The Wilson Journal of Ornithology 120(2):256–267, 2008

SONG VARIATION IN BUFF-BREASTED FLYCATCHERS (*EMPIDONAX FULVIFRONS*)

M. ROSS LEIN[1]

ABSTRACT.—I examined song variation within and among 23 individual Buff-breasted Flycatchers (*Empidonax fulvifrons*) recorded in the Chiricahua and Huachuca mountains of Arizona in 1999. I recorded two distinct song types from each individual during intense pre-dawn singing. I used both spectrographic cross correlation (SPCC) of entire songs and discriminant function analysis (DFA) of temporal and frequency measurements to examine whether songs were individually distinctive, and whether songs differed between the two localities. Similarity values of pairs of songs from SPCC were significantly greater for within-male than for between-male comparisons for both song types. Mean similarity values for the two song types did not overlap between these comparison categories. Similarity values between songs of pairs of males from the same mountain range were not greater than for comparisons between pairs of males from different ranges. All temporal and frequency measures for both song types varied significantly more among than within individuals. DFA of principal component scores derived from these measures assigned 85% of Type 1 and 86% of Type 2 songs to the correct individual. Only three frequency variables measured from Type 1 songs differed significantly between birds from the two mountain ranges. DFA assigned only 61% of songs of either type to the correct mountain range, not significantly greater than expected by chance. Thus, both techniques demonstrate significant individual distinctiveness in songs of this species, and neither suggests any geographic structuring of song variation between the two mountain ranges. However, SPCC is considerably more efficient and has greater potential to assign unknown recordings to known individuals correctly, and to detect recordings of ''new'' individuals not included in the reference sample. *Received 24 April 2007. Accepted 19 July 2007.*

Most literature on song variation deals with oscine birds (Passeriformes, suborder Passeri) (Lovell and Lein 2004). Songs are learned in most or all oscines (Kroodsma 1996), which is a major factor generating song variation at individual, population, and geographic scales. Suboscine birds (suborder Tyranni) appear to show lower levels of song variation. The absence of learning during song ontogeny has been demonstrated for a few species of North America tyrant flycatchers (Tyrannidae) (Kroodsma 1996), and this has been generalized to all suboscines based primarily (and circumstantially) on the limited variation in their songs relative to oscines.

Early studies used song variation in suboscine passerines to help resolve the taxonomic status of populations (e.g., Stein 1963; Johnson 1963, 1980; Lanyon 1978) or described how song variants were used in communication systems (Smith 1969, 1970, 1988). None of these studies focused explicitly on the nature of song variation among individuals within local populations. Song variation among individuals, albeit minor, is apparent in published spectrograms of a variety of suboscine species (e.g., Stein 1963, Payne and Budde 1979, Kroodsma 1984). However, in contrast to the large, often qualitative, differences which may be obvious in spectrograms of different individuals (e.g., Borror 1960) or populations (e.g., Baptista and King 1980) of oscines, examination of suboscine song variation requires quantitative analysis of sufficiently large samples of recordings. The few analyses conducted to date have demonstrated, for example, that songs of Alder Flycatchers (*Empidonax alnorum*) (Lovell and Lein 2004) and Acadian Flycatchers (*E. virescens*) (Wiley 2005) are individually distinctive, and that songs of the endangered Southwestern Willow Flycatcher (*E. traillii extimus*) differ significantly from those of a neighboring subspecies (*E. t. adastus*) (Sedgwick 2001). However, more studies are needed to demonstrate the generality of such patterns, especially for other groups of suboscines.

Most quantitative studies of song variation have used univariate or multivariate analyses of temporal and frequency characters measured from spectrograms. Characters are usually selected on the basis of their ease of measurement with the assumption that if enough characters are measured, the analysis will cap-

[1] Department of Biological Sciences, University of Calgary, Calgary, AB T2N 1N4, Canada; e-mail: mrlein@ucalgary.ca

ture the features that may be important. However, it is often difficult to select characters that can be measured in an objective manner for all songs, and characters are often extremely general (e.g., number of notes, maximum frequency, etc.), capturing little of the ''structure'' of the vocal signal.

Digital processing of acoustic signals provides alternative methods for characterizing variation among songs. Peter Marler and coworkers (Clark et al. 1987) first used spectrographic cross correlation (SPCC) to characterize variation in notes of songs of Swamp Sparrows (*Melospiza georgiana*). SPCC compares two digital spectrograms at successive offsets on the time axis and calculates normalized covariance (ranging from −1 to 1) at each offset. The maximum covariance is used as a measure of similarity between the two signals (Baker and Logue 2003). SPCC has the advantage that distribution of sound energy in both frequency and time are compared in a ''holistic'' manner, probably capturing more of the relevant features of the song than would a small number of quantitative measurements (Khanna et al. 1997). Several bioacoustical analysis software packages include routines that automate such measurements. The CORMAT routine of the SIGNAL digital signal analysis software (Engineering Design, Berkeley, CA, USA) can run SPCC on samples of up to 200 songs simultaneously.

The Buff-breasted Flycatcher (*Empidonax fulvifrons*) is the least known of the 11 species of this genus that breed in Canada or the United States (Bowers and Dunning 1994). It breeds in montane forests through Mexico south to Guatemala, El Salvador, and Honduras (AOU 1998), but has been studied almost exclusively in extreme southeastern Arizona where small numbers of birds, possibly less than 100 pairs in total (Martin 1997), inhabit several isolated mountain ranges. I recorded songs of male Buff-breasted Flycatchers during 1999 in the Chiricahua Mountains and the Huachuca Mountains, the two ranges with the largest populations of this species in Arizona.

Male Buff-breasted Flycatchers use two distinctive song types, which I designate as Type 1 and Type 2, during pre-dawn singing and during strong daytime singing (Fig. 1). Bowers and Dunning (1994) published spectrograms of Type 1 songs, which they describe as ''chee-lick'', but did not describe or illustrate the Type 2 song. Type 2 songs are rarely used during the sporadic daytime singing typical of the breeding cycle following pairing. They are similar to Type 1 songs in sound, but the higher frequency is obvious to the ear.

My objective was to characterize patterns of song variation within and among individuals in the Chiricahua and Huachuca mountains. I conducted both spectrographic cross-correlation and multivariate analyses of song characters for the same sample of songs to answer two questions. First, are songs of Buff-breasted Flycatchers individually distinctive? Second, are there detectable differences in songs of birds between the two mountain ranges?

METHODS

Study Area and Field Methods.—I recorded 23 individual male Buff-breasted Flycatchers at seven sites in two isolated mountain ranges in southeastern Arizona between 7 and 25 June 1999. Specific recording localities in the Chiricahua Mountains (number of individuals in parentheses) included Cave Creek (3), Pinery Canyon (2), Rucker Canyon (6), and West Turkey Creek (2). Localities in the Huachuca Mountains were Carr Canyon (3), Garden Canyon (1), and Sawmill Canyon (6). Males at West Turkey Creek and Sawmill Canyon were recorded on multiple dates whereas individuals at other sites were recorded only on one or two dates. Buff-breasted Flycatchers breed in open woodlands dominated by pines (*Pinus* spp.), live oaks (*Quercus* spp.), and alligator juniper (*Juniperus deppeana*). Sites within the same mountain range were separated by maximum distances of 22 km (Chiricahua) and 7 km (Huachuca), whereas breeding populations in the two mountain ranges are separated by more than 100 km of unsuitable desert habitat. Birds arrived on territories in mid-April and paired immediately. However, nesting did not begin until late May. Most pairs were involved in nest-building or egg-laying during the period of recording.

Recordings of pre-dawn singing were made between 0436 and 0518 hrs MST using a Sony TCD-D10 ProII DAT recorder and a Telinga ProII parabolic microphone, or a Sony TC-D5 ProII cassette recorder and a Audio-

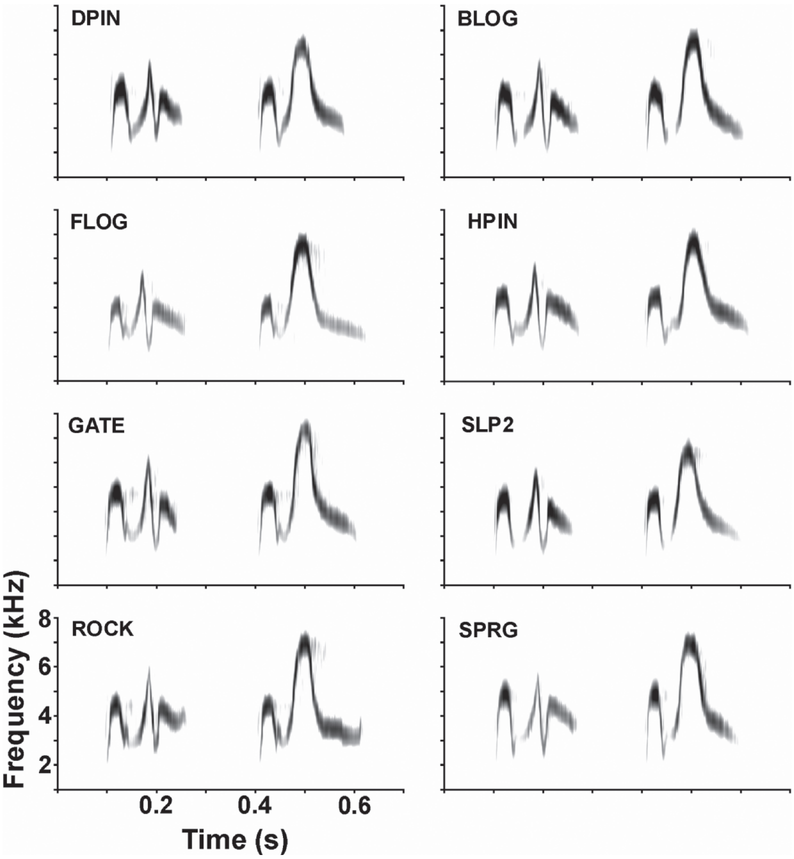

FIG. 1. Audiospectograms of songs of Buff-breasted Flycatchers. Each panel contains examples of Type 1 and Type 2 songs for an individual male. The four males in the left column are from the Chiricahua Mountains and those in the right column are from the Huachuca Mountains, Arizona.

Tech AT815a "shotgun" microphone. I used recordings of pre-dawn singing because each male sang strongly for 15–20 min prior to dawn each day, even during phases of the breeding cycle when daytime singing was rare and sporadic. Males used relatively low perches during pre-dawn singing and could be approached quite closely; recordings made at this time were of high quality.

Males at West Turkey Creek and Sawmill Canyon were banded and color-marked for individual identification; males at other sites were unmarked. However, such markings were not visible before the end of pre-dawn singing. Misidentification of males was highly improbable because there were few individuals at any site and males on neighboring territories could be heard clearly while recording pre-dawn singing of individuals.

Processing of Recordings.—Recordings were acquired as digital files using RTSD Version 2.0 bioacoustical analysis software (Engineering Design, Berkeley, CA, USA) with a sample rate of 20,000 Hz and 16-bit amplitude resolution. Analog signals were filtered during acquisition with a Krohn-Hite Model 3550 filter to avoid aliasing. I selected single recordings for each individual for subsequent sampling based on high signal-to-noise ratios and adequate numbers of songs. Eight examples of each song type were extracted as individual sound files from each recording using SIGNAL 4.0 bioacoustical analysis software (Engineering Design, Berkeley, CA, USA). Many recordings contained hundreds of songs and I selected samples in quasi-random manner. Only one song of each type was selected from each 20-sec segment of the recording with a high signal-to-noise ratio. One or more segments were skipped between each segment from which songs were selected if more than eight segments met the criterion. I selected songs within each 20-sec segment with minimal amounts of reverberation or background noise. SIGNAL 4.0 was used to band-pass filter sampled songs, removing noise outside the frequency range of interest (1,500–7,000 Hz for Type 1 songs and 1,500–8,000 Hz for Type 2 songs), and to normalize all sampled songs to the same root-mean-square amplitude.

Use of single recordings for each individual may underestimate the amount of variation in songs within individuals, but was necessitated by the limited number of recordings available for many males. However, analyses of songs of individual Alder Flycatchers (Lovell and Lein 2004), Willow Flycatchers (*E. traillii*) (MRL, unpubl. data), and Dusky Flycatchers (*E. oberholseri*) (Stehelin 2005) demonstrate little variation among recordings across the breeding season. They also show that patterns of variability within and among individuals are similar when comparisons are made using either single recordings of individuals or multiple recordings made on different dates.

Analysis Using Spectrographic Cross Correlation.—Spectrographic cross correlation (SPCC) was conducted using the CORMAT routine in SIGNAL 4.0. Similarity values calculated by CORMAT are somewhat sensitive to the exact parameters used to generate spectrograms because of the trade-off between time resolution and frequency resolution inherent in the Fast Fourier Transform procedure. Preliminary analyses indicated that spectrograms with a Hanning window (WINDOW = HANN), 64-point transforms (XFTLEN = 64), and 500 steps (XFTSTP = 500) provided maximum similarity values. These produced "wide-band" spectrograms with time and frequency resolutions of 3.2 ms and 312.5 Hz, respectively. Use of a fixed number of steps is justified because CORMAT adds zero-amplitude segments to short signals to bring all component signals to the length of the longest signal. However, because of the different durations of the two types of songs, transform intervals and overlaps varied slightly between song types (0.45 ms and 86% for Type 1 songs, 0.54 ms and 83% for Type 2 songs).

CORMAT produces a lower triangular half-matrix in which each value is the peak cross correlation value for a pair of signals. Values for comparisons of a signal with itself and reciprocal comparisons between each pair of signals are omitted. SPCC of 184 signals resulted in a half-matrix of 16,836 ($n(n - 1)/2$) comparisons of different pairs of songs of each type. I wrote a Fortran program to calculate mean values for the 28 comparisons among the eight songs sampled from each male (within-male similarity) and the 64 comparisons of songs between each pair of males (between-male similarity), creating a 23 × 23 triangular half-matrix for each song type. Mean similarity values in these two matrices were averaged to create a third matrix containing a mean index of similarity for each male-male comparison.

Normal parametric tests could not be used to evaluate differences between within-male and between-male similarity values because the similarity values in the matrices were not independent observations (each song of each male was used in multiple comparisons). Randomization tests (Manly 1997) using Resampling Stats software (Blank et al. 2001) and 1,000 iterations tested whether the difference between mean within-male and mean between-male similarity values in each matrix was greater than expected by chance.

I divided the between-male similarity values into those involving comparisons within a single range and those involving comparisons between ranges to examine differences in songs between the two mountain ranges. I tested for significant differences between within-range and between-range similarity values using Mantel tests (Sokal and Rohlf 1995). Similarity values were converted to dissimilarity values by subtracting each from 1. The half-matrices of dissimilarity values were compared with a design half-matrix containing a 0 in each cell representing a within-range comparison and a 1 in each cell representing a between-range comparison.

Analysis Using Multivariate Comparisons.—A series of temporal and frequency variables was measured or calculated for the same songs used in the SPCC analysis (8 examples of each song type for each of 23 individuals). A representative sample of songs was examined to ascertain which temporal

TABLE 1. Time and frequency variables measured or calculated from audiospectrograms of Type 1 and Type 2 songs of Buff-breasted Flycatchers. Variables indicated with * were measured only for Type 1 songs.

Code	Variable (units)
DURN	Duration of entire song (ms)
PK1T	Duration from start of song of first frequency "peak" (ms)
PK2T	Duration from start of song of second frequency "peak" (ms)
PK3T*	Duration from start of song of third frequency "peak" (ms)
VALT	Duration from start of song of frequency "valley" (ms)
STRF	Frequency at start of song (Hz)
ENDF	Frequency at end of song (Hz)
PK1F	Maximum frequency at first frequency "peak" (Hz)
PK2F	Maximum frequency at second frequency "peak" (Hz)
PK3F*	Maximum frequency at third frequency "peak" (Hz)
VALF	Minimum frequency at frequency "valley" (Hz)
P2VFR	Frequency difference between first frequency peak and frequency valley (Hz)
P1P2FR	Frequency difference between first and second frequency peaks (Hz)
P1PN	Relative time of first frequency peak (PK1T ÷ DURN)
P2PN	Relative time of second frequency peak (PK2T ÷ DURN)
P3PN*	Relative time of third frequency peak (PK3T ÷ DURN)
VPN	Relative time of frequency valley (VALT ÷ DURN)

and frequency characteristics could be measured with objectivity and were repeatable prior to final measurements. Sound in both song types is modulated up and down in frequency repeatedly, producing a series of "peaks" and "valleys" on the spectrogram. The first peak is identical between song types for each individual. Type 1 songs have one more modulation than Type 2 songs resulting in one additional peak and valley. The final set of variables measured included the duration of the song, and the times and frequencies of the peaks (3 for Type 1 songs, 2 for Type 2 songs) and valleys. No measurements were made for the first valley in Type 1 songs because some individuals exhibited a break in sound production at this point with no clear inflection in frequency.

Temporal and frequency measurements were made in a semi-automated fashion using custom programs written in the SIGNAL command language. Points were measured on a spectral contour generated from the spectrogram. The spectral contour tracks the frequency with the maximum sound energy at a given time. Sound amplitude increases gradually at the start of songs and fades out gradually at the end, and the apparent locations of these points can be shifted on spectrograms by modifying the parameters controlling spectrogram intensity. Consequently, start and end times of the song were defined arbitrarily as the points at which the spectral contour exceeded (start) or fell below (end) an amplitude threshold of 20 dB below the maximum amplitude of the spectrogram. This procedure resulted in a consistent approximation of these times because the amplitude of all songs was normalized prior to analysis. Frequencies at the start and end of the song were extracted from the spectral contour at these times. Similarly, times of peaks and valleys were defined as the points of local maxima (peaks) or minima (valleys) of frequency in the spectral contour. Five temporal and six frequency variables were measured for Type 1 songs (Table 1). Only four temporal and five frequency variables were measured for Type 2 songs because these songs have one fewer frequency peak. Additional variables (6 for Type 1 songs, 5 for Type 2 songs) were calculated from the measured variables (Table 1).

I calculated coefficients of variation (CV) for each variable for each song type to quantify the magnitude of variability. I calculated within-male coefficients of variation (CV_w) to measure variation within a single recording and among-male coefficients of variation (CV_a) from the variable means from each male. A one-way ANOVA was conducted on each variable to compare within-male and among-male variation.

I conducted principal components analyses (PCA) on the data sets for each song type be-

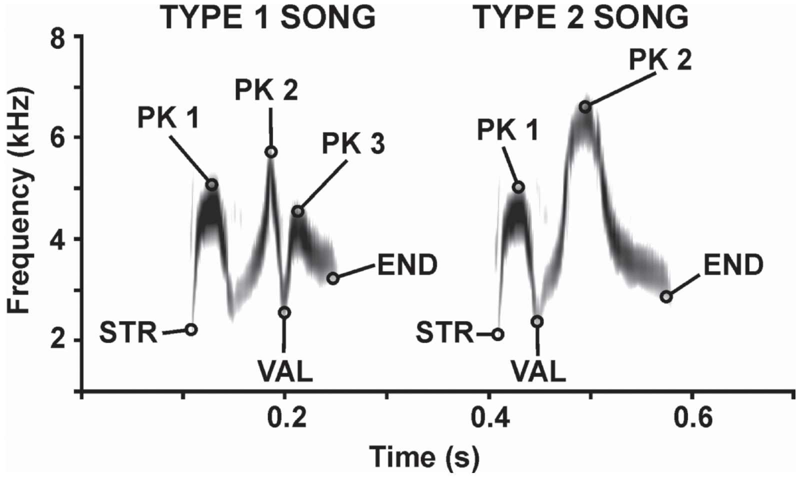

FIG. 2. Audiospectograms of Type 1 and Type 2 songs of Buff-breasted Flycatchers indicating the reference points for temporal and frequency measurements (STR = Start; PK = Peak; VAL = Valley).

cause some of the variables were highly correlated with one another. This produced a smaller number of uncorrelated variables (PC scores) for each song, which were entered into a discriminant function analysis (DFA) to examine whether songs of different individuals could be distinguished. Results of jack-knifed classifications, in which each song was assigned to an individual using discriminant functions calculated from all songs in the data set except the one being classified, are reported as percentages of songs assigned correctly. This is a conservative estimate of the power of the classification procedure (Manly 1994).

I conducted a similar analysis to test whether songs of individuals from the two mountain ranges could be reliably distinguished. This analysis used mean values of variables for the eight songs of each type for each individual to maintain sample independence.

SYSTAT 10.2 was used for parametric statistical analyses. Probability plots were used to check variables for an approximate fit to a normal distribution before parametric statistical tests were applied. I report mean values ± SD and use a criterion of $\alpha < 0.05$ for statistical significance.

RESULTS

Visual examination of spectrograms revealed small differences between songs of different individuals (Fig. 1). These differences were reflected in the time and frequency variables measured (Fig. 2), and were consistent from song to song within a single recording and across recordings made over the course of a season (M. R. Lein, unpubl. data). However, these differences are not detectable by ear in the field or on recordings.

Mean similarity values between pairs of songs, calculated using SPCC (Table 2), were significantly higher for comparisons of songs within individuals than for comparisons between individual males for Type 1 songs (randomization test, $P < 0.001$), Type 2 songs (randomization test, $P < 0.001$), and average

TABLE 2. Similarity values calculated by spectrographic cross correlation for comparison of songs within and between individual Buff-breasted Flycatchers.

	Mean similarity value ± SD (Range)	
	Within-male comparisons (n = 23)	Between-male comparisons (n = 253)
Song Type 1	0.88 ± 0.03 (0.84–0.93)	0.65 ± 0.10 (0.38–0.86)
Song Type 2	0.84 ± 0.04 (0.79–0.93)	0.57 ± 0.10 (0.33–0.79)
Average	0.86 ± 0.02 (0.82–0.91)	0.61 ± 0.08 (0.36–0.81)

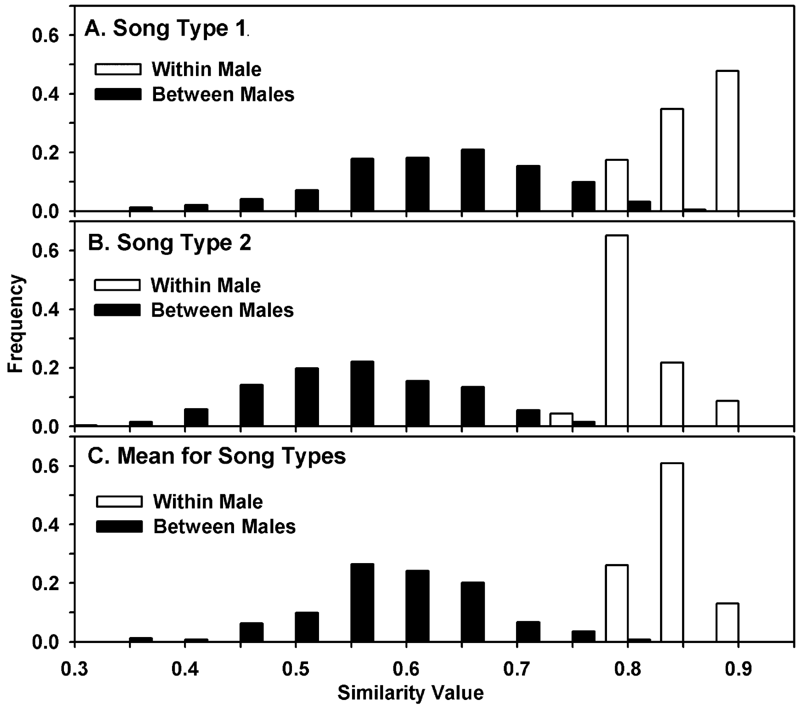

FIG. 3. Similarity values from spectrographic cross correlation of different songs from individual male Buff-breasted Flycatchers (within-male) or of songs from different pairs of males (between-male). A. Song Type 1. B. Song Type 2. C. Average for song types.

similarity values (randomization test, $P < 0.001$). Both within-male and between-male values had considerable variation (Table 2, Fig. 3) with minor overlap in the ranges for both Type 1 and Type 2 songs. However, when similarity values for the two song types were averaged for each comparison (i.e., for each individual male or for each pair of males), there was no overlap in the ranges (Table 2). The maximum average similarity value for between-male comparisons was 0.81 while the minimum value for within-male comparisons was 0.82.

There were no significant differences in similarity values for comparisons between males within a single mountain range and comparisons between males from different mountain ranges (Table 3) for Type 1 songs (Mantel test, $g = 0.714$, $P = 0.20$), Type 2 songs ($g = -0.271$, $P = 0.45$) or average similarity values ($g = 0.295$, $P = 0.34$). The ranges of similarity values for the two types of comparisons overlapped completely (Fig. 4).

Univariate analyses indicated that all variables measured or calculated for individual songs of both types varied significantly more among individuals than within individuals (Tables 4, 5; one-way ANOVAs, all $P <$

TABLE 3. Similarity values calculated by spectrographic cross correlation for comparison of songs between individual Buff-breasted Flycatchers within the same mountain range (either Chiricahua Mountains or Huachuca Mountains) and between mountain ranges.

	Mean similarity value ± SD (Range)	
	Within-range comparisons (n = 123)	Between-range comparisons (n = 130)
Song Type 1	0.65 ± 0.10 (0.38–0.86)	0.64 ± 0.10 (0.39–0.84)
Song Type 2	0.57 ± 0.10 (0.36–0.78)	0.57 ± 0.10 (0.33–0.79)
Average	0.61 ± 0.08 (0.39–0.80)	0.61 ± 0.08 (0.36–0.81)

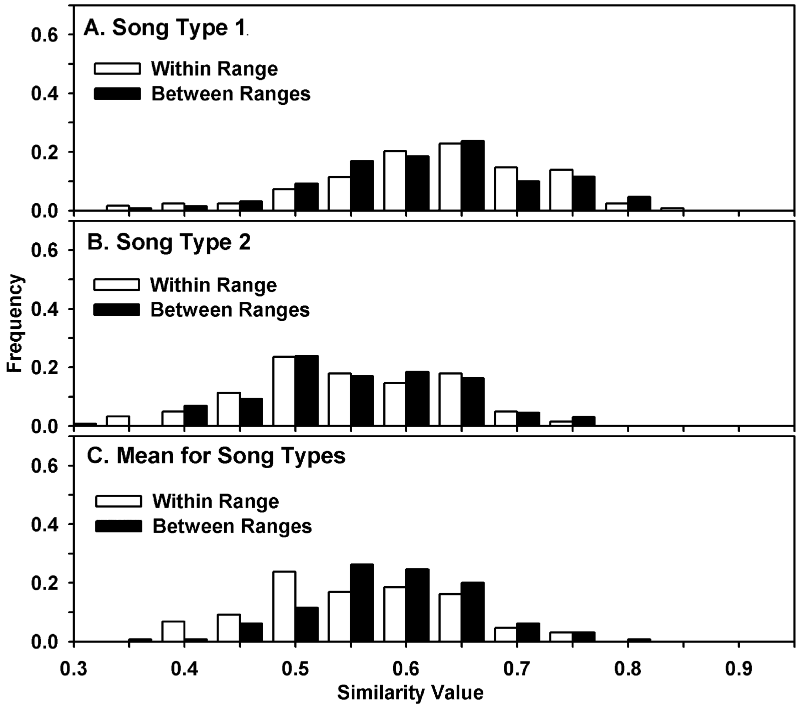

FIG 4. Similarity values from spectrographic cross correlations of songs of pairs of male Buff-breasted Flycatchers from the same mountain range (within-range) or from different mountain ranges (between-range). A. Song Type 1. B. Song Type 2. C. Average for song types.

TABLE 4. Descriptive statistics and coefficients of variation for 17 variables measured or calculated for eight Type 1 songs for each of 23 male Buff-breasted Flycatchers.

Variable	Mean ± SD	Mean CV_w (Range)	CV_a	$F_{22, 161}$[a]
DURN	150.6 ± 11.4	3.0 (1.4–5.5)	7.6	40.0
PK1T	16.7 ± 2.9	11.9 (6.1–20.1)	17.2	9.4
PK2T	81.2 ± 6.7	2.7 (0.9–5.3)	8.3	57.1
PK3T	107.4 ± 7.0	2.8 (1.1–4.7)	6.5	34.9
VALT	94.7 ± 6.7	2.5 (1.2–4.0)	7.1	54.0
STRF	2825.4 ± 374.4	9.3 (3.3–16.6)	13.3	7.1
ENDF	3383.9 ± 267.8	3.1 (1.0–6.9)	7.9	40.4
PK1F	4469.7 ± 172.0	1.0 (0.4–2.3)	3.8	96.6
PK2F	5292.7 ± 288.1	1.3 (0.2–2.8)	5.4	112.9
PK3F	4089.8 ± 166.8	1.2 (0.3–2.5)	4.1	74.3
VALF	2548.5 ± 264.9	2.0 (0.8–3.9)	10.4	187.6
P2VFR	2744.3 ± 371.0	2.7 (0.9–4.1)	13.5	179.9
P1P2FR	823.0 ± 265.8	10.6 (3.2–20.2)	32.3	70.0
P1PN	11.1 ± 2.0	11.7 (5.0–21.4)	18.0	12.1
P2PN	54.0 ± 3.8	2.6 (0.9–4.3)	7.0	47.0
P3PN	71.5 ± 4.5	3.0 (1.0–5.6)	6.2	24.7
VPN	63.0 ± 4.3	2.6 (1.3–4.0)	6.8	43.5

[a] *F*-values for ANOVAs comparing within- and among-male variation for each variable; all $P < 0.001$.

TABLE 5. Descriptive statistics and coefficients of variation for 14 variables measured or calculated for eight Type 2 songs for each of 23 male Buff-breasted Flycatchers.

Variable	Mean ± SD	Mean CV_w (Range)	CV_a	$F_{22,\,161}$[a]
DURN	193.2 ± 15.1	2.3 (0.9–3.6)	7.8	78.2
PK1T	18.8 ± 2.7	8.2 (2.9–16.3)	14.3	12.0
PK2T	93.5 ± 7.4	3.7 (1.8–5.4)	8.0	28.5
VALT	37.8 ± 4.9	3.7 (0.9–7.9)	12.8	65.5
STRF	2288.0 ± 236.9	3.8 (1.4–7.9)	10.4	44.1
ENDF	2912.3 ± 292.6	3.6 (1.4–6.8)	10.0	48.0
PK1F	4463.6 ± 168.3	0.8 (0.4–1.9)	3.8	151.5
PK2F	6871.2 ± 390.3	1.4 (0.3–3.2)	5.7	93.8
VALF	2691.8 ± 203.3	3.1 (0.9–6.4)	7.6	32.0
P2VFR	4179.4 ± 379.0	3.2 (1.5–6.3)	9.1	49.8
P1P2FR	2407.6 ± 333.5	4.4 (0.9–9.0)	13.9	54.8
P1PN	9.8 ± 21.5	8.6 (4.0–28.3)	15.7	15.4
P2PN	48.6 ± 4.7	4.0 (2.5–6.1)	9.6	38.6
VPN	19.7 ± 3.0	4.2 (1.5–6.9)	15.4	86.9

[a] F-values for ANOVAs comparing within- and among-male variation for each variable; all $P < 0.001$.

0.001). The ratio of CV_a/mean CV_w was >1 for all variables.

PCA of the variables generated 5 principal components (PCs) with eigenvalues >1.0 explaining 85.3% of the variation in the original variables for Type 1 songs, and 5 PCs with eigenvalues >1.0, explaining 81.9% of the variation in the original variables for Type 2 songs. MANOVAs on the scores of individual songs on the five PCs, conducted as part of DFA, indicated highly significant differences among multivariate means for different individuals for both Type 1 songs ($F_{110,773} = 54.5$, $P < 0.001$) and Type 2 songs ($F_{110,773} = 56.3$, $P < 0.001$). Jack-knifed classifications assigned 156 of 184 Type 1 (84.8%) and 159 of 184 Type 2 songs (86.4%) to the correct individual.

Univariate analyses of variables averaged over the eight songs of each type for each individual indicated three frequency variables measured or calculated for Type 1 songs differed significantly between males from the Chiricahua and Huachuca mountains (Table 6). There was no difference between mountain ranges in other variables for Type 1 songs (all $F_{1,21} \leq 2.4$, all $P \geq 0.137$) or in any variable for Type 2 songs (all $F_{1,21} \leq 2.9$, all $P \geq 0.101$).

PCA of the individual means for the variables generated 5 PCs with eigenvalues >1.0 for each song type explaining 87.4 and 85.6% of the variation in the original variables for Type 1 and Type 2 songs, respectively. MANOVAs on the scores for individual males on the five PCs indicated no significant differences in multivariate means between the two mountain ranges for either Type 1 ($F_{5,17} = 2.1$, $P = 0.115$) or Type 2 songs ($F_{5,17} = 1.1$, $P = 0.419$). Jack-knifed classifications assigned 14 of 23 individuals (60.9%) to the correct mountain range in DFAs for each song type. This does not differ ($\chi^2 = 0.91$, $P = 0.34$) from the frequency expected if individuals were assigned randomly to mountain ranges.

DISCUSSION

Both analyses indicate that both song types of Buff-breasted Flycatchers are individually

TABLE 6. Variables measured or calculated for Type 1 songs of Buff-breasted Flycatchers that showed significant differences between birds from the Chiricahua Mountains and the Huachuca Mountains, Arizona.

Variable	Mean ± SD (Hz) Chiricahua Mtns (n = 13)	Mean ± SD (Hz) Huachuca Mtns (n = 10)	$F_{1,21}$	P
ENDF	3508.0 ± 70.2	3222.6 ± 133.0	10.4	0.004
VALF	2649.1 ± 253.9	2417.7 ± 228.3	5.1	0.034
P2VFR	2612.8 ± 411.7	2915.1 ± 229.7	4.3	0.050

distinctive. SPCC showed almost no overlap in similarity values for within-male and between-male comparisons (Fig. 3), while DFA showed a high level of accuracy in assignment of songs to the correct individual. Most measured or calculated variables showed little variation within individuals (CV_w values <5%, Tables 4, 5). Many different features of songs had high CV_a/mean CV_w ratios (Tables 4, 5) indicating their potential to differentiate among individuals (Robisson et al. 1993, Bee et al. 2001, Vignal et al. 2004).

The ''errors'' in classification of songs in the DFA are also instructive. Nineteen of 28 Type 1 songs assigned to an incorrect individual were associated with ''reciprocal errors'' in which one song of male A was assigned to male B and one song of male B was assigned to male A. Many of the misclassifications were also consistent with the findings of the SPCC analysis. For example, DFA assigned five Type 1 songs of male SPRG (from Sawmill Canyon in the Huachuca Mountains) to male SNAG (from West Turkey Creek in the Chiricahua Mountains), and three Type 1 songs of SNAG to SPRG. This pair had an extremely high between-male similarity value (0.84), indicating the close similarity of their Type 1 songs despite their geographical separation. Examination of the results for Type 2 songs showed similar patterns.

Both techniques failed to demonstrate significant differences in songs of males between the two mountain ranges. There was no separation of between-male similarity values calculated for comparisons within a single mountain range from those for comparisons between ranges (Fig. 4). Three of 31 variables showed significant univariate differences between the two ranges (Table 6). This did not result in a correct classification of songs by DFA that was significantly better than expected by chance. This finding demonstrates the potential danger of basing analyses of song variation on a small number of variables that are often chosen subjectively. Selected variables may not accurately reflect overall patterns of variation.

The absence of differentiation in songs between the two mountain ranges is also apparent from examination of the SPCC results for pairs of males with between-male similarity values that approach those typical of within-male comparisons. Six of the nine between-male similarity values >0.8 for Type 1 songs, and four of six between-male similarity values >0.75 for Type 2 songs, were for comparisons between males from different ranges. There appears to be no relation between song similarity of pairs of males and geographic proximity.

It is reassuring that both analytical techniques produced qualitatively similar results. However, they approach the question of variation differently and, therefore, each is best suited to address different kinds of questions. SPCC is rapid, but provides only a single value for similarity between each pair of songs. It provides no information about the ways in which the two songs are similar or different from each other. Multivariate analysis of a large set of temporal and frequency measurements is time consuming, and has an element of subjectivity in choosing appropriate features for measurement. However, it does permit identification of the acoustic features in which different samples of songs differ.

It has been suggested that DFA may be used to assign individual identities to ''unknown'' recordings (Terry et al. 2001). Discriminant functions generated from songs of known reference individuals are used to classify the unknown songs. These songs will be assigned to the correct individuals if there is sufficient information in the song features used in the DFA. However, this technique has a serious potential drawback. It is poorly-suited to handle ''true unknowns'', that is, songs of ''new'' individuals that are not included in the sample of reference individuals used to generate the discriminant functions. DFA forces assignment of all songs to one of the pre-existing categories (individuals).

SPCC does not have this drawback. ''Unknown'' songs from an individual in the reference sample will be identified correctly. They will have high similarity values with known songs from that individual, and relatively low similarity values with other individuals in the reference sample. Songs of a ''true unknown'' will have low similarity values with all males in the reference sample and may be identified as coming from a ''new'' individual.

SPCC does have a serious potential problem that must be recognized. It works best for relatively simple sounds without silent intervals separating different elements. If two songs had identical elements, but differed in the timing of the intervals separating them, the elements would show progressively less temporal overlap during the duration of the songs. This would result in a low cross correlation value even though the sound elements were identical (Khanna et al. 1997). This problem is probable with songs more complex than those of *Empidonax* flycatchers. However, it is possible to divide complex songs into a series of elements, to conduct SPCCs on the different elements (Nelson et al. 1995), and to combine or average the cross correlation values for the different elements of two songs into a single measure of similarity. This would parallel the procedure used to generate an average similarity value across the two song types of Buff-breasted Flycatchers.

ACKNOWLEDGMENTS

Valerie Haines provided able assistance in the field and made many of the original recordings. Stephen Russell, Josiah and Valere Austin, and Sheridan Stone provided valuable advice or logistical support for the field work. Tara Stehelin acquired the digital sound files from the recordings. Critical comments from Valerie Haines, Sarah Hechtenthal, and Scott Lovell greatly improved the final manuscript. Field work was conducted under permits from the USDA–Coronado National Forest and the U.S. Army–Fort Huachuca Military Reservation. The research was evaluated and approved by the Life and Environmental Sciences Animal Care Committee of the University of Calgary (Protocol 96-006) and was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

AMERICAN ORNITHOLOGISTS' UNION (AOU). 1998. Check-list of North American birds. Seventh Edition. American Ornithologists' Union, Washington, D.C., USA.

BAKER, M. C. AND D. M. LOGUE. 2003. Population differentiation in a complex bird sound: a comparison of three bioacoustical analysis procedures. Ethology 109:223–242.

BAPTISTA, L. F. AND J. R. KING. 1980. Geographical variation in song and song dialects of montane White-crowned Sparrows. Condor 82:267–284.

BEE, M. A., C. E. KOZICH, K. J. BLACKWELL, AND H. C. GERHARDT. 2001. Individual variation in advertisement calls of territorial male green frogs, *Rana clamitans*: implications for individual discrimination. Ethology 107:65–84.

BLANK, S., C. SEITER, AND P. BRUCE. 2001. Resampling Stats add-in for Excel user's guide. Resampling Stats Inc., Arlington, Virginia, USA.

BORROR, D. J. 1960. The analysis of animal sounds. Pages 26–37 *in* Animal sounds and communication (W. E. Lanyon and W. N. Tavolga, Editors). American Institute of Biological Sciences, Washington, D.C., USA.

BOWERS JR., R. K. AND J. B. DUNNING JR. 1994. Buff-breasted Flycatcher (*Empidonax fulvifrons*). The birds of North America. Number 125.

CLARK, C. W., P. MARLER, AND K. BEEMAN. 1987. Quantitative analysis of animal vocal phonology: an application to Swamp Sparrow song. Ethology 76:101–115.

JOHNSON, N. K. 1963. Biosystematics of sibling species of flycatchers in the *Empidonax hammondi-oberholseri-wrightii* complex. University of California Publications in Zoology 66:79–238.

JOHNSON, N. K. 1980. Character variation and evolution of sibling species in the *Empidonax difficilis-flavescens* complex (Aves: Tyrannidae). University of California Publications in Zoology 112:1–153.

KHANNA, H., S. L. L. GAUNT, AND D. A. MCCALLUM. 1997. Digital spectrographic cross-correlation: tests of sensitivity. Bioacoustics 7:209–234.

KROODSMA, D. E. 1984. Songs of the Alder Flycatcher (*Empidonax alnorum*) and Willow Flycatcher (*Empidonax traillii*) are innate. Auk 101:13–24.

KROODSMA, D. E. 1996. Ecology of passerine song development. Pages 3–19 *in* Ecology and evolution of acoustic communication in birds (D. E. Kroodsma and E. H. Miller, Editors). Cornell University Press, Ithaca, New York, USA.

LANYON, W. E. 1978. Revision of the *Myiarchus* flycatchers of South America. Bulletin of the American Museum of Natural History 161:427–628.

LOVELL, S. F. AND M. R. LEIN. 2004. Song variation in a population of Alder Flycatchers. Journal of Field Ornithology. 75:146–151.

MANLY, B. J. F. 1994. Multivariate statistical methods: a primer. Second Edition. Chapman and Hall, London, United Kingdom.

MANLY, B. J. F. 1997. Randomization, bootstrap and Monte Carlo methods in biology. Second Edition. Chapman and Hall, London, United Kingdom.

MARTIN, J. A. 1997. Distribution, abundance, and habitat characteristics of the Buff-breasted Flycatcher in Arizona. Thesis. University of Arizona, Tucson, USA.

NELSON, D. A., P. MARLER, AND A. PALLERONI. 1995. A comparative approach to vocal learning: intraspecific variation in the learning process. Animal Behaviour 50:83–97.

PAYNE, R. B. AND P. BUDDE. 1979. Song differences and map distances in a population of Acadian Flycatchers. Wilson Bulletin 91:29–41.

ROBISSON, P., T. AUBIN, AND J.-C. BRÉMOND. 1993. Individuality in the voice of the Emperor Penguin *Aptenodytes forsteri*: adaptation to a noisy environment. Ethology 94:279–290.

SEDGWICK, J. A. 2001. Geographic variation in the song of Willow Flycatchers: differentiation between *Empidonax traillii adastus* and *E. t. extimus*. Auk 118:366–379.

SMITH, W. J. 1969. Displays of *Sayornis phoebe* (Aves, Tyrannidae). Behaviour 33:283–322.

SMITH, W. J. 1970. Song-like displays in *Sayornis* species. Behaviour 37:64–84.

SMITH, W. J. 1988. Patterned daytime singing of the Eastern Wood-Pewee, *Contopus virens*. Animal Behaviour 36:1111–1123.

SOKOL, R. R. AND F. J. ROHLF. 1995. Biometry. Third Edition. W. H. Freeman and Company, New York, USA.

STEHELIN, T. E. 2005. Variation in syllable structure and singing behaviour in a population of Dusky Flycatchers (*Empidonax oberholseri*). Thesis. University of Calgary, Calgary, Alberta, Canada.

STEIN, R. C. 1963. Isolating mechanisms between populations of Traill's Flycatcher. Proceedings of the American Philosophical Society 107:21–50.

TERRY, A. M. R., P. K. MCGREGOR, AND T. M. PEAKE. 2001. A comparison of some techniques used to assess vocal individuality. Bioacoustics 11:169–188.

VIGNAL, C., N. MATHEVON, AND S. MOTTIN. 2004. Audience drives male songbird response to partner's song. Nature 430:448–451.

WILEY, R. H. 2005. Individuality in songs of Acadian Flycatchers and recognition of neighbours. Animal Behaviour 70:237–247.