# Increased attenuation and reverberation are associated with lower maximum frequencies and narrow bandwidth of bird songs in cities

## Jennifer N. Phillips, Catherine Rochefort, Sara Lipshutz, Graham E. Derryberry, David Luther & Elizabeth P. Derryberry

Springer

Springer

**ORIGINAL ARTICLE**

# Increased attenuation and reverberation are associated with lower maximum frequencies and narrow bandwidth of bird songs in cities

Jennifer N. Phillips[1,2] · Catherine Rochefort[1,3] · Sara Lipshutz[1,4] · Graham E. Derryberry[5] · David Luther[6,7] · Elizabeth P. Derryberry[5]

## Abstract

As urbanization expands globally, the communication systems of an increasing number of species are affected. Because bird song is a long-distance signal used to attract mates and defend territories, the evolution of bird song is often shaped by habitat structure and background noise. These potential drivers of song evolution are more often studied in natural areas than in urbanized areas, leaving open the question of how anthropogenic changes to the landscape are affecting the evolution of bird song. One songbird that persists in both urbanized and rural areas in North America is the White-crowned Sparrow (*Zonotrichia leucophrys*). Our previous work demonstrates that increased background noise in cities and in natural habitats affects acoustic adaptation in this species. However, we lack information about how sound transmits in urban and rural habitats. Because cities tend to have different physical properties than rural areas, it is pertinent to understand the influence of urban habitat structure on song evolutionary processes. Here, we test the acoustic adaptation theory and hypothesize that differences in the sound transmission properties of urban and rural habitats affect song divergence between urban and rural populations. We conducted sound transmission trials of tones from 1 to 8 kHz on 59 White-crowned Sparrow territories in three urban and three rural locations around San Francisco and Point Reyes, California, at varying heights and rates of note production. We also recorded and analyzed songs of males from each location to see if differences in signal–noise ratio, attenuation, reverberation, and distortion predicted song divergence. We found that urban locations have higher attenuation and reverberation than rural locations and urban songs tend to have short whistles, faster trills, and narrower bandwidth. These findings partially support acoustic adaptation theory, though faster trills in the city may instead be driven by cultural drift or sexual selection. Overall, our results add to a rapidly growing area of research and allow us to better understand the complexity of influences on song divergence.

**Keywords** Bioacoustics · Birdsong · Communication · Sound transmission · Urban ecology

Communicated by S. Kipper.

This article is a contribution to the Topical Collection 27th International Ornithological Congress, Vancouver, Canada, 19–26 August 2018.

✉ Jennifer N. Phillips
jnphilli@calpoly.edu

Extended author information available on the last page of the article

**Zusammenfassung**

**Erhöhte Dämpfung und Nachhall sind mit niedrigeren Maximalfrequenzen und einer engen Bandbreite von Vogelliedern in Städten verbunden**

Da die Urbanisierung weltweit zunimmt, sind die Kommunikationssysteme einer zunehmenden Anzahl von Arten betroffen. Da Vogelgesang ein Fernsignal ist, das verwendet wird, um Partner anzuziehen und Gebiete zu verteidigen, wird die Entwicklung des Vogelgesangs häufig durch die Struktur des Lebensraums und Hintergrundgeräusche geprägt. Diese potenziellen Treiber der Liedentwicklung werden häufiger in natürlichen Gebieten als in städtischen Gebieten untersucht, wobei die Frage offen bleibt, wie sich anthropogene Veränderungen der Landschaft auf die Entwicklung des Vogelgesangs auswirken. Ein Singvogel, der sowohl in städtischen als auch in ländlichen Gebieten Nordamerikas vorkommt, ist der Weißkronenspatz (Zonotrichia leucophrys). Unsere früheren Arbeiten zeigen, dass vermehrte Hintergrundgeräusche in Städten und in natürlichen Lebensräumen die akustische Anpassung dieser Art beeinflussen. Es fehlen jedoch Informationen darüber, wie Schall in städtischen und ländlichen Lebensräumen übertragen wird. Da Städte tendenziell andere physikalische Eigenschaften haben als ländliche Gebiete, ist es wichtig, den Einfluss der städtischen Lebensraumstruktur auf die Entwicklungsprozesse von Liedern zu verstehen. Hier testen wir die Theorie der akustischen Anpassung und stellen die Hypothese auf, dass Unterschiede in den Schallübertragungseigenschaften von städtischen und ländlichen Lebensräumen die Lieddivergenz zwischen städtischen und ländlichen Bevölkerungsgruppen beeinflussen. Wir haben Schallübertragungsversuche mit Tönen von 1 bis 8 kHz in 59 White-Crowned Sparrow-Gebieten in drei städtischen und drei ländlichen Gebieten um San Francisco und Point Reyes, Kalifornien, in unterschiedlichen Höhen und Raten der Notenproduktion durchgeführt. Wir haben auch Songs von Männern von jedem Ort aufgenommen und analysiert, um festzustellen, ob Unterschiede im Signal-Rausch-Verhältnis, der Dämpfung, dem Nachhall und der Verzerrung die Song-Divergenz vorhersagen. Wir haben festgestellt, dass städtische Standorte eine höhere Dämpfung und einen höheren Nachhall aufweisen als ländliche Standorte und städtische Songs tendenziell kurze Pfeifen, schnellere Triller und eine geringere Bandbreite aufweisen. Diese Ergebnisse stützen teilweise die akustische Anpassungstheorie, obwohl schnellere Triller in der Stadt stattdessen durch kulturelle Drift oder sexuelle Selektion angetrieben werden können. Insgesamt tragen unsere Ergebnisse zu einem schnell wachsenden Forschungsbereich bei und ermöglichen es uns, die Komplexität der Einflüsse auf die Songdivergenz besser zu verstehen.

## Introduction

Ecological conditions affect the transmission of signals from sender to receiver (Morton 1975; Wiley and Richards 1978). A number of factors can contribute to the degradation of acoustic signals, including ambient noise levels, wind, microclimate, and the density and texture of surfaces that reflect or absorb sound, such as vegetation. Signals may be modified through loss of intensity (attenuation) or scattering of sound waves (Wiley and Richards 1978), leading to a degraded signal reaching a receiver. Given that receivers are often competitors or potential mates (Andersson 1994), signal degradation may have fitness consequences, where receivers tend to respond less to degraded songs (Peters et al. 2012). Therefore, signals that maximize communication distance and minimize degradation should be favored by natural selection (Wiley 2006). The idea that communication systems adapt to local environments is known as acoustic adaptation (Endler 1992). As environments change, signals are also predicted to change in ways that continue to maximize transmission (Derryberry 2007, 2009). For example, bamboo and *terra firme* habitats in the Neotropics vary in sound transmission properties, and the songs of bamboo-specialist birds vary as predicted from their nearest relative in *terra firme* habitat (Tobias et al. 2010). In many taxa, this process

of signal divergence is an important factor in the evolution of reproductive isolation (Boughman 2002).

Recently, song divergence has been applied in an evolutionarily new environment—cities (Slabbekoorn and den Boer-Visser 2006). Human-generated sensory pollution, in the form of noise levels, is evolutionarily unprecedented (Swaddle et al. 2015). High-amplitude noise levels can obscure or mask a signal occurring within the same frequency range. Empirical studies indicate that animals persisting in urbanized areas are, in some cases, adapting their communication systems to these relatively new selection pressures (Slabbekoorn 2013; LaZerte et al. 2016; Luther et al. 2016). For example, a number of bird species in urban areas sing louder and with higher minimum frequencies than in more rural areas, which reduces masking by high-amplitude, low-frequency anthropogenic noise (Slabbekoorn 2013). Thus, there is evidence that signals are diversifying and even diverging between rural and urban areas (Slabbekoorn et al. 2007).

Although sensory pollution in terms of human-generated noise is one aspect affecting sound transmission in cities, it is not the only factor. Signals degrade in amplitude and in spectral and temporal structure as they propagate through the environment (Wiley and Richards 1982). They are also subject to further degradation due to absorption by impervious

surfaces, such as roads, and reflection by vertical structures in the built environment (Richards and Wiley 1980; Wiley and Richards 1982; Warren et al. 2006). These factors should affect the transmission profiles of urban environments, and thus the selection pressures acting on signals. To understand more fully the role of acoustic adaptation in urbanized environments, there is a need to assess whether populations experiencing different degrees of urbanization have different transmission profiles and correspondingly different communication signals.

Transmission profiles can be described by measuring four key components: the strength of a signal relative to background noise (signal–noise ratio, SNR), reduction in the intensity or purity of a signal (attenuation), increases in the persistence of sound after the original sound is produced (reverberations) and the extent to which reverberations from a previous note affect the fidelity of following notes (distortion) (Wiley and Richards 1978; Richards and Wiley 1980). Previous work on this topic indicates that acoustic signals in more urban areas experience greater reverberation (Katti and Warren 2004; Slabbekoorn et al. 2007; Kight et al. 2013) and lower signal–noise ratios (Kight et al. 2012) than signals in more rural areas. In contrast, at least two studies suggest that urban areas may degrade songs less (e.g., have less excess attenuation and higher SNRs; Slabbekoorn et al. 2002; Nemeth et al. 2006), potentially because hard surfaces (roads, buildings) may enhance propagation of higher frequencies (Mockford et al. 2011; Gall et al. 2012). In either case, transmission studies predict divergence in at least some parameters of song structure in response to divergence in environmental transmission profiles. There is some indication that urban songs transmit better than rural songs in urban habitats, although not vice versa (Mockford et al. 2011). A study investigating dark-eyed junco (*Junco hyemalis*) song transmission found that urban habitats create multiple echoes (e.g., more unpredictable distortion), while forest habitats create reverberations resulting in a tailed reflection of sound, indicating both urban and rural environments can distort signals (Slabbekoorn et al. 2007). Consistent with acoustic adaptation theory that pure tones are less degraded in areas with many reflective surfaces, urban junco songs have narrower bandwidth than rural songs, with both maximum and minimum frequency affected; however, trill rate and song duration were not significantly affected by the urban environment (Slabbekoorn et al. 2007). While it is expected that urban areas generally have transmission properties that degrade sound more than natural areas, and that animal songs have adapted to increase signal transmission given these environmental constraints, transmission properties and song structures are not often studied together. This suggests that we need more information about the transmission properties of various soundscapes (i.e., biotic and abiotic sound in a given environment) and the physical
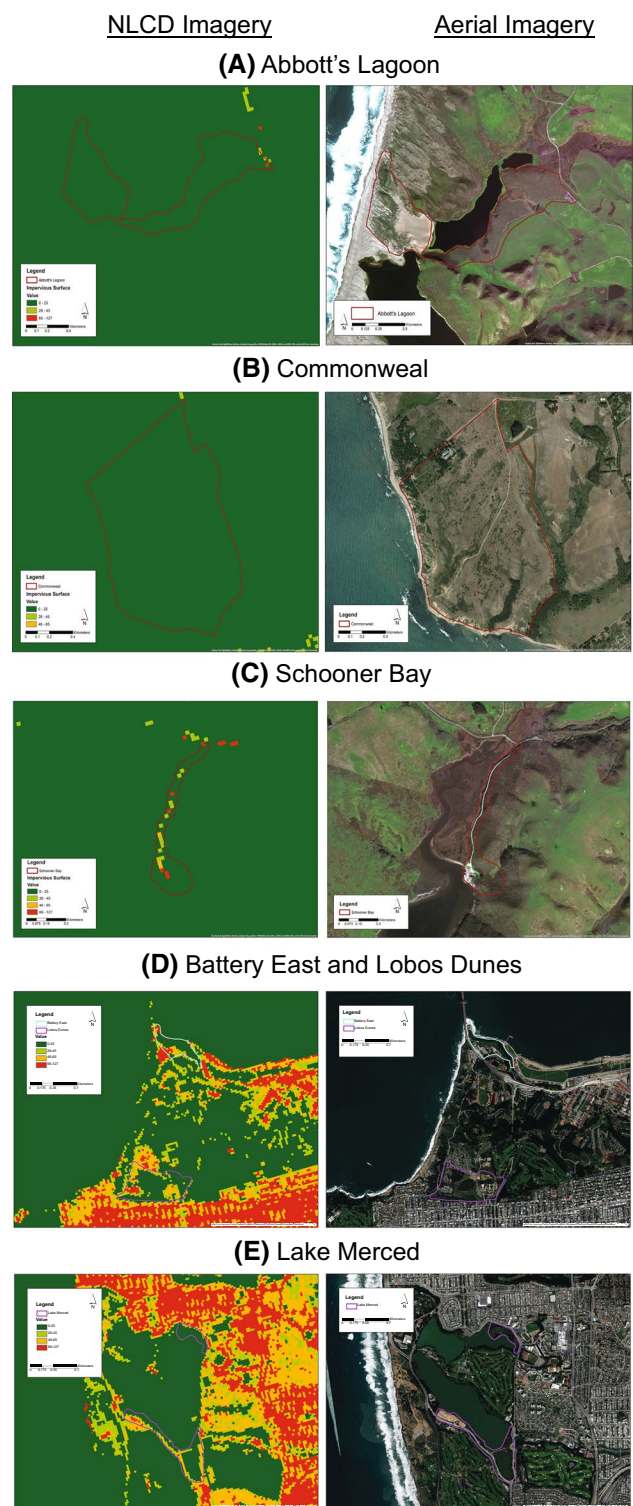


**Fig. 1** Map showing NLCD 2006 cover for rural **a** Abbott's Lagoon, **b** Commonweal and **c** Schooner Bay, contrasting with high impervious cover in urban, **d** Battery East and Lobos Dunes and **e** Lake Merced. Maps built with NLCD layer (left) and ESRI world imagery (right) in ArcMap 10 (ESRI, Redlands, CA, USA; Sources: ESRI, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community) (color figure online)

properties that are associated with differences in song structure to better understand how songs diverge between urban and rural populations.

Here, we investigate how habitat type and the physical landscape affect transmission of Nuttall's White-crowned Sparrow (*Zonotrichia leucophrys nuttalli;* NWCS) songs. NWCS have distinct geographic differences in their song structure termed 'dialects' (Marler and Tamura 1962), and within dialects songs vary with differences in habitat structure, at least in rural locations (Derryberry 2009). Furthermore, over the past 30 years, NWCS songs have increased in song minimum frequency, coinciding with an increase in noise in San Francisco (Luther and Baptista 2010). Songs of current populations vary with the soundscape in minimum frequency, bandwidth, and vocal performance (Derryberry et al. 2016; Luther et al. 2016). In this study, we measured the sound transmission properties of three urban and three rural locations near San Francisco, California using playback of tones at known frequencies. We also analyzed levels of impervious cover in each location. From these data, we extracted information on signal degradation, including signal–noise ratios, attenuation, reverberation and distortion. We predicted that urban locations would have higher levels of attenuation at high frequencies, lower signal–noise ratios, and more reverberation and distortion than rural locations. We also measured temporal and spectral characteristics of songs produced by territorial males to see if songs are diversifying, such that they fill the most appropriate acoustic niche based on each location's transmission profile. Based on our expected transmission profiles, we would expect urban birds to produce songs with narrower bandwidths, higher minimum frequencies, lower maximum frequencies, lower pitched and shorter whistles, and slower trills compared to rural birds. We also expect the gradient of urbanization across locations to affect the degree of these predicted patterns. Our study design allows us to test the theory of acoustic adaptation both within and across cultural units as males from the three urban locations sing the same song type, and males from the three rural locations sing three different song types.

## Study system

We conducted our sound transmission study in six locations. Three locations were within city limits in San Francisco County, San Francisco, CA (Battery East, Lake Merced, and Lobos Dunes) and three locations were in Point Reyes National Seashore, Marin County, CA (Commonweal, Abbott's Lagoon, and Schooner Bay; Fig. 1). All locations were within 50 miles of each other. Battery East is an urban area with ambient noise resulting from traffic of the Golden Gate Bridge, bay water waves, wind, human foot traffic, automobile traffic, and planes overhead. This area is dominated by dense shrub species such as coast buckwheat (*Eriogonum latifolium*), coast angelica (*Angelica hendersonii*), and coyote brush (*Baccharis pilularis*). Impervious surfaces are present in the form of buildings, paths, and the historic battery structures built after the civil war. Lake Merced is the second urban area with ambient noise from busy roads, human foot traffic, and a gun shooting range. The area is dominated by tree species, such as Monterey pine (*Pinus radiata*) and Monterey cypress (*Cupressus macrocarpa*), and an understory composed of poison oak (*Toxicodendron diversilobum*), silver bush lupine (*Lupinus albifrons*), and California blackberry (*Rubus ursinus*). Apartment buildings are present on some sides of the lake. Lobos Dunes is the third urban area with ambient noise resulting from some human foot traffic, automobile traffic, and planes overhead. It is a dune scrub community consisting of plant species such as yellow bush lupine (*Lupinus arboreus*), silver bush lupine (*Lupinus albifrons*), chamisso's lupine (*Lupinus chamissonis*), mock heather (*Ericameria ericoides*), and lizard tail (*Eriophyllum staechadifolium*). A large company building and houses surround half of this area. Point Reyes National Seashore is a rural area in Marin County with ambient noise resulting from ocean waves, wind, cows, and other birds. Some distant noise may occur from farms or rural roads, but in general, this area has an exceptionally low human population. The Commonweal habitat is pastureland with poison oak (*Toxicodendron diversilobum*) and blackberry (*Rubus* spp.) bushes dispersed throughout. A few remnant concrete pads are present and a rural road borders the eastern side. Abbott's Lagoon is a coastal scrub and dune habitat, and is covered with old growth northern coastal scrub species. Schooner Bay has a similar habitat type, but has more variation in elevation and had some minor anthropogenic disturbances due to an oyster farm on the location (Fig. 1).

Several subspecies of White-crowned Sparrows can be found year-round in these locations, but only one remains to breed (*Z. l. nuttalli*) (Blanchard 1936, 1941). Males begin to defend breeding territories and reduce flocking behavior in early March, and females initiate clutches as early as March/April (Mewaldt and King 1977; DeSante and Baptista 1989; Phillips pers. obs.). Breeding continues through August, so juveniles may be found from late April through September foraging in groups and receptive to learning song. We conducted the sound transmission experiments from May to July of 2013, when receptive females, territorial males and juveniles receptive to learning song are all potential receivers of signal information. These potential communication scenarios informed how we collected sound transmission data.

## Transmission: methods

### Habitat structure

To assess habitat structure, we extracted each of the six location's impervious cover using the National Land Cover Database 2006 dataset, which has a resolution of 30 m LandSat imagery (Homer et al. 2011, 2012; Phillips et al. 2018). We used the spatial analyst tool to extract impervious values in ArcGIS 10 (ESRI, Redlands, CA, USA). The total area surveyed for occupancy by NWCS defined each location.

### Sound transmission

We quantified the sound transmission properties of the six locations to determine if they had distinct acoustic environments. We broadcast and rerecorded artificial sound stimuli in each location to quantify the extent to which sounds are differentially degraded. This is a standard approach used to determine transmission profiles in a range of habitat types (Morton 1975; Slabbekoorn et al. 2007; Tobias et al. 2010).

### Sound stimuli

We generated artificial sound stimuli in SIGNAL v5 (Beeman 1998). As NWCS songs are tonal in structure, such that most notes are free of overtones (harmonics), we used pure tones as sound stimuli. We used a standard approach of creating stimuli at a constant frequency. Pure tones of 100-ms duration were played at 1, 1.5, 2, 2.5, 3, 4, 5, 6, 7, and 8 kHz. We included half steps between 1 and 3 kHz to capture how the sound transmission environment would impact minimum frequency degradation. The remaining frequencies span the frequency range of typical White-crowned Sparrow songs in these locations (Baptista 1975). At each frequency, two tones were played at three different intervals: 10 ms, 100 ms, and 300 ms. These values reflect three different rates of note production, 9, 5 and 2.5 notes per second. We selected 9 and 5 as the majority of NWCS trills in rural areas are produced between 5 and 10 notes per second, and we selected 2.5 to assess any limitations on reducing trill rate to minimize reverberation.

### Experiments

Transmission experiments were conducted on 59 male White-crowned Sparrow territories, in Battery East ($n=10$), Lake Merced ($n=14$), Lobos Dunes ($n=10$), Abbott's Lagoon ($n=9$), Schooner Bay ($n=6$), and Commonweal ($n=10$). Experiments took place between 6 AM and 12 PM during the breeding seasons of 2013 and 2014. This time of day is a typical period of peak vocal activity in the breeding

season. We used an I-Pod Nano to broadcast the sound file through an SME-AFS loudspeaker (Saul Mineroff Electronics, Elmont, NY, USA), and recorded it as a 44.1 kHz WAV file using a unidirectional Sennheiser ME67 Microphone and a Marantz PMD-661 MKII digital recorder. The same equipment was used for all transmission experiments. After territories had been determined by mapping song posts, the speaker was placed near the center of a male's territory. The transmission playback was then aligned in a randomly chosen cardinal direction. Stimuli were broadcast at a 1.5-m height and recorded using two unidirectional microphones: one at 1.5-m height and one at 0.5-m height. The two heights simulated three different types of sound receivers: a counter singing male at 1.5-m height and a receptive female or juvenile at 0.5-m height. Recordings of the transmission stimuli were made at a distance of 5 m and 30 m and recorded set gains (5.5 gain at 5 m and 7 gain at 30 m). The recordings at 5 m were used to account for the effects of equipment on signal transmission. The recordings at 30 m simulated the distance between two average White-crowned Sparrow male territories (Chilton et al. 1995). All transmission playbacks were repeated three times on each territory without touching the equipment between playbacks. Volume levels for playback were standardized at 80 dB SPL at 1 m to approximate natural songs using a Radioshack 7-range sound level meter (Fort Worth, Texas, USA).

### Transmission analysis

All sound files were resampled at 25 kHz prior to analysis, and data were extracted using a macro-program in SIGNAL v5 (Beeman 1998). Timing of data collection was synchronized to an initial set of frequency sweeps on the playback file. Three amplitude measurements [root-mean-squared (RMS), values of sound pressure deviations in volts] were taken during 100-ms periods; measurement A was taken after the onset of the first tone; measurement B after the offset of the first tone; measurement C after the onset of the second tone. We took a fourth measure of noise amplitude, measurement N, which was taken during 100 ms prior to the onset of the first tone. All amplitude measurements are relative not absolute. From these measures, we quantified four values to describe environmental degradation of tones. (1) Signal–noise ratio (SNR) is the ratio of the strength of a tone to the strength of background noise. SNR was calculated as A divided by N. Sounds with higher SNR values are easier to differentiate from background noise. (2) Attenuation is the reduction in intensity and purity of a signal, which can result from reduced sound amplitude. Attenuation was calculated as A at 30 m divided by A at 5 m. (3) Reverberations are echoes that increase the persistence of a sound after the original sound is produced (Wiley and Richards 1978; Richards and Wiley 1980). When sounds are reflected off of surfaces,

**Table 1** Effects of signal frequency, habitat, and transmission height on (A) attenuation, (B) reverberation, and (C) signal–noise ratio

| Sources of variation | df | SS | F ratio | P |
|---|---|---|---|---|
| *(A) Attenuation* | | | | |
| Main effects | | | | |
| Frequency | 9 | 27.7 | 57.6 | **< 0.0001** |
| Habitat | 1 | 2.9 | 39.0 | **< 0.0001** |
| Height | 1 | 0.03 | 0.38 | 0.54 |
| Two-way interactions | | | | |
| Frequency × habitat | 9 | 4.8 | 11.21 | **< 0.0001** |
| Frequency × height | 9 | 0.34 | 0.7 | 0.72 |
| Habitat × height | 1 | 0.68 | 9.1 | **0.0026** |
| Final model | | | | |
| $r^2 = 0.41$ | 21 | 36.1 | 36.8 | **< 0.0001** |
| *(B) Reverberation* | | | | |
| Main effects | | | | |
| Frequency | 9 | 52.8 | 90.8 | **< 0.0001** |
| Habitat | 1 | 30.5 | 371.2 | **< 0.0001** |
| Height | 1 | 0 | 0.0 | 0.92 |
| Two-way interactions | | | | |
| Frequency × habitat | 9 | 1.5 | 4.4 | **< 0.0001** |
| Frequency × height | 9 | 0.2 | 0.3 | 0.96 |
| Habitat × height | 1 | 0.0 | 0.4 | 0.54 |
| Final model | | | | |
| $r^2 = 0.66$ | 19 | 84.4 | 119 | **< 0.0001** |
| *(C) Signal–noise ratio* | | | | |
| Main effects | | | | |
| Frequency | 9 | 50.1 | 89.1 | **< 0.0001** |
| Habitat | 1 | 38.4 | 533.0 | **< 0.0001** |
| Height | 1 | 0.0 | 0.0 | 0.93 |
| Two-way interactions | | | | |
| Frequency × habitat | 9 | 2.3 | 9.1 | **< 0.0001** |
| Frequency × height | 9 | 0.2 | 0.3 | 0.97 |
| Habitat × height | 1 | 0.1 | 1.1 | 0.29 |
| Final model | | | | |
| $r^2 = 0.74$ | 19 | 91.1 | 171.8 | **< 0.0001** |

Bold indicates significance at the $\alpha = 0.05$ level

Statistics are from linear models with main and two-way interaction effects. The final model includes only significant terms. Artificial stimuli were broadcast at 10 frequencies. Habitat refers to whether transmission experiments took place in urban or rural habitats. Transmission height was the height of the microphone (either 0.5 m or 1.5 m) above ground level

echoes travel a longer distance before reaching a receiver, and, therefore, affect the temporal structure of a signal by smearing into the silent space between notes. Reverberation was calculated as B at 30 m divided by A at 30 m. (4) Distortion is a measure of the extent to which reverberations from a previous note affect the fidelity of following notes. Distortion was calculated as C at 30 m divided by A at 30 m. Only tones separated by 100 ms and 300 ms were used to calculate

reverberation, as B could not be taken for tones separated by 10 ms. Rapid, fast, and slow paced stimuli were pooled to calculate SNR, attenuation, and reverberation. For each territory, we calculated SNR, attenuation, reverberation and distortion values and then averaged those values across the three replicate playbacks.

### Transmission statistical analyses

The effect of habitat type (e.g., 'urban' and 'rural'), frequency and transmission height on SNR, attenuation and reverberation were tested using linear models with main effects and first-order interactions. We used the same approach to assess the effect of habitat type, frequency and note pace on distortion. Effects of location were assessed using post hoc comparisons (Tukey HSD). Prior to analyses, all variables were transformed as needed to meet assumptions of normality and homogeneity of variance. Statistical analyses were conducted in R 3.6.1 (R Development Core Team 2015) and in JMP 11 (SAS Institute Inc., Cary, NC, 1989–2019).

## Results: transmission

### Habitat structure

We found significant differences between urban and rural habitats in NLCD impervious surface ($R^2 = 0.95$, $F_{1,4} = 84.9$, $P = 0.0007$). Urban locations have significantly more impervious cover than rural locations (impervious % cover mean $\pm$ SD: urban $21.3 \pm 0.01$; rural: $0.02 \pm 0.35$). Impervious surface mostly consists of cement or other anthropogenic surfaces, although hard dirt trails or roads are also counted in this assessment and make up most rural impervious cover.

### Habitat sound transmission properties

#### Signal–noise ratio (SNR)

Habitat type explained most of the variation in signal–noise ratios (SNR), such that all three urban locations had much lower SNR than the three more rural locations (Table 1, Fig. 2), although each location had a unique SNR profile (Table 2). The two most urbanized locations (Battery East and Lake Merced) had extremely low signal–noise ratios at all frequencies, suggesting that very little of the signal may be discriminated or even detected at these distances. Among the three rural locations, Abbott's Lagoon had the highest SNR and Schooner Bay the lowest SNR, consistent with the relative degree of human presence in those locations. Across all locations, frequencies between 1.5 and 2.5 kHz peaked in their signal–noise ratio, again suggesting a sound window at this frequency range.

## Attenuation

We found a significant effect of signal frequency and habitat type (i.e. 'urban' vs. 'rural') on attenuation (Table 1, Fig. 2). Frequency explained most of the variation in attenuation, such that attenuation increased as frequency increased. Attenuation was significantly greater in urban than in rural locations, and increased with frequency much more rapidly in urban than in rural locations. Microphone height did not explain significant variation in attenuation, although in the rural locations, there tended to be more attenuation at higher transmission heights. Location explained significant variation in attenuation (Table 2), with Battery East territories experiencing overall the highest levels of attenuation and Abbott's Lagoon the lowest (Tukey HSD $P < 0.05$). Most of the variation in attenuation among locations occurred at the higher frequencies (6–8 kHz) with overlap among populations. Only Battery East and Abbott's Lagoon were significantly different at 6 and 7 kHz, with Battery East having higher attenuation (Tukey HSD $P < 0.05$). Both Battery East and Lake Merced had significantly higher attenuation than the three rural locations at 8 kHz (Tukey HSD $P < 0.05$). Considering the lower frequencies (1–5 kHz), most locations were similar in attenuation values; however, there was a prominent attenuation window between 3 and 4 kHz in Battery East that was not present in the other locations.

## Reverberation

Signal frequency, habitat type and the interaction between frequency and habitat type had a significant effect on how signals degraded through reverberation (Table 1, Fig. 2). Habitat type explained most of the variation in reverberation, with all three urban locations consistently showing higher levels of reverberation at all frequencies compared to the three rural locations. Lobos Dunes, as the least urbanized location with few structures and little traffic, experienced an intermediate level of reverberation. All of the locations varied in reverberation, with a significant interaction between frequency and location (Table 2). However, at all locations reverberation increased nearly linearly with frequency, except for a drop between 1.5 and 2.5 kHz, which seems to be a particular sound window with the lowest level of reverberation in all locations.

## Distortion

Signal pace, frequency and habitat type all explained significant variation in signal distortion (Table 3, Fig. 3). Distortion was greatest in slow paced signals (Tukey HSD $P < 0.05$) but not different between fast and rapid pace (Tukey HSD $P > 0.05$). This pattern was exhibited in Lobos Dunes and Commonweal and, to a certain extent, in Schooner Bay. The remaining populations did not exhibit differences in distortion at different paces. Distortion also varied significantly with frequency, mainly due to an increase in distortion at the highest (8 kHz) frequency (Tukey HSD $P < 0.05$). Overall, distortion was higher in rural than in urban locations, although there was variation among locations (Tables 2, 3). Commonweal territories experienced significantly higher levels of distortion than either Lake Merced or Battery East territories (Tukey HSD $P < 0.05$). We next examined variation between the rapid and fast paced data more closely by treating the locations as repeated measures of distortion at each of the ten frequencies. Locations varied significantly in distortion for rapid ($F = 9.73$, $df = 8$, $P = 0.01$) and fast paced signals ($F = 8.37$, $df = 8$, $P = 0.02$), with the Commonweal territories yielding the highest levels of distortion for both paces.

# Song: predictions and methods

## Predictions based on transmission profiles

There is significant variation among locations in aspects of signal degradation, yielding predictions for which locations are most likely to be limited in song structure by habitat transmission characteristics. Signals within the frequency range of NWCS songs experience higher levels of attenuation in Battery East than signals at any of the other five locations, followed by the other two urban locations. Thus, we would expect that Battery East songs may be most limited in frequency bandwidth to minimize signal degradation, with lower maximum frequency and higher minimum frequency than other locations. We did not find differences in distortion among urban locations, and therefore do not predict differences in trill rate due to acoustic adaptation. Among rural locations, Schooner Bay has the highest levels of attenuation at higher frequencies. Signals at lower frequencies experience more reverberation in Schooner Bay than in the other two rural locations. Thus, we predict that Schooner Bay songs will tend to have narrower bandwidths, with potentially lower maximum frequencies and/or higher minimum frequencies than the other rural songs. Signals transmitted in the Commonweal experienced the most distortion overall, but the least distortion at the middle pace (5 notes per second). Thus, we can predict that Commonweal songs will tend to be slower than songs from other locations. Lastly, we can also predict that across all locations, urban songs are more frequency limited than rural locations, due to more attenuation at high frequency and urban noise at low frequencies. Based on overall higher distortion in rural areas, we also would expect slower trills in rural birds compared to urban birds.
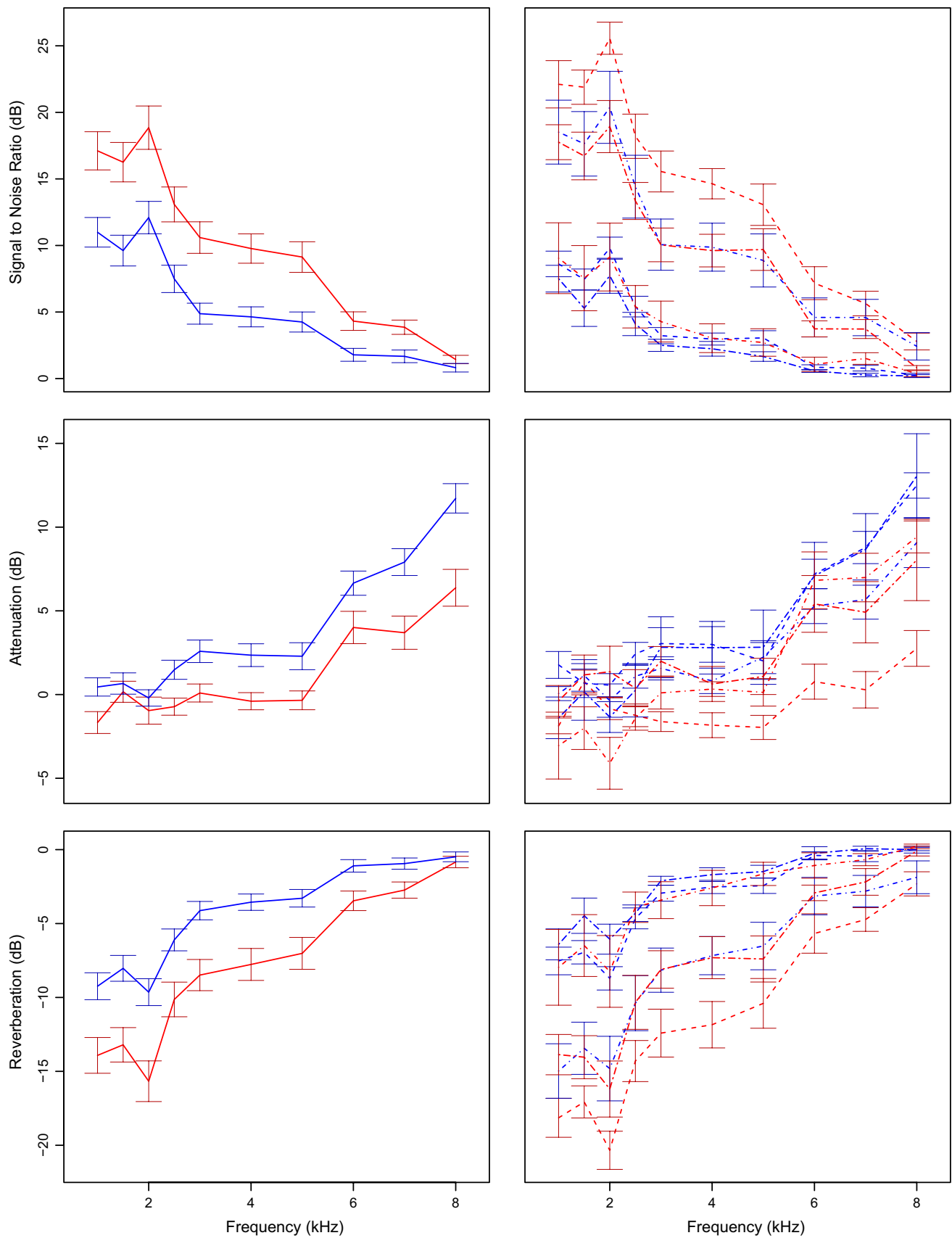
◀**Fig. 2** Relationship between the frequency of artificial sounds and **a** signal–noise ratio (SNR), **b** attenuation and **c** reverberation. Data from the two transmission heights are combined, as height did not explain a significant amount of variation in transmission data. Left panels denote urban (solid blue) and rural (solid red) averages with standard error bars for each frequency transmitted. Right panels show individual locations, with data presented for each rural location: Abbott's Lagoon (dashed red), Commonweal (two dash red), and Schooner Bay (dot dash red) and for each urban location: Battery East (two dash blue), Lake Merced (dashed blue), and Lobos Dunes (dot dash blue). Means and standard errors given for each frequency transmitted. Note that rural locations tend to have higher SNR, lower attenuation and less reverberation across locations as compared to urban locations (color figure online)

## Song methods

### Song data

We recorded songs of 128 territorial males including those defending territories where transmission experiments were conducted, during the breeding seasons of 2013 and 2014. [Battery East ($n=11$), Lake Merced ($n=24$), Lobos Dunes ($n=29$), Abbott's Lagoon ($n=19$), Schooner Bay ($n=7$), and Commonweal ($n=34$)]. We recorded using a Sennheiser omnidirectional ME62 microphone mounted on a SME PR-1000 parabolic microphone reflector, and a Marantz PMD 661 portable solid-state recorder. Songs were recorded with a sampling rate of 44.1-kHz WAV files at 16-bit precision. Birds in Battery East, Lobos Dunes, and Lake Merced sang the same song type, the San Francisco dialect (Baptista 1975). Commonweal birds sang the 'Clear' dialect, Schooner Bay males sang the 'Drake' dialect, and Abbott's Lagoon males sang Drake-McLure dialect hybrid songs (Baker et al. 1982). We analyzed song variation among the sampled locations to test the predictions generated by the sound transmission results.

For each individual, 1–18 songs with good signal–noise ratios were selected for analysis of song characteristics (mean ± SD 8.5 ± 3.4; Derryberry et al. 2016). We assigned songs to dialect by comparing dialect descriptions in Luther and Baptista (2010) and Baptista (1975). Songs were resampled at 25 kHz, high pass filtered at 1500 Hz to eliminate noise below the frequency range of NWCS song, and analyzed using SIGNAL v5 (Beeman 1998). We focused on the introductory whistle and the simple note trill, because these two song phrases are shared across dialects (Baptista 1975). The whistle has been suggested to be particularly salient in the context of song learning (Soha and Marler 2001) and the trill in discrimination of dialect differences in the context of male–male competition (Nelson and Soha 2004). We analyzed six acoustic measurements that captured variation in song among males: whistle dominant frequency, whistle duration, song maximum and minimum frequency, frequency bandwidth, and trill rate. We measured whistle

dominant frequency as the frequency at which the most sound energy was transmitted during whistle production from a smoothed power spectrum of each whistle. Trill rate was calculated as the average number of notes produced per second from oscillograms. Frequency bandwidth was calculated as the difference between the maximum and minimum frequencies (Hz) measured from a spectrogram (256 point fast Fourier transform (FFT), frequency resolution: 98 Hz, time resolution: 10.2 ms) at − 36 dB relative to the maximum level of the signal. We selected the − 36 dB threshold because it captured variation in frequency bandwidth while excluding background noise. Because frequency measures were taken at a fixed number of decibels below the peak amplitude, variation in frequency measures was not due to variation in song amplitude (Zollinger et al. 2012). All measurements were taken on each song and then averaged for each male.

### Song statistical analyses

A Bartlett's test for homogeneity of variance was significant ($P < 0.05$); thus, we reduced song characteristics with a principal components analyses (PCA). Song structure was compared among both habitat (urban/rural) and location using linear models for each principal component, followed by Tukey's HSD posthoc pairwise comparison tests. We analyzed song data for the three urban locations, as these belonged to the same dialect separately from the three rural locations. Finally, we combined song data across all six locations.

Prior to analyses, all variables were transformed as needed to meet assumptions of normality and homogeneity of variance. Statistical analyses were conducted in R 3.6.1 (R Development Core Team 2015).

## Results: song divergence

Song variables reduced down to three principal components with eigenvalues above 1. PC1 is positively loaded with song bandwidth and song maximum frequency, explaining 35.1% of the variation ($\chi^2 = 577.8$, $df = 14.76$, $P < 0.0001$; Table 4). PC2 is positively loaded with trill rate and negatively loaded with whistle length explaining 22.9% of variation ($\chi^2 = 512$, $df = 8.63$, $P < 0.0001$); while PC3 is loaded with song minimum frequency and whistle dominant frequency, explaining 20% of the variation ($\chi^2 = 474.6$, $df = 12.2$, $P < 0.0001$; Table 4).

Songs are significantly different among the three urban locations within the San Francisco dialect (PC1: $F_{2, 60} = 13.5$, $P < 0.0001$; PC3: $F_{2, 60} = 19.4$, $P < 0.0001$). As predicted by differences in transmission profiles among the three urban locations, Battery East has narrower song bandwidth—with

**Table 2** Effects of signal frequency and location on (A) attenuation, (B) reverberation, and (C) signal–noise ratio and (D) distortion of sound stimuli

| Sources of variation | *df* | SS | *F* ratio | *P* |
|---|---|---|---|---|
| (A) Attenuation | | | | |
| Location | 5 | 8.4 | 38.9 | **< 0.0001** |
| Frequency | 9 | 23.9 | 61.1 | **< 0.0001** |
| Location × frequency | 45 | 6.5 | 3.3 | **< 0.0001** |
| (B) Reverberation | | | | |
| Location | 5 | 35.8 | 217.8 | **< 0.0001** |
| Frequency | 9 | 51.5 | 174.4 | **< 0.0001** |
| Location × frequency | 45 | 2.78 | 1.88 | **0.0005** |
| (C) Signal–noise ratio | | | | |
| Location | 5 | 45.9 | 425.2 | **< 0.0001** |
| Frequency | 9 | 50.1 | 258.1 | **< 0.0001** |
| Location × frequency | 45 | 3.4 | 3.5 | **< 0.0001** |
| (D) Distortion | | | | |
| Location | 5 | 0.33 | 4.19 | **0.0008** |
| Frequency | 9 | 1.0 | 7.19 | **< 0.0001** |
| Location × frequency | 9 | 0.24 | 1.71 | 0.08 |

Bold indicates significance at the α = 0.05 level

Statistics are from linear models with main and two-way interaction effects. Ten transmission experiments were run in each urban and rural location, except Lake Merced ($n = 17$) and Schooner Bay ($n = 7$)

**Table 3** Effect of signal pace, frequency, and habitat on signal distortion (signal-to-signal ratio)

| Sources of variation | *df* | SS | *F* ratio | *P* |
|---|---|---|---|---|
| Pace | 2 | 0.42 | 13.38 | **< 0.0001** |
| Frequency | 9 | 0.93 | 6.58 | **< 0.0001** |
| Habitat | 1 | 0.12 | 7.86 | **0.0051** |
| Height | 1 | 0.05 | 3.2 | 0.07 |
| Pace × frequency | 18 | 0.29 | 1.04 | 0.40 |
| Pace × habitat | 2 | 0.03 | 1.04 | 0.35 |
| Pace × height | 2 | 0.02 | 0.72 | 0.49 |
| Frequency × habitat | 9 | 0.24 | 1.7 | 0.08 |
| Frequency × height | 9 | 0.11 | 0.77 | 0.64 |
| Habitat × height | 1 | 0.0 | 0.0 | 0.91 |
| Final model ($r^2 = 0.01$) | 4 | 0.59 | 9.5 | **< 0.0001** |

Bold indicates significance at the α = 0.05 level

Statistics are from linear models with main and two-way interaction effects. Final model includes only significant terms. Pace refers to the rate at which artificial sound stimuli were broadcast (rapid, fast, and slow). Signals were broadcast at ten different frequencies. Habitat refers to whether transmission experiments took place in urban or rural habitats. Transmission height was the height of the microphone (either 0.5 m or 1.5 m) above ground level

decreases in maximum frequency and increases in minimum frequency—compared to Lobos Dunes and Lake Merced (Tukey HSD < 0.05). Consistent with a lack of differences in distortion among urban locations, PC2 (e.g., trill rate) was

not significantly different between urban locations (all Tukey HSD > 0.05). Lake Merced has significantly higher values for PC3 than the other two locations (Tukey HSD < 0.05).

Songs are significantly different among the three rural locations (PC1: $F_{2, 52} = 3.75$, $P = 0.03$; PC3: $F_{2, 52} = 24.4$, $P < 0.0001$). Consistent with predictions based on the transmission profiles, Schooner Bay has a narrower bandwidth—with increased minimum frequency—as compared to Abbott's Lagoon and Commonweal (Tukey HSD ≤ 0.05). Counter to our predictions of slower songs in the Commonweal, PC2 (e.g., trill rate) was not significantly different among rural locations ($F_{2, 52} = 1.07$, $P = 0.35$; all Tukey HSD > 0.05). Consistent with lower frequencies experiencing more distortion in Schooner Bay compared to the other rural locations, we found that all rural locations differed significantly in PC3, such that Schooner Bay has the highest dominant frequency followed by Abbott's Lagoon, then Commonweal (Tukey HSD < 0.05).

We predicted that urban songs would occur in narrower frequency bandwidths and at lower maximum frequencies and higher minimum frequencies compared to rural songs. We did not find this pattern of song divergence for bandwidth and maximum frequency—urban songs are not different overall from rural songs in PC1 ($F_{1, 116} = 0.07$, $P = 0.78$). However, location does significantly predict PC1, where Battery East has the smallest bandwidth, lowest maximum frequency, and highest minimum frequency for urban locations, and Schooner Bay has the same trend for rural locations ($F_{5, 112} = 7.16$, $P < 0.0001$; Fig. 4).

We also predicted that rural songs would tend to be slower in trill rate and longer in whistle duration to minimize the effects of distortion. Rural songs are significantly slower in trill rate and longer in whistle duration (PC2: $F_{1, 116} = 193.4$, $P < 0.0001$) than songs from more urban locations.

If urban locations act like canyons, we would expect whistles to be lower pitched so they can travel further, similar to forest ecosystems. Urban songs are not significantly different from rural songs in PC3 ($F_{1, 112} = 0.01$, $P = 0.91$), but location does have a significant effect ($F_{5, 112} = 17.9$, $P < 0.0001$; Fig. 4), where Abbott's Lagoon and Schooner Bay have the highest dominant frequencies compared to other locations (Tukey HSD < 0.05). However, the loading of PC3 with both whistle dominant frequency and song minimum frequency may be masking a biologically significant effect. When whistle dominant frequency is examined alone, urban whistles are lower pitched than rural whistles ($F_{51, 126} = 19.4$, $P < 0.0001$), matching our prediction.

## Discussion

Environmental selection pressures on acoustic signals in urban habitats are not typical of natural habitats, such as increased flutter echo in urban canyons, leading to slower
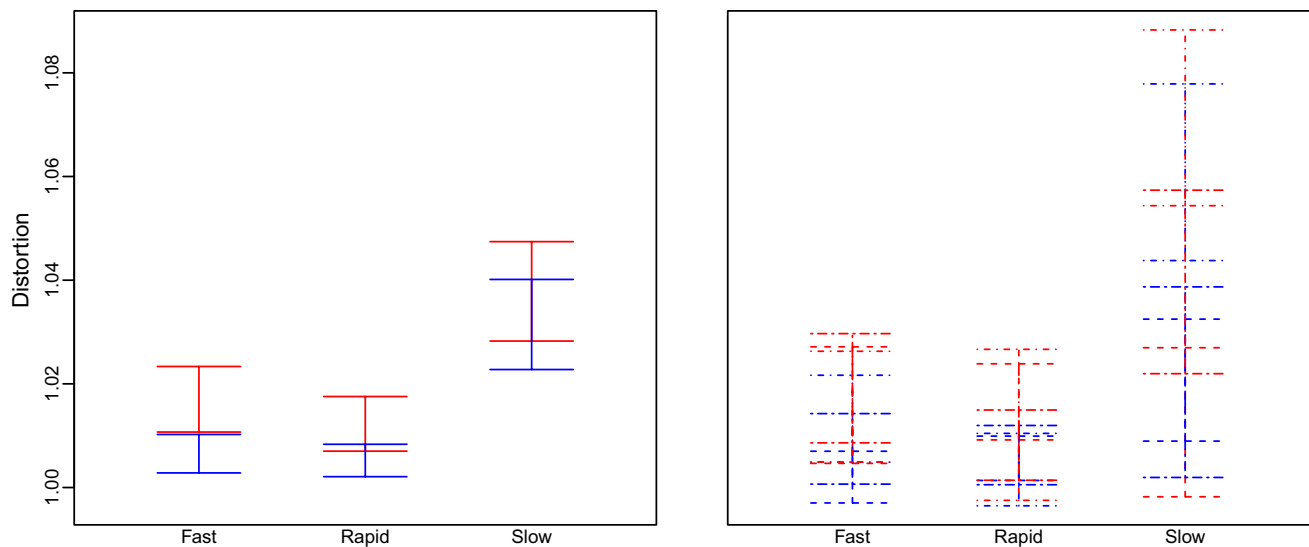
**Fig. 3** Left panel: rural locations (solid red) have more distortion than urban locations (solid blue; means and standard deviation), particularly at fast trill rates. Right panel: slow paced notes have the highest distortion across locations: Abbott's Lagoon (dashed red), Common-weal (two dash red), and Schooner Bay (dot dash red) and for each urban location: Battery East (two dash blue), Lake Merced (dashed blue), and Lobos Dunes (dot dash blue) (color figure online)

**Table 4** Song PCA loadings, where PC1 is positively loaded with song bandwidth, song maximum frequency, and negatively with minimum frequency

|  | PC1 | PC2 | PC3 |
|---|---|---|---|
| Song bandwidth | **0.99** | − 0.04 | − 0.001 |
| Song maximum frequency | **0.87** | 0.16 | 0.36 |
| Trill rate | 0.06 | **0.77** | 0.32 |
| Whistle length | 0.10 | **− 0.45** | 0.71 |
| Song minimum frequency | **− 0.57** | 0.38 | **0.60** |
| Whistle dominant frequency | − 0.18 | − 0.63 | **0.34** |

Bold indicates significance at the $\alpha = 0.05$ level

PC2 is positively loaded with trill rate and negatively loaded with whistle length, while PC3 is positively loaded with minimum frequency and whistle dominant frequency

attenuation and more reverberation of vocal signals in cities (Katti and Warren 2004; Dowling et al. 2011). A few studies have investigated the transmission properties of sound in urban areas at limited locations (Slabbekoorn et al. 2007; Gall et al. 2012); however, ours is one of the first to also investigate sound transmission within both urban and rural locations along a gradient of human activity (Mockford et al. 2011; Lazerte et al. 2015). Previous studies show that mountain and black-capped chickadee (*Poecile gambeli* and *atricapillus*) wideband and whistled call transmission is affected more by ambient noise rather than habitat type, where song SNR decreases with increasing noise, minimum frequency increases, and maximum frequency decreases (Lazerte et al. 2015). Rural great tit (*Parus major*) songs transmit

less effectively in urban areas (Mockford et al. 2011). We find that higher frequency signals suffer greater attenuation, reverberation and loss of SNR in urban than in rural locations; whereas, rural locations experience more distortion. Thus, acoustic adaptation predicts that urban songs should occur in a narrower bandwidth with lower maximum frequencies than rural songs and that rural songs should be slower than urban songs. We find that songs are diversifying in song structure between both habitats and locations in accordance with acoustic adaptation theory. Urban songs have shorter whistles and faster trills, while rural songs have longer whistles and slower trills. However, song evolution is not merely dichotomous—each location has a different effect on transmission and song divergence. While we quantified percent impervious cover, which is significantly higher in urban locations, other anthropogenic factors could affect song evolution. For example, we found reduced bandwidth, higher minimum frequency, and lower maximum frequency trends in urban locations, but one rural location also follows this pattern. Overall, our results suggest complex interactions between abiotic, biotic, and anthropogenic factors that shape acoustic divergence of songs.

Human development can vary greatly in its effect on the landscape. We tested multiple locations within both urban and rural areas to attempt to capture the diversity of habitats in each. We used a coarse measure of impervious surface to differentiate urban and rural locations. We provide evidence that urban and rural habitats do not all produce environmental pressures to the same degree, rather that locations produce pressures along a gradient. Within urban locations,
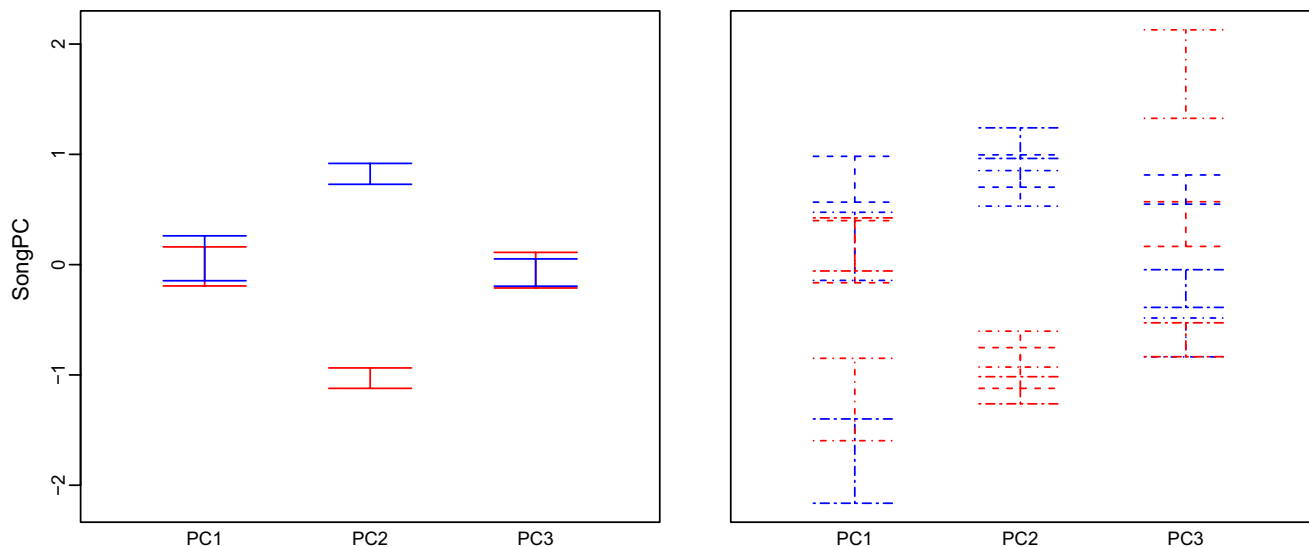
**Fig. 4** Left panel: urban locations have significantly faster trills and shorter whistles (PC2) than rural locations, while PC1 and PC3 do not significantly differ between urban (solid blue) and rural (solid red) locations. Right panel: PC1 (Song maximum frequency and band- width) are reduced in both urban Battery East (two dash blue) and rural Schooner Bay (dot dash red). There is significant divergence of songs between locations for PC3 (minimum frequency and whistle dominant frequency) (color figure online)

transmission properties of Battery East and Lobos Dunes produced different magnitudes of effects. Transmission properties of Lobos Dunes generally fell between Battery East and Commonweal data, suggesting that the degree of urbanization in a habitat can moderate the magnitude of signal–noise ratios, attenuation, and reverberation present within a signal. While our coarse (30 m), publicly available data on impervious surface reflect the dichotomy of urban and rural habitats, finer-scale habitat features and abiotic features (e.g. different vegetation types) also likely play a big part in why transmission properties differ between locations, and will be important to investigate in future studies.

Our study provides detailed quantification of sound transmission properties along an urban to rural gradient. In all locations, SNR ratios decrease as frequency increases, attenuation increases as frequency increases, and reverberation is greater at higher frequencies. These results generally follow patterns expected from classic sound transmission studies (Richards and Wiley 1980; Wiley and Richards 1982). Contrary to our predictions, faster paced tones were less distorted. Typically, faster paced notes exhibit the highest distortion due to reverberations from the first note increasing the amplitude of the second note. Our finding of the opposite pattern was unexpected and suggests that more work is needed on transmission of extremely slow paced signals. However, this finding may be related to faster trills evolving in our city locations.

Urbanization places multiple pressures on an acoustic signal resulting from unnatural structural components and increased human-generated ambient noise. Previous research has shown that the ambient noise levels and song minimum frequency of birdsongs increase with the level of urbanization (Slabbekoorn et al. 2007; Mockford et al. 2011; Slabbekoorn 2013; Derryberry et al. 2016). In our study, signal–noise ratios were higher in rural locations compared to urban locations across all frequencies, but especially at lower frequencies (Fig. 2) suggesting that acoustic signals have the potential to be masked to a greater degree by ambient noise in urban areas. This finding is consistent with previous measurements of background noise in this system, where Point Reyes has roughly a 10-dB decrease in background noise compared to San Francisco overall (Lee and MacDonald 2011, 2013), and locations vary individually within both urban and rural areas (Derryberry et al. 2016).

In natural environments less affected by human activity, higher frequencies are expected to be more attenuated (i.e., decreased amplitude) in closed canopy or windy conditions, such that the high range of modulated notes would be selected against (Richards and Wiley 1980; Wiley and Richards 1982; Price et al. 1988). Vocalizations in closed habitats typically have low frequency, narrow bandwidth notes (Morton 1975; Boncoraglio and Saino 2007). For example, calls within three bird families have 5% lower frequency in closed habitats than open habitats (Billings 2018). Our results show that attenuation was greatest in Battery East, an especially windy and anthropogenically noisy location, across all frequencies above 3000 Hz compared to the two other urban locations and all rural locations. Similar to Slabbekoorn, Yeh, and Hunt's (2007) findings that urban songs have higher minimum frequencies and reduced maximum

frequencies, Battery East songs had a reduced frequency bandwidth compared to all locations other than Schooner Bay, another particularly windy location, which also has a comparatively narrow bandwidth. These data suggest that birds in more urban habitats reduce the maximum frequency of their signals to reduce the effects of greater attenuation at higher frequencies. These data taken together with our SNR and attenuation measures also provide evidence that urban birds have a reduced frequency bandwidth to avoid the effects of reduced SNR occurring at lower frequencies in urban locations.

While forests are expected to absorb sound, "urban canyons" of cement and glass are more likely to further reverberate sounds (Warren et al. 2006), allowing them to travel further than in open habitats. While this habitat structure may enhance the spread of urban background noise, it may also benefit the first note of White-crowned Sparrow song, the pure-toned whistle, by increasing transmission distance. At least one study suggests that the communication distance of tonal notes may be further in open urban areas with impervious surfaces acting as sound channels that increase signal propagation (Gall et al. 2012). Urban areas also tend to produce multiple discrete echoes, reverberations, over concrete and glass in comparison to forested habitats (Slabbekoorn et al. 2007). These echoes have the potential to fill silent intervals between notes with sound and interfere with other notes; hence, we predicted that urban songs would be shorter in duration, the initial whistle may be lower pitched, and trill rates would be slower to avoid the effects of flutter echo. We found that urban whistle length was shorter and whistle pitch was lower in urban birds than in rural birds, matching our prediction. However, trill rates in urban areas are faster than rural areas, which does not match typical acoustic adaptation theory. Battery East songs had the fastest trill rates, followed by Lake Merced, linearly declining along the urbanization gradient to the Commonweal, which had the slowest rates. The Commonweal had the highest rates of distortion, which may explain why songs have diverged to have slower trills in the Commonweal as compared to the other locations.

However, increased trill rates for urban birds might instead be explained by the loss of bandwidth in urban locations. Both bandwidth and trill rate play a part in a sexually selected performance tradeoff in White-crowned Sparrows (Luther et al. 2015; Phillips and Derryberry 2017a, b). Therefore, if increased attenuation and reverberation favor reduced bandwidth in urban areas, increasing trill rate may help to recover vocal performance, such that a male can still signal vocal ability to potential mates and territorial rivals. However, a playback study in San Francisco found that males did not significantly differentiate between songs of different trill rates alone (Phillips and Derryberry 2018). Alternative explanations may be that fast trills reverberate

and distort such that they are perceived as long tonal note without internote intervals, which could increase signal detection distance. Once intended receivers move closer to distinguish trill notes, they may be able to properly assess performance. Lastly, trills may be faster due to random cultural drift, but further research is needed to disentangle the evolutionary drivers of fast trills in cities.

Interestingly, rural Schooner Bay often followed patterns exhibited by urban Battery East songs. While Schooner Bay is not urban, it was the most affected by human activities of the rural locations; during the time of these experiments, it had an active oyster farm with a public access hard pack road, houses, buildings, storefront, and boat traffic. Additionally, Schooner Bay may have microclimatic differences based on the shape of the lagoon, funneling wind and fog differently than either the Commonweal or Abbott's Lagoon locations, which were further inland from the coast. While future research should investigate the cause of these differences within rural locations, we suggest that these other factors, such as detailed impervious cover or other human disturbance to the landscape can influence transmission and song characteristics. That is, even small-scale anthropogenic changes to the environment have the potential to affect local signal divergence.

Our results suggest that NWCS in urban habitats have been under environmental selection pressures that have altered their characteristics from rural songs. Songs in urban areas are currently matched (i.e., transmit with greater clarity) to the transmission properties of an urban environment (Derryberry et al. 2016), overall providing support for acoustic adaptation to both urban noise and landscapes.

## Alternative explanations

Alternative explanations for song divergence should be considered. The acoustic adaptation hypothesis predicts that sound transmission, as well as habitat differences in ambient noise, may drive divergence (Slabbekoorn and Smith 2002). Ambient noise levels are different along the urban to rural gradient in this study with Battery East having the greatest and Commonweal having the least ambient noise (Derryberry et al. 2016). It is possible that the observed differences in song could reduce degradation by ambient noise in urban areas. In addition a higher minimum frequency in locations with increased low-frequency ambient noise should reduce masking from the anthropogenic noise and increase signal–noise ratios.

Nonadaptive, stochastic processes, such as genetic drift, could in part explain the signal divergence between urban and rural songs as well. Genetic drift affecting beak morphology could alter a bird's ability to create certain song characteristics, such as frequency bandwidth and speed of trills. However, beak morphology has been assessed along

this urban to rural gradient and no significant differences in bill morphology were found (Luther et al. unpublished data). Additionally, sexual selection via female mate choice could enhance the speed of divergence. NWCS use song as a primary mating signal, and therefore, differences in song play a role in female mate choice. The specific details of which song components female NCWS prefer have not been assessed, and there is certainly room here for future research.

Cultural evolution could also play a significant role in song divergence. Cultural evolution plays a part in NWCS song divergence in urban habitats (Moseley et al. 2018). For example, hand-reared NWCS learn songs less masked by noise, which is likely a main mechanism affecting cultural evolution, especially of masked frequencies (Moseley et al. 2018). Cultural evolution is one mechanism by which songs may become adapted to a local environment, and therefore, is consistent with the acoustic adaptation hypothesis. Further research is needed to determine the weights of acoustic and physical environments in song learning and ultimately song divergence.

## Conclusions

Numerous species, especially birds, rely on sound as a primary means of communication. Urbanization can detrimentally affect communication ability by interfering with signal saliency. Urban habitats put different selection pressures on acoustic signals, possibly leading to signal divergence and a lack of mate recognition. Signal divergence between urban and rural populations has the potential to lead to premating reproductive isolation on an evolutionary timescale, and thus, the factors that drive this divergence, such as urban structure and noise are important to understand. While the past 15 years have seen an uptick in the research in this area, a comprehensive understanding within each system and species should be made to understand the impact of human communities on the ecology and evolution of species that persist in urban environments, and whether patterns of sound transmission and acoustic adaptation are consistent across the world.

## Compliance with ethical standards

**Conflict of interest** The authors declare no conflict of interest.

**Ethical approval** All studies adhered to ethical guidelines approved by Tulane Institutional Animal Care and Use Committee and the National Park Service Institutional Animal Care and Use Committee. Permits were provided by U.S. Fish and Wildlife Service (MB679782-

1), the Bird Banding Laboratory (23900), and the California National Resources Agency (SC-6799), Golden Gate National Recreation Area (SCI-0007), San Francisco Parks and Recreation (041415), and by Point Reyes National Seashore (SCI-0016).

## References

Andersson MB (1994) Sexual selection. Princeton University Press, Princeton

Baker MC, Thompson DB, Sherman GL et al (1982) Allozyme frequencies in a linear series of song dialect populations. Evolution (N Y) 36:1020–1029. https://doi.org/10.2307/2408079

Baptista LF (1975) Song dialects and demes in sedentary populations of the white-crowned sparrow (*Zonotrichia leucophrys* nuttalli). Univ Calif Publ Zool 105:20

Beeman K (1998) Digital signal analysis, editing, and synthesis. In: Hopp SL, Owren MJ, Evans CS (eds) Animal acoustic communication: sound analysis and research methods. Springer, Berlin

Billings AC (2018) The low-frequency acoustic structure of mobbing calls differs across habitat types in three passerine families. Anim Behav 138:e74–e74. https://doi.org/10.1016/j.anbehav.2018.02.001

Blanchard BD (1936) Continuity of behavior in the Nuttall white-crowned sparrow. Condor 38:145–150

Blanchard BD (1941) The white-crowned sparrows (*Zonotrichia leucophrys*) of the Pacific seaboard: environment and annual cycle. University of California Press, Oakland, California, USA

Boncoraglio G, Saino N (2007) Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Funct Ecol 21:134–142. https://doi.org/10.1111/j.1365-2435.2006.01207.x

Boughman JW (2002) How sensory drive can promote speciation. Trends Ecol Evol 17:571–577

Chilton G, Baker MC, Barrentine CD, Cunningham MA (1995) White-crowned sparrow (*Zonotrichia leucophrys*). In: Poole A, Gill FB (eds) The birds of North America, vol 183. Academy of Natural Sciences. Philadelphia, Pennsylvania, USA, pp 1–28

Derryberry EP (2007) Evolution of bird song affects signal efficacy: an experimental test using historical and current signals. Evolution (N Y) 61:1938–1945. https://doi.org/10.1111/j.1558-5646.2007.00154.x

Derryberry EP (2009) Ecology shapes birdsong evolution: variation in morphology and habitat explains variation in white-crowned sparrow song. Am Nat 174:24–33. https://doi.org/10.1086/599298

Derryberry EP, Danner RM, Danner JE et al (2016) Patterns of song across natural and anthropogenic soundscapes suggest that white-crowned sparrows minimize acoustic masking and maximize signal content. PLoS One 11:e0154456. https://doi.org/10.1371/journal.pone.0154456

DeSante D, Baptista L(1989) Factors affecting the termination of breeding in Nuttall's White-crowned Sparrows. Wilson Bull 101(1):120–124

Dowling JL, Luther DA, Marra PP (2011) Comparative effects of urban development and anthropogenic noise on bird songs. Behav Ecol 23:201–209. https://doi.org/10.1093/beheco/arr176

Endler J (1992) Signals, signal conditions, and the direction of evolution. Am Nat139:125–153

Gall MD, Ronald KL, Bestrom ES, Lucas JR (2012) Effects of habitat and urbanization on the active space of brown-headed cowbird song. J Acoust Soc Am 132:4053–4062. https://doi.org/10.1121/1.4764512

Homer CG, Dewitz JA, Yang L, Jin S, Danielson P, Xian G et al (2011) Completion of the 2006 national land cover database for

the conterminous United States. Photogramm Eng Remote Sens 77:858–566

Homer CH, Fry JA, Barnes CA (2012). The national land cover database. U.S. Geological Survey Fact Sheet 2012-3020, p 4. https://doi.org/10.3133/fs20123020

Katti M, Warren PS (2004) Tits, noise and urban bioacoustics. Trends Ecol Evol 19:109–110. https://doi.org/10.1016/j.tree.2003.12.006

Kight CR, Hinders MK, Swaddle JP (2013) Chapter 5: acoustic space is affected by anthropogenic habitat features: implications for avian vocal communication. Ornithol Monogr 74:47–62. https://doi.org/10.1525/om.2012.74.1.47.3

Kight CR, Saha MS, Swaddle JP (2012) Anthropogenic noise is associated with reductions in the productivity of breeding Eastern Bluebirds (*Sialia sialis*). Ecol Appl 22:1989–1996. https://doi.org/10.1890/12-0133.1

Lazerte SE, Otter KA, Slabbekoorn H et al (2015) Relative effects of ambient noise and habitat openness on signal transfer for chickadee vocalizations in rural and urban green-spaces. Bioacoustics 24:233–252. https://doi.org/10.1080/09524622.2015.1060531

LaZerte SE, Slabbekoorn H, Otter KA et al (2016) Learning to cope: vocal adjustment to urban noise is correlated with prior experience in black-capped chickadees. Proc R Soc B 283:e27052–e27052. https://doi.org/10.1098/rspb.2016.1058

Lee C, MacDonald J (2013) Golden gate national recreation area: acoustical monitoring 2007/2008. Fort Collins, CO

Lee C, MacDonald J (2011) Baseline ambient sound levels in Point Reyes National Seashore. Los Angeles

Luther D, Baptista L (2010) Urban noise and the cultural evolution of bird songs. Proc R Soc B 277:469–473. https://doi.org/10.1098/rspb.2009.1571

Luther DA, Phillips J, Derryberry EP (2016) Not so sexy in the city: urban birds adjust songs to noise but compromise vocal performance. Behav Ecol 27:332–340. https://doi.org/10.1093/beheco/arv162

Luther DA, Phillips J, Derryberry EP (2015) Not so sexy in the city: urban birds adjust songs to noise but compromise vocal performance. Behav Ecol. https://doi.org/10.1093/beheco/arv162

Marler P, Tamura M (1962) Song "Dialects" in three populations of white-crowned sparrows. Condor 64:368–377

Mewaldt RL, King JR (1977) The annual cycle of white-crowned sparrows (*Zonotrichia leucophrys nuttalli*) in coastal California. The Condor 79:445–455

Mockford EJ, Marshall RC, Dabelsteen T (2011) Degradation of rural and urban great tit song: testing transmission efficiency. PLoS One 6:e28242. https://doi.org/10.1371/journal.pone.0028242

Morton ES (1975) Ecological sources of selection on avian sounds. Am Nat 109:17–34

Moseley DL, Derryberry GE, Phillips JN et al (2018) Acoustic adaptation to city noise through vocal learning by a songbird. Proc R Soc B 285:20181356

Nelson DA, Soha JA (2004) Perception of geographical variation in song by male Puget Sound white-crowned sparrows, *Zonotrichia leucophrys* pugetensis. Anim Behav 68:395–405. https://doi.org/10.1016/j.anbehav.2003.08.027

Nemeth E, Dabelsteen T, Pedersen SB, Winkler H (2006) Rainforests as concert halls for birds: are reverberations improving sound transmission of long song elements? J Acoust Soc Am 119:620. https://doi.org/10.1121/1.2139072

Peters S, Derryberry EP, Nowicki S (2012) Songbirds learn songs least degraded by environmental transmission. Biol Lett 8:736–739. https://doi.org/10.1098/rsbl.2012.0446

Phillips JN, Berlow M, Derryberry EP (2018) The effects of landscape urbanization on the gut microbiome: an exploration into the gut of urban and rural white-crowned sparrows. Front Ecol Evol 6:148. https://doi.org/10.3389/fevo.2018.00148

Phillips JN, Derryberry EP (2017a) Equivalent effects of bandwidth and trill rate: support for a performance constraint as a competitive signal. Anim Behav 132:209–215. https://doi.org/10.1016/j.anbehav.2017.08.012

Phillips JN, Derryberry EP (2017b) Vocal performance is a salient signal for male–male competition in White-crowned Sparrows. Auk 134:564–574. https://doi.org/10.1642/AUK-17-2.1

Phillips JN, Derryberry EP (2018) Urban sparrows respond to a sexually selected trait with increased aggression in noise. Sci Rep 8:7505. https://doi.org/10.1038/s41598-018-25834-6

Price MA, Attenborough K, Heap NW (1988) Sound attenuation through trees: measurements and models. J Acoust Soc Am 84:1836–1844. https://doi.org/10.1121/1.397150

R Development Core Team R (2015) R: a language and environment for statistical computing

Richards DG, Wiley RH (1980) Reverberations and amplitude fluctuations in the propagation of sound in a forest: implications for animal comunication. Am Nat 115:381–399

Slabbekoorn H (2013) Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds. Anim Behav 85:1089–1099. https://doi.org/10.1016/j.anbehav.2013.01.021

Slabbekoorn H, den Boer-Visser A (2006) Cities change the songs of birds. Curr Biol 16:2326–2331. https://doi.org/10.1016/j.cub.2006.10.008

Slabbekoorn H, Ellers J, Smith (2002) Birdsong and sound transmission: the benefits of reverberations. Condor 104:564–573

Slabbekoorn H, Yeh P, Hunt K (2007) Sound transmission and song divergence: a comparison of urban and forest acoustics. Condor 109:67–78

Soha JA, Marler P (2001) Cues for early discrimination of conspecific song in the white-crowned sparrow (*Zonotrichia leucophrys*). Ethology 107:813–826. https://doi.org/10.1046/j.1439-0310.2001.00713.x

Swaddle JP, Francis CD, Barber JR et al (2015) A framework to assess evolutionary responses to anthropogenic light and sound. Trends Ecol Evol 30:550–560. https://doi.org/10.1016/j.tree.2015.06.009

Tobias JA, Aben J, Brumfield RT et al (2010) Song divergence by sensory drive in Amazonian birds. Evolution (N Y) 64:2820–2839

Warren PS, Katti M, Ermann M, Brazel A (2006) Urban bioacoustics: it's not just noise. Anim Behav 71:491–502. https://doi.org/10.1016/j.anbehav.2005.07.014

Wiley RH (2006) Signal detection and animal communication. Adv Study Behav 36:217–247. https://doi.org/10.1016/S0065-3454(06)36005-6

Wiley RH, Richards DG (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. In: Kroodsma DE, Miller EH, Ouellet H (eds) Acoustic communication in birds. Academic Press, New York

Wiley RH, Richards DG (1978) Physical constraints on acoustic communication in the atmosphere: implications for the evolution of animal vocalizations. Behav Ecol Sociobiol 3:69–94

Zollinger SA, Podos J, Nemeth E et al (2012) On the relationship between, and measurement of, amplitude and frequency in birdsong. Anim Behav 84:e1–e9. https://doi.org/10.1016/j.anbehav.2012.04.026

## Affiliations

**Jennifer N. Phillips[1,2]** · **Catherine Rochefort[1,3]** · **Sara Lipshutz[1,4]** · **Graham E. Derryberry[5]** · **David Luther[6,7]** · **Elizabeth P. Derryberry[5]**

[1] Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, LA 70118, USA

[2] Department of Biological Sciences, California Polytechnic State University, San Luis Obispo, CA 93407, USA

[3] Department of Psychology, Southern Methodist University, Dallas, TX 75275-0240, USA

[4] Department of Biology, Indiana University, Bloomington, IN 47405, USA

[5] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA

[6] Biology Department, George Mason University, Fairfax, VA 22030, USA

[7] Smithsonian Mason School of Conservation, 1500 Remount Road, Front Royal, VA 22630, USA