

## SINGING BEHAVIOR OF SOUTHWESTERN WILLOW FLYCATCHERS IN ARIZONA

HELEN K. YARD AND BRYAN T. BROWN

**Abstract.** We studied the singing behavior of a small population of Southwestern Willow Flycatchers (*Empidonax traillii extimus*) along the Verde River in central Arizona from May–August, 1996–1998. Our objectives were to document patterns of daily and seasonal song rates and to evaluate sex differences. Most *fitz-bew* songs were given from perches, as only 2% of male songs and 7% of female songs were given in flight. Males were never detected singing from or within 2 m of nests, but females occasionally sang from or near nests. Rates of male *fitz-bew* song appeared greatest during courtship/nest-building; female song rates appeared greatest during incubation. Rates of male *fitz-bew* song during morning and midday appeared similar, but both appeared greater than evening; female song rates also appeared to differ, with morning > midday > evening. Mean rates of *fitz-bew* song for all years, nesting phases, and times combined appeared similar between sexes. Females gave an adjusted 20% of all *fitz-bew* songs detected (four male songs for each female song). Male rates of *creet* song appeared greater during both courtship/nest-building and incubation than during both nestling and post-fledging phases, while female rates did not appear to differ by nesting phase. Male rates of *creet* song appeared greater in morning than evening, while female rates of *creet* song appeared greater during midday than evening. Determination of sex and estimation of population size based solely on singing behavior should be considered unreliable for this subspecies.

**Key Words:** Arizona; behavior; *Empidonax traillii extimus*; riparian; song; Southwestern Willow Flycatcher; Verde River; vocal behavior.

Passerine song functions to attract mates, establish and defend territories, maintain pair bonds, and initiate other breeding phenomena (Catchpole and Slater 1995, Kroodsma and Miller 1996). Song rates in particular can provide information on the status of the singer (Staicer 1989), the presence of potential predators (Preston et al. 1998), and nest-site quality (Hoi-Leitner et al. 1995). Quantitative evaluations of singing behavior and song rates can be useful in addressing management concerns about parameters such as the potential influence of human intrusion (Gutzwiller et al. 1994, 1997) and determining the relative accuracy of auditory survey protocols (McShea and Rappole 1997), yet this information does not exist for most bird species.

The advertising-song performance of the Willow Flycatcher (*Empidonax traillii*) consists of a variable sequence of three vocalizations: *fitz-bew*, a similar *fitz-bew*, and *creet* (Kroodsma 1984, Sedgwick 2000). Where singing behavior and sex have been correlated through observations of color-banded individuals, males produce most songs (Sedgwick 2000). Females have been reported to sing uncommonly (Seutin 1987) to regularly (Sogge 2000), but quantitative data on female song are lacking. The only evidence of daily or seasonal differences in song frequency is qualitative (e.g., McCabe 1991, Sedgwick 2000, Sogge 2000b).

We studied the singing behavior of a small population ( $\leq 20$  individuals) of the endangered Southwestern Willow Flycatcher (*E. t. extimus*)

to document this little-known aspect of its natural history. Our objectives were to describe some aspects of singing behavior, investigate patterns of daily and seasonal song rates, evaluate sex differences in singing behavior, and discuss the implications of these findings with respect to determining sex and estimating breeding population size during field surveys. More information on the singing behavior of Southwestern Willow Flycatchers would be of value because the current protocol for presence-absence surveys relies primarily on vocal detection techniques (Sogge et al. 1997a,b).

### METHODS

Our study was conducted at Camp Verde, central Arizona, along the Verde River (elevation ca. 1070 m), a perennial stream in a broad floodplain. The river at our site exhibited several small braided channels, and discharge from an irrigation ditch periodically caused standing water to occur in low-lying areas. Beaver (*Castor canadensis*) activity resulted in occasional inundation of part of the site. Riparian vegetation was dominated by discrete patches of native gallery forest composed of Fremont cottonwood (*Populus fremontii*) and Goodding willow (*Salix gooddingii*) which formed a discontinuous canopy up to ca. 22 m in height. Dense stands of non-native saltcedar (*Tamarix ramosissima*) up to 8 m in height occurred both as understory associated with gallery forest and as discrete patches between stands of gallery forest. Other associated vegetation included *Baccharis* spp., boxelder (*Acer negundo*), netleaf hackberry (*Celtis reticulata*), and honey mesquite (*Prosopis glandulosa*). The site received regular, but apparently low, use by livestock, hikers, bird-

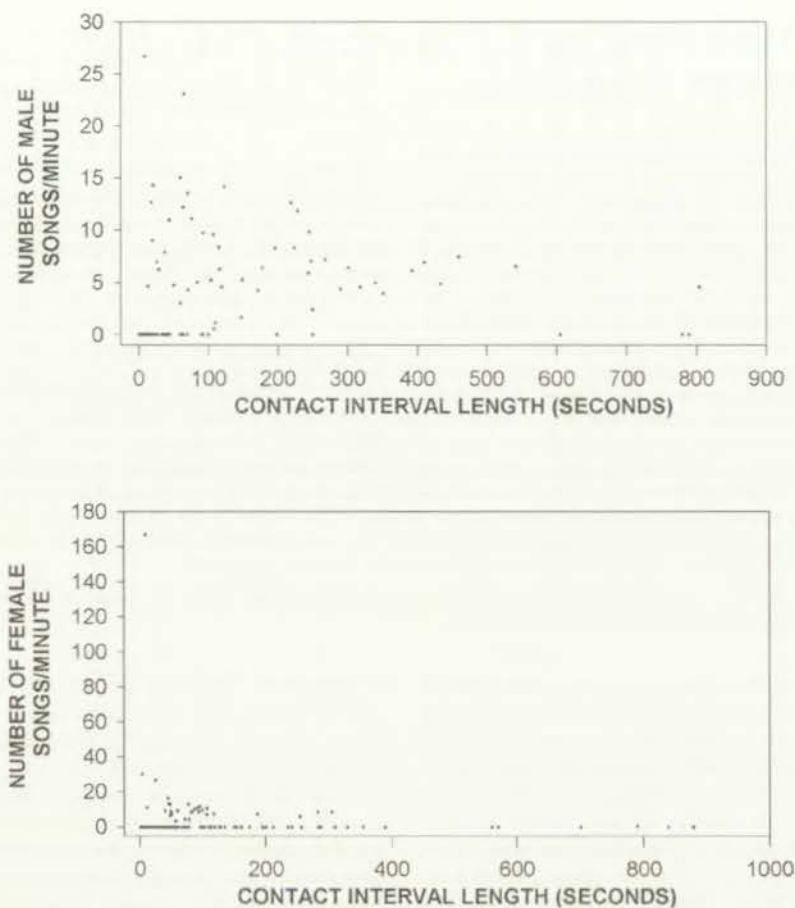


FIGURE 1. Number of songs/min by contact interval length, for male (top) and female (bottom) Southwestern Willow Flycatchers at Camp Verde, Arizona, May–August, 1996–1998.

watchers, fishermen, all-terrain-vehicle enthusiasts, and others.

The focal-animal continuous-observation technique (Bradley 1985) was used to sample singing behavior from mid-May through early August, 1996–1998. Vocalizations of flycatchers visible to observers were recorded during 15-min sample intervals (with few exceptions) occurring from 0500 to 2045 hrs (Mountain Standard Time), for up to four days/week. Contact intervals, in contrast, were those times during the sample interval when potentially singing flycatchers were actually in sight. Observers attempted to keep potentially singing flycatchers in sight throughout the sample interval; most observations were made from approximately 20–30 m.

We recorded the following variables for each song event: date and time, band color-combination, location (in nest, within 2 m of nest, or > 2 m from nest), position of individual (perched or in air/flight), and song type (*fitz-bew* or *creet*). Our “*fitz-bew*” song type included both *fitz-* and *fizz-bews*, which are difficult to differentiate in the field. Both are given at our study site, with *fitz-bews* given about twice as often as *fizz-bews* (Sedgwick 2000; J. Sedgwick, pers. comm.). We

selected 2 m for our nest proximity threshold because preliminary observations suggested very little flycatcher activity not directly related to maintenance of the nest or eggs/young (i.e., foraging, perching, singing, vigilance) took place within this distance. We were simultaneously monitoring flycatcher nest and territory locations, nesting activity, population dynamics, and other behaviors, which were the source of our information on banded birds and nesting phase (H. Yard and B. Brown, unpubl. data). Sex was determined for individually color-banded birds by brood patch (females) and cloacal protuberance (males), and by a genetic technique (Busch et al. 2000). Unbanded birds, birds whose band color-combinations could not be accurately determined, and birds for which both the physical examination and genetic techniques failed to precisely identify sex were categorized as unknown sex.

Scatter plots of male and female *fitz-bew* songs/min by contact interval length were made to evaluate visibility bias (Fig. 1). We defined visibility bias as the brief observation of an individual (<5 sec) during which several songs were uttered in quick succession. Subsequent conversion to songs/min would produce

TABLE 1. SOUTHWESTERN WILLOW FLYCATCHER SONG TYPES BY SEX AND LOCATION FOR ALL COMBINATIONS OF NESTING PHASE AND TIME OF DAY, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Song type	Number of male songs (452.2 contact min)				Number of female songs (920.2 contact min)				Total
	Perched >2 m from nest	Perched on or <2 m from nest	In air	Total	Perched >2 m from nest	Perched on or <2 m from nest	In air	Total	
<i>Fitz-bew</i>	1,225	0	23	1,248	560	40	45	645	1,893
<i>Creet</i>	257	0	8	265	147	8	15	170	435
Total	1,482	0	31	1,513	707	48	60	815	2,328

inflated values for these observations, which were not produced during longer contact intervals. Visual inspection of outliers revealed no apparent bias in rates of male song and only one instance of female bias, which was eliminated from further analysis.

We compared song rates (i.e., frequencies) by nesting phase (courtship/nest-building, egg-laying/incubation, nestling, and post-fledging) and time of day (morning, 0500–0959 hrs; midday, 1000–1459 hrs; and evening, 1500–2045 hrs). Song rates were calculated by computing songs/min over all contact periods, not just those in which songs were observed. Statistical tests for differences between and within parameters have not been presented due to serious concerns regarding pseudoreplication. Most individuals were sampled multiple times within years, resulting in a lack of independence between observations. All years were combined for analysis to maximize sample sizes for comparison, which further compounded pseudoreplication because some individuals were present and sampled in >1 year. Therefore, different mean values between or within sampled variables have been described as "apparently different." Data summaries and song rates were calculated using SPSS software (Norusis 1993). Some sample sizes differed due to missing data.

## RESULTS

Nineteen flycatchers were banded and their sex determined from 1996–1998; song data were gathered on 18 of these during 723 sample intervals for a mean of 40 intervals/individual (SD = 41, range = 1–141, median = 24). Three individuals were present and sampled in both 1996 and 1997, seven in 1997 and 1998, and one individual in all three years.

Most *fitz-bew* songs given by males and females were from perches. Of all male songs de-

tected, 2% were given in flight; 7% of all female songs were given in flight (Table 1). Aerial song by males, females, or individuals of unknown sex occurred throughout the nesting season. No male *fitz-bew* songs and only 6% of female songs were given on or <2 m from nests (Table 1), so all following rates of male and female song were calculated for those given from perches >2 m from nests (males: N = 1225 songs, females: N = 560 songs).

Male and female *fitz-bew* songs were detected in 66% (N = 96) and 46% (N = 76), respectively, of total contact intervals of perched individuals >2 m from active nests for which nesting phase and time of day was known. Males and females uttered 1–20 songs during 41% and 37% of total contact intervals, respectively, and gave  $\geq 21$  songs during 25% and 9% of intervals, respectively. Most (71%) of the total (N = 66) female song bouts occurred earlier in the nesting season, during either May or June (Table 2).

Mean rates of male *fitz-bew* song appeared to differ between nesting phases, with the greatest difference in means between courtship/nest-building (Table 3) and nestling phases. Mean rates of female song also appeared to differ by nesting phase, with egg-laying/incubation and post-fledging both noticeably greater than courtship/nest-building or nestling phases (Table 3). Mean rates of male *fitz-bew* song by time of day appeared to differ between morning/midday and evening (Table 4). Mean female song rates also appeared different by time of day, with morning > midday > evening (Table 4).

TABLE 2. NUMBERS OF FEMALE SOUTHWESTERN WILLOW FLYCATCHER *FITZ-BEW* SONG BOUTS OBSERVED AT ALL LOCATIONS BY MONTH AND NESTING STAGE, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Nesting phase	May	June	July	August	Total
Courtship/nest-building	2	1	1	0	4
Egg-laying/incubation	0	25	4	1	30
Nestling	0	1	6	0	7
Post-fledging	0	0	5	0	5
Unknown	18	0	2	0	20
Total	20	27	18	1	66

TABLE 3. MEANS ( $\bar{X}$ ), STANDARD DEVIATIONS (SD), AND SAMPLE SIZES (N = NUMBER OF CONTACT INTERVALS) FOR SOUTHWESTERN WILLOW FLYCATCHER *FITZ-BEW* SONGS/MIN BY SEX AND NESTING PHASE, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Nesting phase	Male songs/min				Female songs/min			
	N	$\bar{X}$	SD	Range	N	$\bar{X}$	SD	Range
Courtship/nest-building	29	5.2	6.3	0–27	5	0.1	0.2	0–0.5
Egg-laying/incubation	46	2.1	5.3	0–30	54	2.3	4.0	0–16
Nestling	18	0.1	0.2	0–1	79	0	0	0
Post-fledging	15	2.2	4.7	0–15	45	0.5	2.1	0–12
Total	108	2.6	5.3	0–30	183	0.8	2.6	0–16

Mean ( $\pm$  SD) *fitz-bew* song rates for males ( $2.9 \pm 5.2$ ; Table 4) and females ( $1.4 \pm 4.0$ ) appeared similar when perched individuals >2 m from nests were compared. However, males sang during a greater proportion of their contact intervals than females, and produced a greater overall number of songs than females despite their contact time of <50% that of females (Tables 1 and 4). Female contact time was 2.035 times more than male contact time when all locations, times of day, and nesting phases were combined (Table 1). To correct for discrepancies in contact time by sex, we multiplied the total number of male songs (N = 1248) by 2.035. This resulted in 645 female songs and an adjusted 2540 male songs, for a total of 3185 songs by individuals of known sex. Females gave an adjusted 20% of all *fitz-bew* songs detected, or, males gave an adjusted four songs for each female song.

Males gave the *creet* song 3.6 times more than females when the abundance of this call was corrected for overall contact minutes by sex (Table 1; songs >2 m from nest). Males were never observed giving this song from or near nests, but females occasionally gave it from nests (Table 1). Both sexes gave this song in flight on rare occasions (Table 1).

Male *creet* song was relatively common during courtship/nest-building and egg-laying/incubation but was not detected during nestling and post-fledging phases (Table 5). Female rates of *creet* song appeared highest during egg-lay-

ing/incubation (Table 5), a phase when the mean rate of female *creet* song appeared much higher than any male phase. Both male and female rates of *creet* song appeared to decrease with time of day (Table 6).

#### DISCUSSION

Sedgwick (2000) and Sogge (2000b) reported that Willow Flycatcher song (presumably by males) was most common shortly after arrival on the breeding grounds and early in the nesting cycle, declining with the season and particularly after pairing. Our findings are consistent with this general pattern for both *fitz-bew* and *creet* songs by males. An exception was the high rate of male *fitz-bew* song in the post-fledging phase, which, due to the prevalence of second clutches, was probably early in the following nesting cycle when male song would be expected to exhibit another peak.

Sedgwick (2000) and Sogge (2000) also reported that males sang throughout the day, but that the highest rates of song were in the early morning. In contrast, McCabe (1991) found that male song was not as consistent or vigorous in morning compared to late evening. Our findings were consistent with the former pattern, as the males at Camp Verde sang throughout the day, and mean song rates of both *fitz-bews* and *creets* were greater for males during morning. However, our need to combine data from all years and nesting phases for analysis may have masked subtle differences between phases that

TABLE 4. MEANS ( $\bar{X}$ ), STANDARD DEVIATIONS (SD), AND SAMPLE SIZES (N = NUMBER OF CONTACT INTERVALS) FOR SOUTHWESTERN WILLOW FLYCATCHER *FITZ-BEW* SONGS/MIN BY SEX AND TIME OF DAY FOR ALL NESTING PHASES COMBINED, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Time of day	Male songs/min				Female songs/min			
	N	$\bar{X}$	SD	Range	N	$\bar{X}$	SD	Range
Morning (0500–0959 hrs)	88	3.1	5.3	0–30	83	3.0	5.7	0–30
Midday (1000–1459 hrs)	57	3.4	5.5	0–27	100	0.8	2.6	0–16
Evening (1500–2045 hrs)	18	0.8	1.9	0–6	48	0.0	0.0	0
Total	163	2.9	5.2	0–30	231	1.4	4.0	0–30

TABLE 5. MEANS ( $\bar{X}$ ), STANDARD DEVIATIONS (SD), AND RANGE OF WILLOW FLYCATCHER CREET SONGS/MIN ( $N$  = SAMPLE SIZE OF CONTACT INTERVALS) BY SEX AND NESTING STAGE FOR ALL TIMES OF DAY, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Nesting phase	Male songs/min				Female songs/min			
	N	$\bar{X}$	SD	Range	N	$\bar{X}$	SD	Range
Courtship/nest-building	29	0.6	0.6	0–2	5	0	0	0
Egg-laying/incubation	46	0.5	1.2	0–6	55	2.8	16.1	0–120
Nestling	18	0	0	0	79	0.1	0.4	0–3
Post-fledging	15	0	0	0	46	0.1	0.4	0–2
Total	108	0.4	0.9	0–6	185	0.9	8.8	0–120

we were not able to evaluate. We observed male song in late evening but could not evaluate the evening song peak described by McCabe (1991). Flycatcher activity was reduced during evenings, making it much more difficult for us to locate individuals for observation of singing behavior.

Our study of a small population of flycatchers resulted in a high rate of pseudoreplication due to repeated samples of the same individuals. This lack of independence between most of our samples was a potential source of bias, as it is in all studies of endangered species or small populations (Mannan and Meslow 1984, Noon and Block 1990). The necessity of netting, color-banding, and then accurately determining sex of individuals to be sampled for singing behavior will limit sample sizes and hinder efforts to duplicate this study elsewhere. We recommend that future studies of Southwestern Willow Flycatcher singing behavior attempt to gather samples from many geographic areas in order to determine if the descriptive (i.e., not statistically confirmed) patterns we observed exist rangewide. Until this study can be duplicated elsewhere, our findings should be interpreted cautiously.

The finding that females sang commonly at rates similar to males was contrary to findings of other investigators. Females gave an adjusted 20% of all *fitz-bew* songs detected during our study, suggesting a probability of 0.2 that any song detected at random during presence-absence surveys (with or without a simultaneous visual contact) would be a female. Flycatcher plumage is sexually monomorphic and few re-

liable behavioral cues exist to differentiate sex. With an estimated probability of only 0.8 that a randomly-encountered, singing individual is a male, determining sex based on song alone should be considered unreliable.

Our examination of female song identified several issues which should be considered in similar future studies. First, the probability that a randomly-encountered singing individual is a female might be even greater than our data indicated. This is because while female contact intervals were three times more numerous than males', their contact minutes were only two times greater. Thus, female contact intervals were on average shorter than males', which could have led to bias against detection of female song. Second, we were unable to determine the extent to which potential pseudoreplication might have influenced our determination of female song rate (i.e., only a few females might sing frequently while most might rarely or never sing), and future studies should consider this issue. Finally, our study was not designed to determine the purpose and function of female song, and this topic awaits examination.

The implications of relatively high rates of female song in an endangered subspecies whose population size is estimated by auditory techniques are profound. Males and females regularly sang while in flight and we occasionally observed both members of a pair singing simultaneously, behaviors which could cause observers to conclude that more singing males were

TABLE 6. MEANS ( $\bar{X}$ ), STANDARD DEVIATIONS (SD), AND RANGE OF WILLOW FLYCATCHER CREET SONGS/MIN ( $N$  = SAMPLE SIZE OF CONTACT INTERVALS) BY SEX AND TIME OF DAY FOR ALL NESTING PHASES COMBINED, CAMP VERDE, ARIZONA, MAY–AUGUST, 1996–1998

Time of day	Male songs/min				Female songs/min			
	N	$\bar{X}$	SD	Range	N	$\bar{X}$	SD	Range
Morning (0500–0959 hrs)	88	1.2	3.1	0–23	85	2.2	13.1	0–120
Midday (1000–1459 hrs)	57	0.4	0.6	0–2	100	0.2	0.8	0–5
Evening (1500–1845 hrs)	18	0.3	0.6	0–2	48	0	0	0

present on more territories than were actually present at a site. Quick estimates of population size at breeding sites with  $\geq 5$  pairs may be particularly unreliable because of unavoidable simultaneous contacts with multiple singing birds. Intensive nest and population monitoring would produce more reliable estimates of population size, but such monitoring is expensive and time-consuming. Marshall (2000) pointed out that only 25 sites with  $\geq 5$  pairs exist rangewide out of a total of 109 Southwestern Willow Flycatcher breeding sites, but they support almost one-half of all known territories. Therefore, even a small element of unreliability could substantially influence rangewide estimates of population size that are employed to measure recovery.

#### ACKNOWLEDGMENTS

This study was funded by the Phelps Dodge Development Corporation (PDDC) as a result of an Endangered Species Act Section 7 consultation with the U.S. Fish and Wildlife Service (USFWS) for the Verde Valley Ranch development. M. Fitzgerald, B. Kammerle, E. Kerwin, and D. Meidinger of PDDC and J. Thomas of SWCA provided administrative and logistical support. S. Ackers, B. Marshall, and M. Murov provided field assistance. Original study design elements were suggested by R. Marshall (USFWS). We thank S. Langridge, T. Littlefeather, M. Moore, J. Owen, E. Paxton, and M. Sogge of Northern Arizona University for banding flycatchers and determining sex. C. Gazak of the University of Utah helped to summarize data. Comments by B. Kus, J. Sedgwick, and an anonymous reviewer improved a draft of this manuscript.