

Frequency and amplitude of fine AM and FM vary as a function of fundamental frequency (Fig. 8). For example, fine AM (rms mean=497 Hz) and fine FM (rms mean=456 Hz) both increase in frequency as a function of fundamental frequency (coarse FM) between 2 and 3 kHz. Both frequencies then decrease above \(\sim 3\) kHz and frequency measurements are significantly higher at 3.25 kHz than at 3.75 kHz \(t\)-test, \(p<0.001, n=1270\); Figs. 8(a) and (b)]. Fine AM (rms mean=3.1%) and fine FM (frequency excursion; rms mean=59 Hz) also both increase in amplitude as a function of fundamental frequency between 2 and 3 kHz. Both amplitudes then decrease above \(\sim 3\) kHz and amplitude measurements are significantly higher at 3.25 kHz than at 3.75 kHz \(t\)-test, \(p<0.001, n=1270\); Figs. 8(c) and (d)]. Modulation measurements may not be independent and thus statistical comparisons that are drawn as a function of fundamental frequency should again be viewed with caution. Nevertheless, decreases in both fine AM and FM amplitude are expected to contribute to the rapid decrease in sideband mag-
FIG. 7. Upper and lower sideband magnitudes appear to be determined, in part, by the phase relationships that exist between fine amplitude and frequency modulations. (a) Average phase relationship (r) between fine AM and fine FM as a function of fundamental frequency. (b) Average magnitude of the first upper sideband (f + m) as a function of fundamental frequency. (c) Average magnitude of the first lower sideband (f - m) as a function of fundamental frequency. (d) Relationship between phase measurements presented in (a) and upper sideband magnitudes displayed in (b). (e) Relationship between phase measurements presented in (a) and lower sideband magnitudes displayed in (c). All averages are calculated within 50-Hz bins. Error lines depict one standard deviation and the gray line in (a) illustrates the number of average values obtain within each 50-Hz bin.

magnitude that occurs near $\sim 3.5$ kHz [Fig. 7(b)].

Calls increase and then decrease in amplitude (Fig. 6) and positive increases in amplitude tend to occur over relatively low sound frequencies as a result. For example, Gaussian functions fit to amplitude envelopes tend to increase in amplitude over relatively low fundamental frequencies but decrease in amplitude over higher frequencies [Fig. 9(a)]. These factors make it difficult to evaluate whether differences in fine AM and FM (Fig. 8) correlate best with changes in fundamental frequency or amplitude. For example, amplitude of fine FM is not only significantly correlated with fundamental frequency [$r = -0.38, N = 212,721, p < 0.001$, Fig. 8(b)] but is also significantly correlated with the rate to which amplitude is changing within each Gaussian function [$\Delta dB/S$ ms; $r = 0.37, N = 212,721, p < 0.001$, Fig. 9(b)]. The same correlation with amplitude ($\Delta dB/S$ ms) is weaker when calculated using coarse amplitude envelopes due to small fluctuations in amplitude that can be attributed to several factors (Fig. 4; $r = 0.20, N = 212,721, p < 0.001$). Nevertheless, the forces that regulate amplitude are likely to be the same forces that regulate fundamental frequency (e.g., Laje et al., 2002). As a consequence, it may not be surprising that fundamental frequency and changes in amplitude both correlate with fine modulation amplitude and frequency.

If fine AM and FM are both produced by the same physical forces, or by forces that are coupled (integrated) during phonation, then similar modulation frequencies should be observed whenever there is a strong phase relationship between fine AM and FM (see above). A positive correlation exists between frequency of fine AM and frequency of fine FM [$r = 0.31, n = 212,721$; Fig. 10(a)]. Similarly, a positive correlation exits between amplitude of fine AM and amplitude (frequency excursion) of fine FM [$r = 0.63, n = 212,721$; Fig. 10(b)]. All correlations are highly significant ($p < 0.001$) and correlations become stronger as the phase relationship between AM and FM increases [Figs. 10(c) and (d)]. For example, the correlation coefficient (r) between frequency of fine AM and frequency of fine FM increased from 0.31 to 0.83 ($n = 69,538$) when measurements were limited to those obtained when the correlation coefficient describing the phase relationship between AM and FM was greater than 0.75. Similarly, the same correlation coefficient between AM and FM decreased from 0.31 to 0.17 ($n = 54,667$) when measurements were limited to those obtained when correlation coefficients (r) describing the same phase relationship were negative [Figs. 10(e) and (f)]. A weak but distinct function (slope) appears to define this correlation when AM and FM are out of phase with one another [$r < 0$; Figs. 10(e) and (f)].

Different modulation frequencies and phases are often observed to each side of $\sim 3.5$ kHz (Figs. 7 and 8) and thus similar differences in correlation coefficients should occur as a function of fundamental frequency. These results are not illustrated, however, plots produced from measurements obtained over different sound frequency ranges resemble those shown in Fig. 10. For example, figures produced using all
D. Contributions to sidebands from AM and FM

Sidebands in time-frequency spectrograms often increase in width or diverge into distinct spectral peaks whenever fine AM and FM diverge in modulation frequency. These changes demonstrate that AM and FM both contribute to sideband structure but are difficult to quantify due to time-frequency tradeoffs (Beecher, 1988). Modulations in towhee calls are sinusoidal and, as a consequence, contributions to sidebands that are due to both AM and FM can be approximated using estimates of modulation indices derived from time-series data (Fig. 4). An average raw FM index of 0.21 ± 0.06 (mean ± sd; \( N = 1067 \)) was obtained over the duration of each call and an adjusted FM index of 0.26 was obtained after compensating for amplitude losses that are due to methodology [i.e., use of a sliding window (Fig. 3) and calculations of rms amplitude versus peak-to-peak amplitude]. Sidebands predicted using an FM index of 0.26 and the Bessel function have a relatively narrow bandwidth (Fig. 11; \( f ± m = -17.8 \text{ dB}, \quad f ± 2m = -41.5 \text{ dB}, \quad f ± 3m = -68.8 \text{ dB} \)).

To allow for comparisons with frequency modulations, AM indices were calculated from time intervals in recorded calls where FM indices equaled 0.21 ± 0.005 (mean ± sd; 6320 indices from 878 calls). An average raw AM index of 0.056 ± 0.002 was obtained under these conditions and an adjusted AM index of 0.21 was obtained after compensating for amplitude losses that are due to methodology (see...
A. Coarse modulation functions

It is important to understand the control parameters that are required to generate vocalizations (Suthers and Margoliash, 2002) and it is, therefore, interesting that complex birdsong can be modeled using only a few vocal control parameters (Laje et al., 2002; Mindlin et al., 2003; Elemans et al., 2003). Birdsong is often highly stereotyped and, as a result, vocal control parameters are not likely to be described well using continuous functions. Instead, control parameters might best be viewed as discontinuous functions and discontinuities might, in practice, define combinations of parameter values that are rarely observed during phonation. For example, in a relatively extreme case, if a given species only produces upward frequency sweeps in its repertoire, then combinations of parameter values that are required for producing upward frequency sweeps might be physically “disconnected” from stable parameter values that are required for producing downward frequency sweeps.

Dynamic systems are rarely stable over a broad range of conditions and thus correlations between modulation functions [Figs. 5(c), 6(d), and 9 (Nelson, 2000)] might be produced by stable interactions that occur between vocal control parameters. Constraints on vocal production mechanisms have been described previously (e.g., Podols, 2001; Zollinger and Suthers, 2001) and constraints are expected to impose limitations on available parameter values. Nevertheless, different combinations of parameter values may not have equal stability within a dynamic system and, thus, relatively unstable combinations of parameter values may have implications similar to those proposed for constraints. The difference is that a constraint defines what subjects cannot physically accomplish while variation in stability describes what subjects normally accomplish as a result of interactions between several dynamic physical forces.

Towhee calls are highly variable in acoustic structure. In addition, several discrete measurements obtained from time-frequency spectrograms correlate well with source-SPL and towhees appear to use at least one of these variables when estimating source-SPL (Nelson, 2000). Nevertheless, vocalizations are dynamic by nature and it remains unclear whether towhees pay attention to these variables (e.g., lowest frequency, bandwidth, etc.) or the dynamics that are summarized by these variables. Coarse modulation frequency decreases with modulation phase (Fig. 5) and thus subjects may, for example, more likely pay attention to modulation frequency or phase when estimating source-SPL. This hypothesis has not yet been tested, however lowest frequency measurements and other variables associated with coarse FM functions correlate well with source-SPL. Towhees reliably estimate source-SPL after hearing only the first, low frequency (~3 kHz) half of synthetic call stimuli (Nelson, 2002) and these results suggest that subjects either (1) assess coarse modulation frequency and phase or (2) somehow resolve discrete variables when estimating source-SPL.
B. Fine modulations

Fine amplitude and frequency modulations in towhee calls are strong, coherent, and sinusoidal. Fine amplitude and frequency modulations in human speech are often described as “shimmer” and “jitter” and these two measurements are often associated with laryngeal pathologies (e.g., Hansen et al., 1998; Jiang et al., 2003; Vieira et al., 2002). Unlike measurements of shimmer and jitter in speech, however, modulations in towhee calls do not correlate well with fundamental frequency (do not occur on a period by period basis). Instead, modulation frequency remains relatively constant even as fundamental frequency varies between ~2 and 4.5 kHz (i.e., modulation periods are at least three times longer than fundamental periods and are sometimes ten times longer).

AM can be produced whenever there is a direct or indirect (multiplicative) interaction between two signals (e.g., Nowicki and Capranica, 1986b). Unlike AM, however, FM can be produced only when there is a direct integrative interaction between two signals. It is therefore surprising that towhees are able to vary fundamental frequency (~2 to 4.5 kHz) and, at the same time, generate a coherent modulating signal that does not vary directly with fundamental frequency.

Physical models that might account for the modulation patterns that exist in towhee calls have yet to be described. Towhees, like other songbirds, have a duplex syrinx and are able to produce two independent sounds simultaneously in their vocalizations (Greenwalt, 1968; Stein, 1968; Suthers, 1990). Nowicki and Capranica (1986b) proposed that acoustic interactions between these two sounds might sometimes produce AM. Nevertheless, acoustic interactions are not expected to produce FM. Physical mechanisms for producing fine AM have subsequently been proposed for mammals (e.g., Neubauer et al., 2001; Brown et al., 2003) and budgerigars, Melopsittacus undulatus (Banta Lavenex, 1999). Nevertheless, neither of these models incorporates an explicit mechanism for generating FM and neither model has sought to account for how two independent sinusoidal FM signals might integrate to produce a single vocalization with coherent modulations.

Rapid oscillations (>400 Hz) in subbronchial air pressure (pressure below the medial tympaniform membranes, MTM) might generate the fine frequency and amplitude modulation patterns that are evident in towhee calls (Fletcher, 1988; Laje et al., 2002). However, if the medial labia (ML) and lateral labia (LL) are the primary sound generating structures in the songbird syrinx (Goller and Larsen, 2002, 1997), then rapid changes in sublabial air pressure or labial elasticity might also be responsible for producing modulations in vocalizations (as well as modulations in air pressure measurements obtained below the MTM). The MTM are expected to vibrate in a sinusoidal manner as long as they do not contact the opposite wall of each bronchus (Fletcher, 1988) and are ideally positioned to produce fluctuations in both sublabial air pressure and labial tension (Fee, 2002). Vibrations of the MTM might therefore also potentially generate the frequency and amplitude modulation patterns that are evident in towhee calls. Fee (2002) demonstrated that there is an interaction (integration) between the MTM and ML, however, it remains unclear whether vibrations of the MTM might produce the fine sinusoidal modulations that are evident in towhee calls or whether the MTM might simply vibrate together with the ML to produce a stronger fundamental frequency.

Other songbird species such as Carolina chickadees, Poecile carolinensis, and zebra finches, Taeniopygia guttata, also produce fine frequency modulations in their vocalizations in addition to fine amplitude modulations (not illustrated). These modulations do not necessarily correlate with fundamental frequencies in these vocalizations and, as a consequence, results presented for the eastern towhee may have relatively broad implications. For example, amplitude modulations in chickadee “Dee” syllables are believed to carry important information (e.g., Nowicki, 1989; Freeberg et al., 2003). Furthermore, zebra finches and birds in general appear to resolve fine temporal patterns surprisingly well and may often be able to resolve fine envelope patterns (modulations) with periods as short as 1–2 ms (Dooling et al., 2002).

Pigeons, Columba livia, and European starlings, Sturnus vulgaris, are able to discriminate low levels of sinusoidal FM (Bräucker and Schwartzkoff, 1986; Langemann and Klump, 1992). High modulation rates are expected to produce spectral cues that span more than one analysis channel (i.e., critical band), and thus spectral cues may help to explain how starlings are able to discriminate low levels (<15 Hz) of high frequency FM (640 Hz, Langemann and Klump, 1992). Similar experiments have yet to be conducted with towhees. Nevertheless, both humans and starlings are able to discriminate smaller frequency limens when frequency changes are imposed upon a low frequency carrier (fundamental). Critical bands generally increase in width as sound frequency increases (e.g., Langemann et al., 1995) and FM discrimination may therefore help to explain why towhees tend to modulate sound frequencies below ~3.5 kHz more strongly than higher sound frequencies.

C. Sound frequency channels

Towhees use SPL as an auditory distance cue when played call stimuli produced with sound frequencies below ~3.5 kHz but use a cue that has yet to be identified when played call frequencies above ~3.5 kHz (Nelson, 2002). Sound frequencies below ~3.5 kHz also attenuate more reliably over distance in comparison with higher sound frequencies, suggesting that towhees might use a duplex sound localization “strategy” as a mechanism for improving communication in their natural environment (Nelson, 2003). Given these results, it is interesting that fine AM and FM both decrease in frequency and amplitude above ~3.5 kHz and that the phase relationship between fine AM and FM becomes weak near and above ~3.5 kHz.

Towhees may generate sound frequencies to each side of ~3.5 kHz using separate sides of their bipartite syrinx, as has been observed in other songbird species (e.g., Suthers, 1999), and differences in fine modulation patterns (e.g., phase relationships between fine AM and FM) might exist simply because the left and right sides of the syrinx behave differently.
More specifically, distinct forces operating within the left and right sides of the syrinx might respond differently to a constant driving force (e.g., air pressure), or each side of the syrinx (i.e., each bronchus) might experience a distinct driving force. In fact, residuals from sine functions fit to coarse FM waveforms vary predictably as a function of fundamental frequency and smaller residuals can nearly always be obtained by fitting two separate sine functions to frequencies above or below ~3.5 kHz. As an alternative, the MTM may vibrate more strongly or coherently when amplitude is increasing (i.e., when sublabial air pressure is increasing).

Proximate explanations for these data will require additional investigations into vocal production mechanisms. Nevertheless, sound frequencies to each side of ~3.5 kHz appear to be processed differently by the auditory system (see above) and thus a functional explanation for why subjects often modulate sound frequencies below ~3.5 kHz more strongly than higher sound frequencies in their calls seems to be warranted. Furthermore, it is interesting that towhees should use SPL as a distance cue in the first place since attenuation over distance is often unpredictable (although see Nelson, 2003) and because it has been suggested that some species may vary source-SPL unpredictably (e.g., Brumm and Todt, 2002; Cynx et al., 1998; Manabe et al., 1998), using a compensatory mechanism where subjects are presumably able to counter increases or decreases in subysringeal air pressure (and thus source-SPL) by decreasing or increasing tension on the ML and LL.

Towhees may vary the physical forces that are required for vocal production by coordinating several distinct neural pathways and individuals may call similarly by generating similar motor control programs (Suthers and Margoliash, 2002). Alternatively, individuals might call similarly if only certain combinations of physical forces are dynamically stable within the syrinx and if extraordinary supplemental forces are required to overcome these stable combinations of forces. These two hypotheses are unlikely to be mutually exclusive and thus it remains interesting that towhees do not normally vary source-SPL unpredictably when calling (Nelson, 2000).

One possible explanation for why towhees do not normally vary source-SPL unpredictably may be that towhees are physically unable to simultaneously produce both fine modulations in vocalizations and the supplementary forces that are required to overcome the stable dynamic forces that normally function to generate highly stereotyped coarse modulation patterns. For example, if rapid changes in sublabial air pressure or labial elasticity function to generate fine modulations, then correlative relationships between source-SPL and spectral and temporal dynamics (fine modulations) might be mediated by interactions between the MTM and the ML. Alternatively, if rapid changes in subbronchial air pressure function to generate fine modulations, then these same relationships might be mediated by the (hypothetical) forces that function to vary subbronchial pressure.

That a close relationship between source-SPL and spectro-temporal patterns might be mediated by stable dynamics within the syrinx remains hypothetical. Nevertheless, a stable dynamic relationship between coarse and fine modulation patterns might help to explain why towhees tend to modulate fundamental frequencies below ~3.5 kHz more strongly than higher fundamental frequencies and pay attention to sound frequencies below ~3.5 kHz when using SPL as a distance cue (Nelson, 2002).

Location is often disassociated from other forms of information that birds might convey when vocalizing (e.g., individual identity, signaler “quality,” etc.). Nevertheless, this disassociation likely occurs because researchers often choose to simplify their experiments and does not exclude the possibility that towhees might convey additional information within these same two sound frequency channels. For example, towhees do not simply judge incident-SPL but appear to also estimate source-SPL and thus may perceive a variable comparable to loudness or vocal effort in addition to, or in spite of, distance (e.g., Zahorik and Wightman, 2001). Humans appear to use low frequency cues when judging vocal effort (e.g., Eriksson and Traumuller, 2002) and thus it may not be surprising that towhees also pay attention to relatively low sound frequencies when judging source-SPL.

ACKNOWLEDGMENTS

I thank Roderick Suthers, Troy Smith, Sumitra Raj Dhar, Ellen Ketterson, William Rowland, and Erica Goss for reading earlier versions of the manuscript and Gabriel Beckers for numerous discussions. Support was provided by NSF Grant No. BIR-9413220 to the Center for the Integrative Study of Animal Behavior at Indiana University and by NIH Grant No. NS-29467 to R. Suthers. Archbold Biological Station provided access to research facilities.


