Syntactic properties of male courtship song in the lesser short-tailed bat, *Mystacina tuberculata*

Kathleen Collier [1]* and Stuart Parsons [1,2]

[1]School of Biological Sciences, University of Auckland, Auckland, New Zealand, [2]School of Science, Technology and Engineering, University of the Sunshine Coast, Maroochydore, QLD, Australia

Bats (Mammalia: Chiroptera) have sophisticated acoustic abilities adapted to many uses, including both echolocation and social vocalisations. Social vocalisations are used in a wide variety of contexts and vary greatly in acoustic arrangement and complexity. Among the most intricate are the courtship songs that males of certain species use to attract mates and advertise their qualities. To date, however, few studies have examined the phonological construction of bat songs or made a detailed assessment of the syntax used to combine different song elements. Here, we examine the syntactic construction of courtship songs in the New Zealand lesser short-tailed bat, *Mystacina tuberculata*, a highly vocal, lek-breeding species with exceptionally high song-output rates. We hypothesised that song construction in this species is both hierarchical and non-random, and demonstrates a high degree of individual variation, potentially allowing for individual recognition. We recorded the courtship songs of nine male bats and used manual classification of song components to examine the song structure. Here we examine whether the deployment of different song components is dependent on their position, and also determine the transition probabilities between different components. We find that the frequency of song-element production and the distribution of elements within songs are non-random at both the individual and population level, and that the number of elements used per phrase differs between individuals. Overall, we demonstrate that *M. tuberculata* songs are hierarchically constructed and employ phonological syntax to build syllables and phrases. We further show that bats employ high structural similarity and conservatism in the construction of syllables, while retaining a capacity for versatility and innovation that allows for considerable individual variation and, likely, individual recognition.

KEYWORDS

Chiroptera (bats), social communication, courtship song, phonological syntax, *Mystacina tuberculata*

Introduction

Bats (Mammalia: Chiroptera) have long been recognised for their highly developed acoustic abilities. While they are better known for their use of echolocation than for social communication, they do in fact produce a wide variety of vocalisations, many of which occur within the human-audible spectrum (though they may also contain ultrasound components). Social vocalisations in bats have now been recorded in a wide range of contexts, and include calls for resource defence (Barlow and Jones, 1997); infant isolation (Balcombe, 1990; Bohn et al., 2007, 2008; Prat et al., 2016); conspecific contact (Bohn et al., 2008, 2013); territorial aggression (Bohn et al., 2008; Behr et al., 2009; Prat et al., 2016); human interaction (Bohn et al., 2008); and male courtship vocalisations, including song (Behr and von Helversen, 2004; Davidson and Wilkinson, 2004; Russ and Racey, 2007; Bohn et al., 2008, 2009; Toth and Parsons, 2018). Moreover, both echolocation and social calls have been shown to communicate aspects of individual identity such as sex, age, and breeding status, and may even be directly addressed from an individual to an intended recipient (Prat et al., 2016).

Although it has long been under-studied by comparison with other taxa, attention is increasingly turning to the singing behaviour of bats. Singing in bats is likely to have been fostered by the same factors that caused its proliferation amongst birds – namely, aerial habits, long travel distances, and large social groups (Smotherman et al., 2016). As such, it is likely that many bat species engage in singing behaviour to a greater or lesser extent. While the number of species known to produce song is currently low, it is highly likely that more will be added as the resources available for monitoring and recording bat behaviour improve (Smotherman et al., 2016).

Of those species which *are* known to sing, many produce songs that rival those of songbirds in terms of output and complexity (Smotherman et al., 2016; Vernes and Wilkinson, 2020). While the complexity of vocalisations that are considered "song" varies greatly both within and between taxa, it is indisputable that many bat songs are highly sophisticated, requiring considerable vocal plasticity (Scharff and Petri, 2011; Vernes, 2017). The distinction between song, language and other vocalisations such as calls is poorly defined, but is generally agreed to be behavioural rather than mechanistic. Songs are often described as "performative," involving a display of some kind, and are typically linked to a specific behavioural function, such as a courtship or territorial display (Catchpole and Slater, 2008).

The most complex songs are combinatorial: – that is, they are composed of hierarchical elements which are combined in different ways to generate meaning (Marler, 1977; Hailman and Ficken, 1986; Berwick et al., 2012; Engesser and Townsend, 2019). The rules governing the order in which elements can be arranged are known as syntax (Marler, 1977). Combinatorial

syntax allows a multitude of possible meanings to be generated from a relatively small number of base sounds, allowing a wide range of concepts to be expressed, as in human language (Berwick et al., 2012). There is a distinction between phonological syntax (the rules governing the assembly of smaller vocal units into larger ones) and lexical syntax (the corresponding changes in meaning; Marler, 1998). Phonological syntax is common in animal communication systems, but evidence for lexical syntax is scarcer (Marler, 1998). In other words, while vocalisations may be constructed according to particular rules (e.g., note B must follow note A) there is little evidence that variations on this formula will change the meaning of the vocalisation (i.e., A-B is unlikely to have a significantly different meaning from A-B-C, or A-B-A-B). While there is relatively little evidence for either kind of syntax in bats, it is likely that phonological syntax, at least, is relatively widespread (Bohn et al., 2009, 2013; Smotherman et al., 2016).

Combinatorial syntax allows a wide scope for individual expression, and may thus contain features that honestly signal a singer's attributes, aid individual recognition, and facilitate mate choice (Catchpole and Slater, 2008). Female preference for a particular type of song construction has been demonstrated in the greater sac-winged bat, *Saccopteryx bilineata*, where males that use a higher proportion of unique composite syllables have more females roosting in their territories (Davidson and Wilkinson, 2004). Consistent differences in syllable construction between individuals have also been shown in a number of species, including *S. bilineata* (Behr and von Helversen, 2004), *Pipistrellus nathusii* (Russ and Racey, 2007; Jahelková et al., 2008), and *Mystacina tuberculata* (Toth and Parsons, 2018). Individual-specific differences are thought to enable females to identify particular males, thus facilitating mate choice.

The courtship songs of the New Zealand short-tailed bat (*M. tuberculata*) provide an excellent opportunity for the study of syntactic complexity in free-living bats. Bat song is relatively under-studied, and little is known about how it evolved, its proximate mechanisms, how it is learned or acquired, or how it is constructed phonologically (Smotherman et al., 2016). *Mystacina tuberculata* is a particularly interesting species in which to study song construction, as it employs a great diversity of combinatorially constructed syllables and has one of the highest sustained song outputs ever recorded in a bat (Toth and Parsons, 2018). *Mystacina tuberculata* are lek breeders, with males selecting small cavities in trees as display roosts (Carter and Riskin, 2006) and performing complex courtship songs to attract females (Toth and Parsons, 2018). Display sites are aggregated in the vicinity of maternity colonies, and males display there between ∼November and May by singing and scent marking (Toth et al., 2015).

In this paper, we examine the phonological structure of male *M. tuberculata* courtship song. We aim to demonstrate that *M. tuberculata* song is hierarchically constructed according to syntactic rules and that male *M. tuberculata* possess unique

repertoires that may allow female bats to identify and assess their individual attributes. To do this, we test the following hypotheses at the population and individual level: (i) that the frequency of song element production is non-random; (ii) that the number of component elements used to create larger constructs (syllables and phrases) differs between males; (iii) that the position of song elements within larger constructs is non-random; and (iv) that the transitions between song constructs are non-random.

Materials and methods

Male *M. tuberculata* were recorded at their display roosts in the Pikiariki Ecological Area of Pureora Forest Park (38°26′S, 175°39′E), central North Island, New Zealand during the summer breeding seasons of 2017 and 2018. Recordings were made using a Bruel and Kjaer $\frac{1}{4}$" Free-field Microphone (Type 4939) linked to a Sound Devices 722 digital recorder (Sound Devices, Reedsburg, WI, United States). Songs were recorded within 10 m of the singing roost tree using a sampling rate of 192 kHz with 24-bit precision. The majority of individuals in the population are PIT-tagged, and the identity of each recorded male was confirmed using a biomark HPR Plus automatic PIT tag reader (Biomark, Boise, ID, United States) mounted over the entrance to the singing roost. Recordings from nine individual males were selected as suitable for syntactic analysis. Each recording was >10 min in length, but owing to the exceedingly high song output of *M. tuberculata,* only 2 min from each song were analysed in detail. We used RavenPro 1.5 (Cornell Lab of Ornithology, Ithaca, NY, United States) to visually classify song elements. Spectrograms were generated using 1,024-sample discrete Fourier transformations with 95% overlap and a Hann window (frequency resolution 188 Hz, temporal resolution 2.7 ms).

The terminology used to describe song components varies widely both within and between biological disciplines, and also differs from that used by linguists. The terms used by birdsong researchers, for example, are rarely applicable to the songs of insects, amphibians, cetaceans, primates, or bats. For this reason, we here present a table in which some of the more commonly used terms are given along with examples of their use (**Table 1**). This is not intended as an exhaustive list, but we hope that it will offer some clarity to aid future interpretation.

Mystacina tuberculata song includes four basic elements: – trills (A), tones (B), upsweeps (C), and downsweeps (D; **Figure 1**). These were originally described by Toth and Parsons (2018), though using the term "notes" rather than elements. We also include here a fifth element, clicks (E). Clicks are of very short duration, appearing in spectrograms as vertical or near-vertical frequency bands, and typically occur in rapid bursts, or singly as a tail attached to the end of another element sequence. In addition to these five elements, we include a further

category (F) for elements that did not fit easily into any other classification or were too indistinct to make out.

Elements are acoustically distinct components that may be combined to form syllables (e.g., upsweep-tone) or may occur independently as syllables in their own right (**Figure 1**). Syllables are discrete units of song, surrounded by silences of at least 1 ms (Kanwal et al., 1994). Syllables may likewise occur independently, or may be combined with others to form phrases. A phrase is a segment of one or more syllables in which the silent period between any two syllables is roughly similar, and always less than the total duration of those two syllables (Kanwal et al., 1994; Bohn et al., 2008, 2013; Wiley, 2018). In *M. tuberculata,* phrases are generally separated by silences of ∼20 ms (Toth and Parsons, 2018).

As the basic elements of *M. tuberculata* song have already been described (Toth and Parsons, 2018), we have concentrated on describing the rules used to combine song components (Mitani and Marler, 1989; Bohn et al., 2009). Our primary unit of analysis is the phrase (n = 4767), but these are annotated throughout so as to allow for smaller-scale analyses by element or syllable.

We assessed the structure of songs using four methods. First, we calculated the occurrence frequency with which song elements were produced across all our study individuals, irrespective of the elements' position or acoustic characteristics. We used Pearson's chi-square tests to compare the rates at which the four most common elements occurred (observed frequency) to the rates expected by chance (i.e., if all elements were produced with equal likelihood; expected frequency). Less common elements were excluded as they typically occurred at frequencies lower than 5%, which would have violated statistical assumptions. Because certain of these elements were still used at low rates by particular individuals, we also employed Fisher's exact tests (simulated *P*-value based on 2,000 replicates). We considered the observed proportions to be the natural rate of element production among the population. We then compared this natural rate to the frequency with which song elements were produced by each particular individual.

Second, we used Kruskal–Wallis tests to determine whether individuals differed in the number of elements and syllables they used to construct phrases. We then conducted pairwise comparisons using Wilcoxon rank sum tests with continuity corrections to identify how individuals differed from one another.

Third, we examined the positioning of elements within syllables and phrases. Pearson's chi-squared tests were conducted using the four most common elements to test for associations between song position and element frequency (Mitani and Marler, 1989). We also employed Fisher's exact tests (simulated *P*-value based on 2,000 replicates). The proportion of times an element occurred in a particular position was compared to the proportion anticipated if elements were distributed equally (random) or according to their distribution

within the population as a whole (natural rate). The positions used in analyses were first (the element used to begin a syllable or phrase), last (the element used to end a syllable or phrase), isolated (elements occurring alone and surrounded by silence), and intermediate (all elements occurring in a non-start/end position within a composite syllable or phrase).

Finally, we examined the transition probabilities between element, syllable, and phrase types within each song (i.e.,

the frequency with which a particular song construct was followed by another; Mitani and Marler, 1989; Bohn et al., 2009). We used Mantel *r*-test comparisons of structural similarity to examine first-order (element-element) transitions, but did not statistically compare the second-order (syllable-syllable) or third-order (phrase-phrase) transitions owing to the enormous quantity of syntactic possibilities and the high degree of individual variation expressed. Matrices were set up

TABLE 1 Definitions of terms.

Term	Definition	Synonyms and sub-categories
Syntax	A set of rules for assembling units into larger units. For example, the temporal arrangements of acoustic units within a song (Marler, 1977).	*Phonological syntax* (rules governing the assembly of smaller vocal units into larger ones); *lexical syntax* (rules governing the corresponding changes in meaning; Marler, 1977).
Element	Acoustically distinct components which may be combined to form syllables (e.g., upsweep-tone) or may occur independently as syllables in their own right (e.g., upsweep; Toth, 2016)	*Simple syllable, subsyllable* and *composite syllable* (Kanwal et al., 1994); *note* (Bohn et al., 2009); *phoneme* in linguistics (Hailman and Ficken, 1986; Wiley, 2018)
Syllable	A discrete unit of song, surrounded by silences of at least 1 ms (Kanwal et al., 1994; Behr and von Helversen, 2004; Bohn et al., 2008, 2009, 2013)	*Note* (Mitani and Marler, 1989); *call type* (Melendez et al., 2006); *morpheme* in linguistics (Hailman and Ficken, 1986; Wiley, 2018)
Phrase	A segment of one or more syllables in which the silent period between any two syllables is roughly similar and always less than the total duration of those two syllables (Kanwal et al., 1994; Bohn et al., 2008, 2013; Wiley, 2018). For our purposes, separated in *M. tuberculata* by silences of ∼20 ms (Toth and Parsons, 2018)	*Syllable train* (Kanwal et al., 1994) or *simple phrase* (Bohn et al., 2008, 2009) when all syllables in the phrase are of the same type; *complex phrase* (Bohn et al., 2008, 2009) when syllables are of different types; *segment* (Kroodsma, 1977); *motif* (Bohn et al., 2009; Scharff and Petri, 2011); *call* (Behr and von Helversen, 2004)
Song-type	The sequence in which phrases are combined. Sequential repetitions of phrases are omitted, so the phrase sequence chirp-trill-trill-chirp belongs to the chirp-trill-chirp song type (Bohn et al., 2013).	As distinct from a *song-variant*, in which sequential repetitions are not omitted (Bohn et al., 2009)
Call	A syllable or phrase, usually brief, which serves a social function such as alerting or locating conspecifics; usually innate (Barón Birchenall, 2016)	Alarm call, social call, etc.; vocalisation.
Song	A vocal display in which multiple types of syllables and phrases are delivered in sustained performances; usually learned (Mitani and Marler, 1989; Bohn et al., 2008; Barón Birchenall, 2016).	Territorial song, courtship song, etc.
Language	A system of communication consisting of elements combined according to syntactic rules; used for exchanges of information between individuals (Wiley, 2018).	Spoken language, sign language, etc.
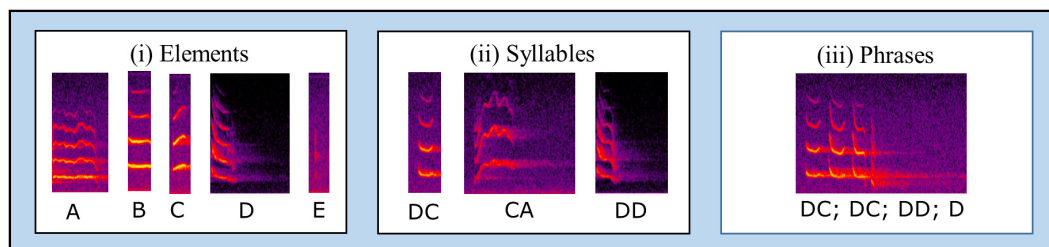
FIGURE 1
The construction of *Mystacina tuberculata* song. There are five basic elements (i): trills (A), tones (B), upsweeps (C), downsweeps (D), and clicks (E). These can be combined in a variety of ways to form syllables (ii) such as "downsweep-upsweep" (DC), "upsweep-trill" (DA), or "downsweep-downsweep" (DD); elements may also stand alone as syllables in their own right. Syllables can be combined to make phrases (iii), such as "downsweep-upsweep; downsweep-upsweep; downsweep-downsweep; downsweep" (DC; DC; DD; D), though syllables may equally stand alone as phrases in their own right. Phrases are typically separated from one another by ∼20 ms of silence. The phrase is the basic unit of analysis, and a series of phrases delivered in sequence is a song.

as "first element by second element" and compared between individuals.

Analyses were conducted in R version 3.5.3 using base packages (R Core Team, 2019). Individual bats were given names during tracking in the field, which we retain here for ease of reference. Unless stated otherwise, means are given ± Standard Error, and the significance threshold is set at $P < 0.05$. In cases where we conducted multiple tests of a shared hypothesis (e.g., that element production by each of nine individuals differed from a particular rate), Bonferroni corrections were applied.

Results

The mean number of phrases annotated per individual from a 2-min song sample was 529.7 ± 26.3. The breakdown of phrases per individual was as follows: Antoninus, $n = 732$; Commodus, $n = 744$; Domitian, $n = 547$; Hadrian, $n = 466$; Lucius, $n = 701$; Marcus, $n = 860$; Nero, $n = 335$; Otho, $n = 258$; and Trajan, $n = 154$.

Frequency of song element production

The frequency with which song-elements were produced at the population level differed from that expected if elements were produced with equal likelihood ($\chi^2 = 13.9$, $df = 3$, $P = 0.003$; Fisher's $P = 0.003$). A total of 12,740 song elements were classified, with upsweeps and downsweeps being the most commonly used (Figure 2). These made up almost exactly equal proportions of total sounds analysed, with downsweeps accounting for 4,060 of all classified sounds (31.8%) and upsweeps for 4,055 (31.8%). Trills were the next most-common song element, accounting for 20.7% of all classifications. Other elements contributed relatively little overall, though were more heavily used by some individuals than others. Tones contributed 4.6% overall and clicks 4.2%. The final 6.8% was made up of elements that either did not fit the existing classification or were too indistinct to label with certainty.

In six of the nine individuals analysed, upsweeps and downsweeps were the two most commonly used element types (Figure 2). Each of these elements typically contributed 20–40% of an individual's total repertoire. In three bats (Commodus, Hadrian, and Nero), the upsweep was the most commonly used element (35.8–41.4%) and the downsweep the second-most common (24.9–36.1%). In one bat (Otho), upsweeps and downsweeps both contributed 36.4% of the total repertoire, while in two others (Trajan and Antoninus), the downsweep was the most common (29.8 and 31.8%), and the upsweep the second-most common (22.6 and 29.8%). In all but one of these individuals the third most-commonly used element was the trill (12.8–16.1% of total repertoire). The one exception
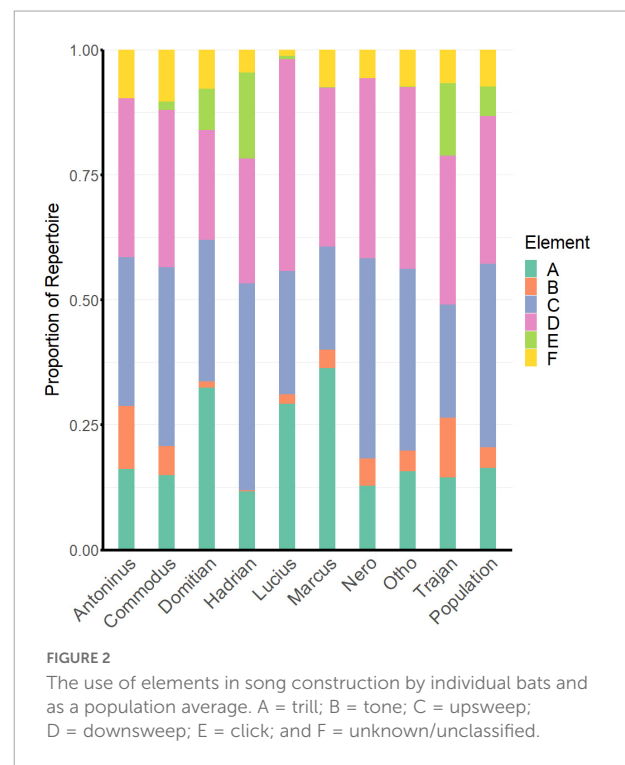
was Hadrian, whose song contained a high proportion of clicks (17.2%), with trills (11.6%) falling into fourth place.

The three remaining bats displayed different song patterns. In one (Lucius), downsweeps were the most commonly used element (42.4%), but trills were the second-most common (29.1%). In two others (Domitian and Marcus), trills made up the highest proportion of the song (32.4 and 36.3%, respectively), followed by upsweeps (Domitian; 28.3%) and downsweeps (Marcus; 31.8%).

While tones, clicks, and unclassified elements did not contribute a large amount to the analysis overall, they did add considerable variation to individuals' repertoires. Antoninus and Trajan used a higher proportion of tones (12.6 and 12.0%, compared to 0.3–5.9% for other bats), while Hadrian, Trajan, and Domitian included clicks in their songs at far higher rates (17.2, 14.4, and 8.3%, respectively), than other bats did (0.0–1.8%). The proportion of indistinct or unclassified elements in an individual's song ranged from 1.3% (Lucius) to 10.4% (Commodus).

Number of song components used in phrases

The mean number of elements used per phrase was 2.7 ± 0.03 (Figure 3A), while the mean number of syllables per phrase was 1.9 ± 0.0 (Figure 3B). The greatest proportion of sounds produced at the population level were single elements

FIGURE 2
The use of elements in song construction by individual bats and as a population average. A = trill; B = tone; C = upsweep; D = downsweep; E = click; and F = unknown/unclassified.

delivered in isolation (33.6% of all sounds). Two-element and three-element phrases made up 28.6 and 17.2%, respectively. Four-element (7.1%), five-element (4.2%), and six-element (2.9%) phrases were all used at least once by every individual analysed, and a range of more complex sounds were also employed, though at relatively low frequencies. The maximum number of elements employed in the construction of a single phrase was 16 (0.08% of all sounds analysed). It is worth noting that single-element phrases may still be of long duration and acoustically dominant; *ergo*, the syntactic complexity of a phrase does not necessarily have any effect on a bat's duty cycle.

The mean number of elements used per phrase by an individual ranged from 1.7 ± 0.09 (Trajan) to 4.3 ± 0.13 (Hadrian; **Figure 3A**), while the mean number of syllables used ranged from 1.2 ± 0.04 (Trajan) to 3.5 ± 0.1 (Hadrian; **Figure 3B**). Kruskal–Wallis tests revealed significant differences in the number of elements ($\chi^2 = 576.9$, *df* = 8, *P* < 0.0001) and syllables ($\chi^2 = 699.9$, *df* = 8, *P* < 0.0001) used by individuals.

Elements per phrase

Of the 36 possible pairwise comparisons between individuals, 26 pairs showed significant differences (*P* < 0.001) in the number of elements used to construct phrases. Hadrian differed from all other individuals, while Commodus, Lucius, and Nero all shared similarities with one another, and Otho shared similarities with the remaining four individuals (Antoninus, Domitian, Marcus, and Trajan). While the

differences between these five individuals appear relatively slight (**Figure 3A**), they are still dissimilar enough to yield statistical significance in three cases (Antoninus-Trajan; Trajan-Domitian; and Antoninus-Marcus).
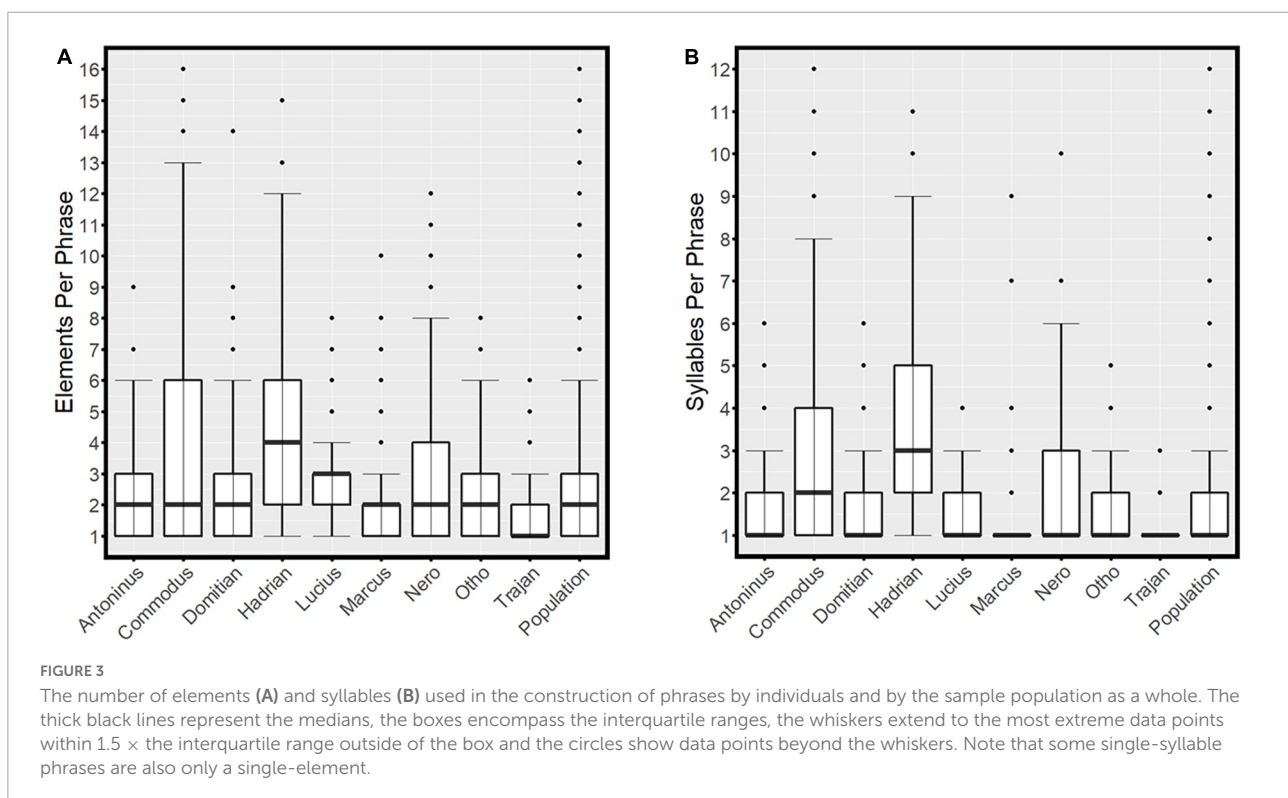
Syllables per phrase

Of the 36 possible pairwise comparisons, 29 pairs showed significant differences (*P* < 0.001) in the number of syllables used to construct phrases. Commodus and Hadrian differed from all other individuals, while Marcus shared similarities with Trajan, Antoninus with Nero, Domitian with Lucius, and Otho with three other individuals (Antoninus, Domitian, and Lucius).

Positioning of song elements within larger constructs

Syllables

The positioning of song elements controls the range of syllables or phrases used by a given individual. At the population level, the distribution of first ($\chi^2 = 25.2$, *df* = 3, *P* < 0.0001; Fisher's *P* = 0.0005), last ($\chi^2 = 31.7$, *df* = 3, *P* < 0.0001; Fisher's *P* = 0.0005), and isolated ($\chi^2 = 11.9$, *df* = 3, *P* = 0.0075; Fisher's *P* = 0.0095) elements in syllables was non-random (i.e., certain elements occurred more often in a particular position than expected by chance). The distribution of intermediate elements, however, did not differ from random ($\chi^2 = 4.5$, *df* = 3, *P* = 0.2; Fisher's *P* = 0.2). When compared to the natural rate of element

FIGURE 3
The number of elements **(A)** and syllables **(B)** used in the construction of phrases by individuals and by the sample population as a whole. The thick black lines represent the medians, the boxes encompass the interquartile ranges, the whiskers extend to the most extreme data points within 1.5 × the interquartile range outside of the box and the circles show data points beyond the whiskers. Note that some single-syllable phrases are also only a single-element.

expression by the population, however, only the distribution of last elements was significantly different (χ^2 = 11.4, df = 3, P = 0.0095; Fisher's P = 0.009); the difference was driven by high rates of upsweeps and downsweeps and the very low occurrence of trills in the final position.

Trills accounted for the majority (26.2%) of single-element syllables, a higher proportion than anticipated by their rate of expression overall (20.7%; Table 2i). In syllables involving multiple elements, downsweeps accounted for the majority of syllable-starts (50.9%) and syllable-ends (47.3%), a higher proportion than anticipated by their overall rate of expression (31.9%). The frequencies with which elements were recorded in an intermediate (non-start/end) position in a multi-element syllable were: trills (32.4%); upsweeps (30%) downsweeps (23.5%); tones (12.8%); unclassified (0.9%); and clicks (0.4%).

At an individual level, the positioning of elements within syllables differed from that expected by chance (Table 2i; distribution of first and last elements all P < 0.005), and also differed to some extent from the natural ratio of element production by the population (distribution of first elements P < 0.005 in seven out of nine individuals; distribution of

last and intermediate elements P < 0.005 in eight out of nine individuals).

Phrases

The positioning of song elements within phrases was a more obvious source of patterning than their positioning within syllables (Table 2). The distribution of first (χ^2 = 24.6, df = 3, P < 0.0001; Fisher's P = 0.0005), last (χ^2 = 27.5, df = 3, P < 0.0001; Fisher's P = 0.0005); isolated (χ^2 = 14.5, df = 3, P = 0.002; Fisher's P = 0.002); and intermediate elements (χ^2 = 18.5, df = 3, P = 0.0003; Fisher's P = 0.0005) were all non-random (Table 2ii). Moreover, the distribution of first (χ^2 = 15.5, df = 3, P = 0.001; Fisher's P = 0.001), last (χ^2 = 17.1, df = 3, P = 0.0006; Fisher's P = 0.0015); and isolated (χ^2 = 18.9, df = 3, P = 0.0003; Fisher's P = 0.0015) elements all differed from the natural rate of element expression by the population.

Trills accounted for the majority (36.4%) of single-element phrases – a higher proportion than anticipated by their rate of expression overall (20.7%). In phrases involving multiple elements, upsweeps accounted for the majority (50.3%) of phrase-starts, and downsweeps for the majority (60.2%) of phrase-ends – likewise higher proportions than anticipated by their overall rates of expression (31.8 and 31.9%, respectively).

The pattern of phrase construction was remarkably consistent between individuals (Table 2ii). Upsweeps were the element most commonly used to begin multi-element phrases (seven of nine individuals; 43.4–76.6% of their phrase-starts). One individual used downsweeps (Trajan; 65.3%) and one used trills (Marcus; 52.9%) as their preferred phrase-starts. Downsweeps were the most commonly used final element (seven of nine individuals; 36.2–87.8% of their phrase-ends), while one individual favoured upsweeps (Trajan; 40.8%), and one trills (Domitian; 48.9%). Trills were the element most commonly used in isolation (six of nine individuals; 27.1–73.3% of single-element phrases), while two individuals favoured clicks (Trajan, 26.7%; and Domitian, 33.8%). In one bat, the majority of single-element phrases were unclassified (Commodus, 42.1%); if unclassified elements were discounted, the next most common were trills (38.6%). Patterns of element use by individuals differed from those expected by chance (distribution of first elements all P < 0.005), and also differed somewhat from the natural ratio of element production by the population (distribution of first and intermediate elements P < 0.005 in seven out of nine individuals; distribution of last elements P < 0.005 in six out of nine individuals).

TABLE 2 The element most commonly used in the first, intermediate or last position along with elements most like to be found as stand-alone/isolated within (i) syllables and (ii) phrases.

i) Bat name	First	Intermediate	Last	Isolated
Antoninus	D (62.0%)	C (50.8%)	D (45.6%)	C (26.8%)
Commodus	D (89.5%)	C (37.2%)	C (80.2%)	A (28.9%)
Domitian	C (53.6%)	A (46.3%)	D (58.8%)	A (36.7%)
Hadrian	D (74.6%)	A (55.6%)	C (67.8%)	C (42.9%)
Lucius	A (39.8%)	A (43.0%)	D (75.7%)	A (36.3%)
Marcus	A (63.3%)	A (37.9%)	D (85.8%)	A (38.9%)
Nero	D (52.5%)	A (41.4%)	C (52.1%)	C (36.6%)
Otho	D (68.4%)	C (40.0%)	C (64.1%)	C (30.5%)
Trajan	D (77.8%)	D (40.0%)	C (53.3%)	A/E (22.3%)
Population	D (50.9%)	A (32.4%)	D (47.3%)	A (26.2%)

ii) Bat name	First	Intermediate	Last	Isolated
Antoninus	C (55.1%)	D (40.6%)	D (49.8%)	A (34.9%)
Commodus	C (43.4%)	C (40.5%)	D (36.2%)	F (42.1%)
Domitian	C (76.6%)	A (44.6%)	A (48.9%)	E (33.8%)
Hadrian	C (46.8%)	C (48.0%)	D (48.0%)	A (73.3%)
Lucius	C (49.3%)	D (38.7%)	D (87.8%)	A (69.1%)
Marcus	A (52.9%)	A (49.0%)	D (79.7%)	A (33.2%)
Nero	C (69.8%)	C (48.0%)	D (65.5%)	A (35.0%)
Otho	C (52.9%)	C (41.4%)	D (55.0%)	A (27.1%)
Trajan	D (65.3%)	C/D (34.3%)	C (40.8%)	E (26.7%)
Population	C (50.3%)	C (38.1%)	D (60.2%)	A (36.4%)

Results are listed for each individual bat and for the population as a whole. Elements are listed by annotation as: trill (A), tone (B), upsweep (C), downsweep (D), click (E), and unknown/unclassified (F). Each is listed with the proportion of times the element was found in that position, and different elements are colour coded for ease of reference.

Transitions between song constructs

Element-element

The most common element-element (first-order) transitions across all individuals were downsweep to upsweep (D-C; 37.2%);

trill to downsweep (A-D; 22%); upsweep to downsweep (C-D; 14.5%); and upsweep to trill (C-A; 7.6%; **Figure 4A**). Of the remaining possible element-element combinations, six had occurrence frequencies between 1–5%, 17 had occurrence frequencies between 0–1%, and nine possible transitions never occurred at all.

While the most common element-element transition overall was the downsweep-upsweep, its use among different individuals varied considerably, comprising between 7 and 73.2% of element-element transitions within an individual's repertoire (**Figure 4A**).

First-order (element-element) transitions were highly conserved, with all individuals having either downsweep-upsweep (D-C) or trill-downsweep (A-D) as their most-frequently used transition. Given that there were 25 possible first-order combinations (more if we allow for the possibility of unknown/unclassified elements), this is strong evidence of syntactic patterning underlying the construction of song. Mantel tests of structural similarity revealed a high degree of correlation between first-order transition matrices in the majority of individuals analysed (of 36 possible comparisons, 11 had R values $\geq$ 0.9, and 20 had R values $\geq$ 0.5), though one individual (Domitian) had distinctly different element transition patterns (See **Supplementary Material**).

Syllable-syllable

The most common syllable-syllable (second-order) transitions across all individuals were upsweep to upsweep (C-C; 11.4%); downsweep upsweep to downsweep upsweep (DC-DC; 10.6%); upsweep to downsweep upsweep (C-DC; 7.1%); trill to trill (A-A; 5.5%); upsweep to trill (C-A; 5.2%); upsweep to downsweep (C-D; 4.8%); and click to click (E-E; 4.8%; **Figure 4B**). Of the remaining possible syllable-syllable combinations, 13 had occurrence frequencies between 1–4%, 257 had occurrence frequencies between 0–1%, and 3,642 possible transitions never occurred at all. Syllable-syllable transitions were highly varied between individuals. While the most commonly used second-order transition overall was upsweep-upsweep (C-C), its use among different individuals ranged from 0.3% (Lucius) to 27.6% (Hadrian).

Phrase-phrase

At the level of transition from phrase to phrase (third order), the enormous variety of possible combinations ensured that the most common transitions were likely to be simple (single-element transitions) and contribute only a small proportion to the overall repertoire (**Figure 4C**). These were often recognisable in spectrograms as rapidly repeated single elements (such as click trains) or as repeated sequences of a common phrase. The most common phrase-phrase transitions across all individuals were trill to trill (A-A; 3.1%); trill downsweep to trill downsweep (AD-AD; 3%); click to click (E-E; 2%); and trill to trill downsweep (A-AD; 1.1%), along with

three transitions containing unknown/undefined elements (F-F, 3.05%; F-A, 1.3%; A-F, 1.2%). Of the remaining possible phrase-phrase transitions, 13 had occurrence frequencies between 0.5–1%, 81 had occurrence frequencies between 0.1–0.49%, and 2,430 possible transitions occurred at least once, but individually represented ≤0.09% of total transitions. Considering only known phrases, 222,900 possible transitions never occurred at all.

While the more complex polysyllabic phrases typically did not have a high repeatability rate, some were repeated often enough across the population to be notable. The most highly ranked polysyllabic phrase-phrase transitions predominantly involved upsweeps and trills: C A – C A (0.5%); C D – C D (0.2%); C AD – C AD (0.2%); A – C A (0.2%); CA – C A (0.2%); and C AD – C A (0.2%).

Transition sequences increased in complexity across increasing hierarchic level (i.e., from first- to third-order; **Figure 4**). Of the element-element transitions, 37.2% were accounted for by a single transition; this increased to 59.2% with the addition of the second-most common transition. By contrast, at the level of the phrase, the most-common transition accounted for only 3.1%, while over 200 possible transitions were required to reach 50% of the total. These results are indicative of expanding complexity and increasing individual variation with increasing combinatorial possibility.
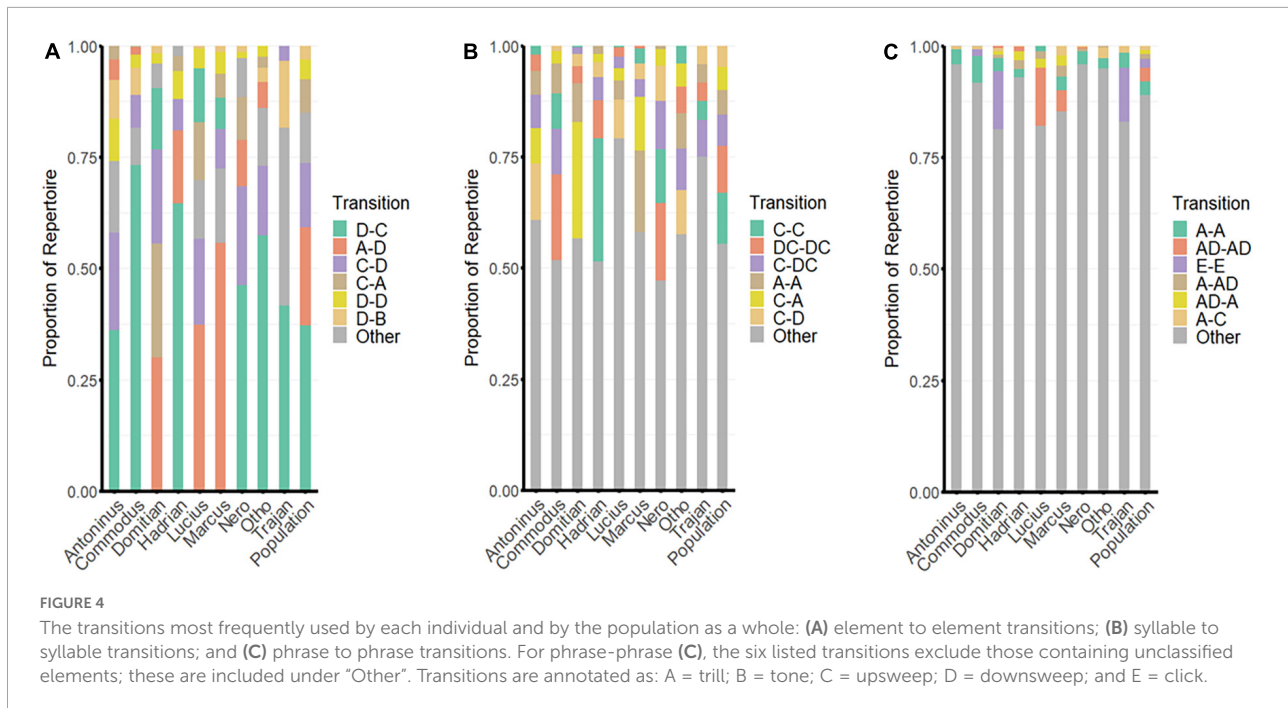
Visual inspection of spectrograms shows that strings of similar syllables or phrases often occur together, as do sequences that "morph" gradually from one into another (e.g., C > C > C A > CA > CAD > AD > AD > A D). This phenomenon is not easily defined or quantified, as there are numerous possible variations. For example, CA > CAD > AD is indicative of "morphing," but so is CA > C AD > AD, or DC > DCD > CD, or an enormous range of other possibilities. As such, it is beyond our power to quantify here, but we do provide examples of spectrograms in which this kind of morphing is demonstrated (**Figure 5**).

Discussion

Male *M. tuberculata* have at their disposal a highly versatile song structure that allows for considerable individual variation. As we have demonstrated, *M. tuberculata* songs are hierarchically constructed and employ phonological syntax to build their repertoires. Syllables and phrases are arranged with high structural similarity and conservatism in terms of their start/end elements and element-element transitions, but still retain a capacity for versatility and innovation *via* differences in the number and arrangement of intermediate components and in the transitions between syllables and phrases.

Our first hypothesis investigated the frequency of element production in *M. tuberculata* song. The frequency of element production by individuals is non-random, and elements

FIGURE 4

The transitions most frequently used by each individual and by the population as a whole: **(A)** element to element transitions; **(B)** syllable to syllable transitions; and **(C)** phrase to phrase transitions. For phrase-phrase **(C)**, the six listed transitions exclude those containing unclassified elements; these are included under "Other". Transitions are annotated as: A = trill; B = tone; C = upsweep; D = downsweep; and E = click.

are not used with equivalent frequency in the course of a song. Upsweeps, downsweeps and trills make up the majority of the song, with other elements perhaps providing an opportunity for individuals to distinguish themselves by deviating from the common theme (e.g., Behr and von Helversen, 2004; Russ and Racey, 2007; Jahelková et al., 2008; Chaverri et al., 2018; Toth and Parsons, 2018). The use of these less common song elements also plays a role in increasing individuals' overall repertoire diversity. In species with open-ended vocal learning, repertoire diversity commonly functions as an indicator of male age and experience (Ballentine, 2009; Knörnschild et al., 2010; Chabout et al., 2015), making it a common predictor of female preference (Davidson and Wilkinson, 2004), territory quality (Mcgregor et al., 1981; Catchpole, 1986), or lifetime reproductive success (Mcgregor et al., 1981).

Increasing the number of discrete elements used to construct syllables and phrases is another avenue males may exploit to add diversity to their repertoires (Davidson and Wilkinson, 2004; Behr et al., 2006; Catchpole and Slater, 2008; Chabout et al., 2015). Our second hypothesis examined this possibility, and revealed considerable variation in syllable and phrase construction between different males, with some individuals favouring simple constructs such as single-element phrases and others using higher rates of complex multi-element phrases. There does, however, appear to be an upper limit to the possible complexity of phrases. The most complex we recorded contained 16 distinct elements (or 12 distinct syllables), but most had between one and three.

Our third hypothesis related to the positioning of elements within larger song constructs. While our interpretation of the syntactic rules governing song must be prefaced by the caveat that this was a small study of relatively few individuals, nonetheless, some general conclusions may be drawn: downsweeps are the element most commonly used to end both syllables and phrases; if used to begin a syllable, a downsweep is usually followed by an upsweep; upsweeps are the element most commonly used to start a phrase; and trills are the element most commonly found in isolation. The use of particular start/end elements is found in a wide range of taxa, and is often more rigidly controlled than the distribution of intermediate elements (Galeotti et al., 1997; Chabout et al., 2015; Knörnschild et al., 2017).

Our fourth hypothesis investigated the structuring of transitions between song constructs. While transitions between the component elements of a composite syllable are non-random, it is less clear precisely how much structural dependence exists in the transitions between the component syllables of a composite phrase, or in the transition from one phrase to another. The unit of the phrase seems, functionally, to be the independent unit of *M. tuberculata* song (in that it occurs surrounded by silence; compare Kanwal et al., 1994; Behr and von Helversen, 2004; Davidson and Wilkinson, 2004; Bohn et al., 2009), and its production should theoretically not be physiologically constrained by the phrase that preceded it. Thus, there would seem to be some reliance on "patterning" for the building of syllables and phrases but, once constructed, phrases appear able to be used more or less at will. This is supported
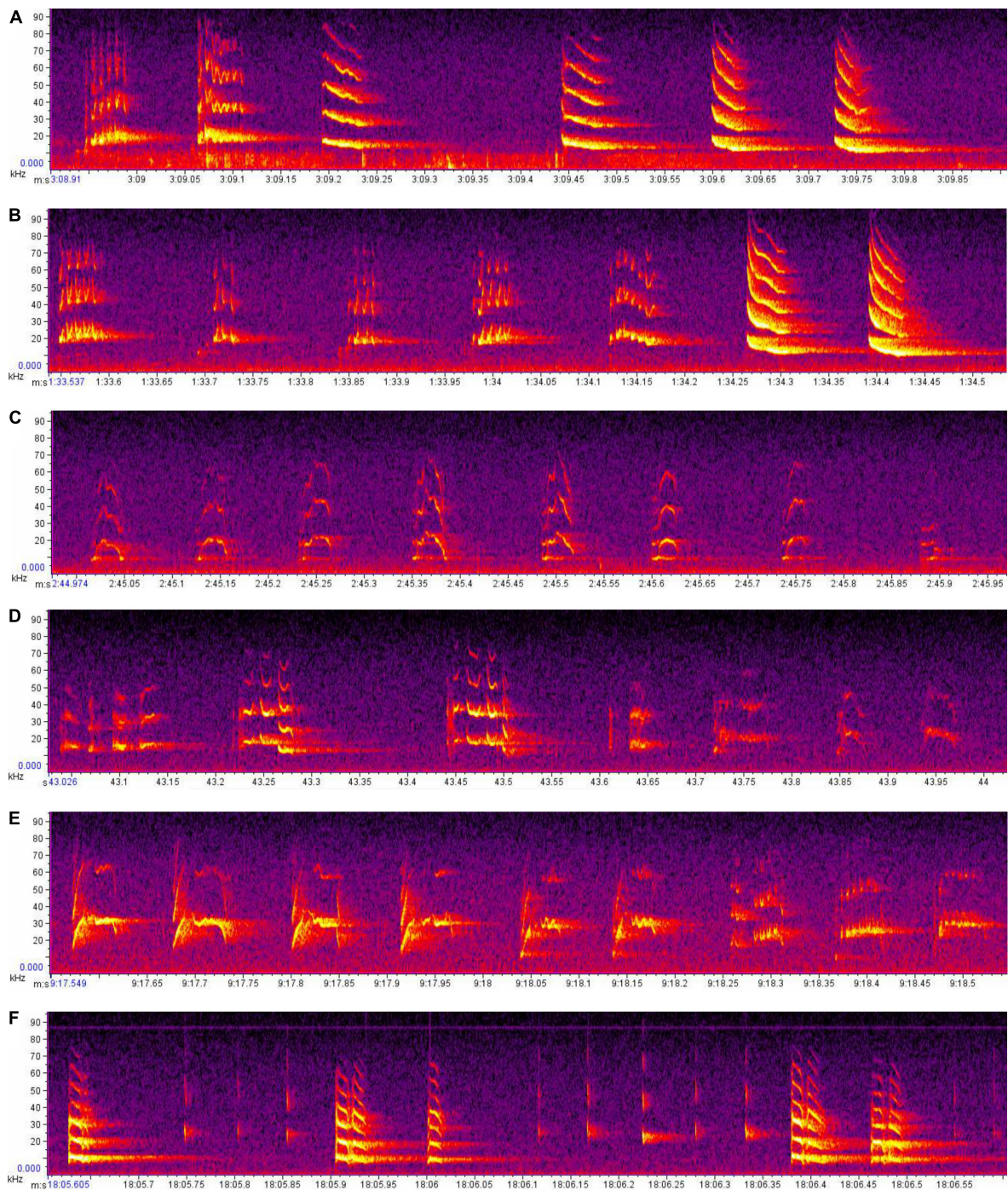
FIGURE 5

Spectrogram images of demonstrating variation in call structures within and between the songs of different *Mystacina tuberculata* males. Three excerpts **(A–C)** are taken from the song of the same male (Nero); and three from different males (**D**, Domitian; **E**, Marcus; and **F**, Otho). Each excerpt is one second in duration. Note the repeated use of sequential phrases (e.g., **C**) and the instances in which one phrase-type morphs into another [e.g., from trill to downsweep **(A)** or from upsweep-trill-downsweep to upsweep-trill to trill **(E)**, but also instances in which phrases are followed by others that are totally dissimilar (e.g., **B,D**)].

by the high degree of variation and randomness exhibited in the transitions between phrases. Nonetheless, strings of similar syllables or phrases occurring in sequence are relatively

common, as are sequences that "morph" gradually from one into another. We have not attempted to quantify this phenomenon at this stage, but it is noteworthy that similar patterns are found

in sac-winged bats (*S. bilineata*; Behr and von Helversen, 2004), and also in mice (Holy and Guo, 2005).

Broadly speaking, the findings of this study are similar to those regarding song structure in free-tailed bats (*Tadarida brasiliensis*; Bohn et al., 2009) and greater sac-winged bats (*S. bilineata*; Behr and von Helversen, 2004). While our methods differed in some respects (including nomenclature and level of analysis, which makes direct comparison difficult), all three species appear to have relatively conserved methods of syllable construction, with more diversity apparent at the higher levels. Certain preferred broad-scale patterns, however, remain apparent in each species. For instance, we found that phrases are more likely to end with downsweeps than any other element, just as *T. brasiliensis* songs are most likely to end with buzzes (Bohn et al., 2009). In all three species, males have repertoires that vary considerably in the frequency and ordering of phrases, which may provide an opportunity to encode individual identity (Behr and von Helversen, 2004; Russ and Racey, 2007; Jahelková et al., 2008; Chaverri et al., 2018; Toth and Parsons, 2018).

Given the scope that combinatorial syntax allows for individual expression, and the song's assumed role in attracting females for mating (Lloyd, 2001; Carter and Riskin, 2006; Toth et al., 2015; Toth and Parsons, 2018), it is highly likely that aspects of song construction are used by females to assess male traits. This has been shown in *S. bilineata*, where males that use a greater number of unique composite syllables retain more females in their territories (Davidson and Wilkinson, 2004), and males show consistent individual differences in song construction that may aid individual recognition and facilitate female mate choice (Behr and von Helversen, 2004). Numerous similar examples exist in birdsong studies in which aspects of song construction have been correlated with female preference or male reproductive success (e.g., Mcgregor et al., 1981; Catchpole, 1986; Lampe and Espmark, 2003; Ballentine, 2009). Some correlates between song construction and male traits in *M. tuberculata* have already been discovered, including a negative correlation between the duration of trill-downsweep syllables and forearm length (Toth and Parsons, 2018). Given the complexity of the song structures in question, we anticipate that there is further information to be gleaned from courtship songs regarding males' individual attributes, energetic investment (e.g., Behr et al., 2006; Collier et al., 2022), and possibly even social context (Bohn et al., 2013).

While we have restricted this analysis to basic categorisation of song elements, it is important to note that the acoustic properties of these elements (such as the peak amplitude of a downsweep or the number of fluctuations in a trill) may also hold considerable significance for a receiver. As an example, the elements that we have termed "upsweeps" are typically either convex (r-shaped) or concave (j-shaped), but may also be relatively straight (/-shaped) or have a wavering, trill-like quality. There is thus considerable variance in energy distribution and duration even between elements that we have given the same classification. At a bare minimum, the four most commonly used syllables in *M. tuberculata* song (trills, upsweep-trills, trill-downsweeps and upsweep-trill-downsweeps) are produced with sufficient between-individual variation to allow for the identification of individual singers (Toth and Parsons, 2018). This implies that acoustic variation in song element production may be of substantial biological significance.

It is currently unclear how *M. tuberculata* song is learned (if indeed it *is* learned). If vocal learning occurs during ontogeny and is not open-ended, then the question arises as to how a male pup learns courtship song when raised predominantly by a female parent. In *S. bilineata,* song is learned *via* vocal imitation during ontogeny by pups of both sexes (Knörnschild et al., 2006; 2010). The breeding systems of the two species differ, however, in that *S. bilineata* live year-round in harems, meaning that pups have close contact with the harem male and are regularly exposed to his courtship songs (Knörnschild et al., 2006; 2010). While *M. tuberculata* males are known to day-roost at maternity colonies and to display in their vicinity, it is unclear whether this on its own is sufficient for pup vocal learning to take place. The alternative, that *M. tuberculata* are open-ended learners, suggests the possibility of eavesdropping and repertoire sharing between close territory holders. Given that *M. tuberculata* males display in close proximity to one another (and indeed may in some instances "timeshare" roosts; Toth et al., 2018), the potential for eavesdropping and repertoire-sharing would seem high. Toth and Parsons (2018) investigated the possibility briefly, but failed to find any evidence for it. While the spatial aspect of such a question was beyond the scope of our analysis, we did note certain visual similarities in phrase types common to two individuals known to share a timeshare roost (Otho and Lucius); these similarities, however, are not quantifiable at this stage.

The songs of male short-tailed bats are sophisticated, intricate, and display remarkable syntactic complexity. We posit that *M. tuberculata* song utilises a form of phonological syntax, in which smaller vocal units are combined into larger ones in accordance with certain rules. Whilst this study may have been able to shed some light on the nature of those rules, much remains to be discovered regarding the limits they may impose on individual performance and the amount of information that males are able to convey within those bounds.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Ethics statement

Author contributions

Both authors were involved in conception and design of the experiment. KC conducted the field work and analysed the data. Both authors were involved in conception and design of the experiment and contributed to the manuscript.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fevo.2022.907791/full#supplementary-material

References

Balcombe, J. P. (1990). Vocal Recognition of Pups by Mother Mexican Free-Tailed Bats, Tadarida brasiliensis mexicana. *Anim. Behav.* 39, 960–966.

Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size in male swamp sparrows, Melospiza georgiana. *Anim. Behav.* 77, 973–978. doi: 10.1016/j.anbehav.2008.12.027

Barlow, K. E., and Jones, G. (1997). Function of pipistrelle social calls: field data and a playback experiment. *Anim. Behav.* 53, 991–999. doi: 10.1006/anbe.1996.0398

Barón Birchenall, L. (2016). Animal communication and human language: an overview. *Int. J. Comput. Psychol.* 29:28000.

Behr, O., and von Helversen, O. (2004). Bat serenades - Complex courtship songs of the sac-winged bat (Saccopteryx bilineata). *Behav. Ecol. Sociobiol.* 56, 106–115. doi: 10.1007/s00265-004-0768-7

Behr, O., Knörnschild, M., and Von Helversen, O. (2009). Territorial Counter-Singing in Male Sac-Winged Bats (Saccopteryx bilineata): low-Frequency Songs Trigger a Stronger Response. *Behav. Ecol. Sociobiol.* 63, 433–442. doi: 10.1007/s00265-008-0677-2

Behr, O., Von Helversen, O., Heckel, G., Nagy, M., Voigt, C. C., and Mayer, F. (2006). Territorial songs indicate male quality in the sac-winged bat Saccopteryx bilineata (*Chiroptera*, *Emballonuridae*). *Behav. Ecol.* 17, 810–817. doi: 10.1093/beheco/arl013

Berwick, R. C., Beckers, G. J. L., Okanoya, K., and Bolhuis, J. J. (2012). A bird's eye view of human language evolution. *Front. Evol. Neurosci.* 4:5. doi: 10.3389/fnevo.2012.00005

Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M., and Pollak, G. D. (2009). Versatility and stereotypy of free-tailed bat songs. *PLoS One* 4:e6746. doi: 10.1371/journal.pone.0006746

Bohn, K. M., Schmidt-French, B., Ma, S. T., and Pollak, G. D. (2008). Syllable acoustics, temporal patterns, and call composition vary with behavioral context in Mexican free-tailed bats. *J. Acoust. Soc. Am.* 124, 1838–1848.

Bohn, K. M., Smarsh, G. C., and Smotherman, M. (2013). Social context evokes rapid changes in bat song syntax. *Anim. Behav.* 85, 1485–1491. doi: 10.1016/j.anbehav.2013.04.002

Bohn, K. M., Wilkinson, G. S., and Moss, C. F. (2007). Discrimination of infant isolation calls by female greater spear-nosed bats, Phyllostomus hastatus. *Anim. Behav.* 73, 423–432. doi: 10.1016/j.anbehav.2006.09.003

Carter, G. G., and Riskin, D. K. (2006). Mystacina tuberculata. *Mamm. Species* 790, 1–8. doi: 10.1644/790.1

Catchpole, C. (1986). Song Repertoires and Reproductive Success in the Great Reed Warbler Acrocephalus arundinaceus. *Behav. Ecol. Sociobiol.* 19, 439–445. doi: 10.1111/j.1558-5646.1998.tb03712.x

Catchpole, C., and Slater, P. J. B. (2008). *Bird Song: Biological Themes and Variations (Second)*. Cambridge: Cambridge University Press.

Chabout, J., Sarkar, A., Dunson, D. B., and Jarvis, E. D. (2015). Male mice song syntax depends on social contexts and influences female preferences. *Front. Behav. Neurosci.* 9:76. doi: 10.3389/fnbeh.2015.00076

Chaverri, G., Ancillotto, L., and Russo, D. (2018). Social communication in bats. *Biol. Rev.* 93, 1938–1954. doi: 10.1111/brv.12427

Collier, K., Parsons, S., and Czenze, Z. J. (2022). Thermal energetics of male courtship song in a lek-breeding bat. *Behav. Ecol. Sociobiol.* 76, 1–8. doi: 10.1007/s00265-022-03141-5

Davidson, S. M., and Wilkinson, G. S. (2004). Function of male song in the greater white-lined bat, Saccopteryx bilineata. *Anim. Behav.* 67, 883–891. doi: 10.1016/j.anbehav.2003.06.016

Engesser, S., and Townsend, S. W. (2019). Combinatoriality in the vocal systems of non-human animals. *Wiley Interdiscip. Rev. Cogn. Sci.* 10:e1493.

Galeotti, P., Saino, N., Sacchi, R., and MØller, A. P. (1997). Song correlates with social context, testosterone and body condition in male barn swallows. *Anim. Behav.* 53, 687–700. doi: 10.1006/anbe.1996.0304

Hailman, J. P., and Ficken, M. S. (1986). Combinatorial animal communication with computable syntax: chick-a-dee calling qualifies as "Language" by structural linguistics. *Anim. Behav.* 34, 1899–1901. doi: 10.1016/S0003-3472(86)80279-2

Holy, T. E., and Guo, Z. (2005). Ultrasonic Songs of Male Mice. *PLoS Biol.* 3:e386. doi: 10.1371/journal.pbio.0030386

Jahelková, H., Horáèek, I., and Bartonièka, T. (2008). The advertisement song of *Pipistrellus nathusii* (*Chiroptera*, *Vespertilionidae*): a complex message containing acoustic signatures of individuals. *Acta Chiropterol.* 10, 103–126. doi: 10.3161/150811008x331144

Kanwal, J. S., Matsumura, S., Ohlemiller, K., and Suga, N. (1994). Analysis of acoustic elements and syntax in communication sounds emitted by mustached bats. *J. Acoust. Soc. Am.* 96, 1229–1254. doi: 10.1121/1.410273

Knörnschild, M., Behr, O., and Von Helversen, O. (2006). Babbling behavior in the sac-winged bat (Saccopteryx bilineata). *Naturwissenschaften* 93, 451–454. doi: 10.1007/s00114-006-0127-9

Knörnschild, M., Blüml, S., Steidl, P., Eckenweber, M., and Nagy, M. (2017). Bat songs as acoustic beacons - Male territorial songs attract dispersing females. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-14434-5

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., and Von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biol. Lett.* 6, 156–159. doi: 10.1098/rsbl.2009.0685

Kroodsma, D. E. (1977). Correlates of song organization among North American wrens. *Am. Nat.* 111, 995–1008. doi: 10.1086/283228

Lampe, H. M., and Espmark, Y. O. (2003). Mate choice in Pied Flycatchers Ficedula hypoleuca: can females use song to find high-quality males and territories? *Ibis* 145, E24–E33. doi: 10.1046/j.1474-919X.2003.00144.x

Lloyd, B. D. (2001). Advances in New Zealand mammalogy 1990–2000: short-tailed bats. *J. R. Soc. N. Z.* 31, 111–125. doi: 10.1080/03014223.2001.9517642

Marler, P. (1977). "The structure of animal communication sounds," in *Recognition of Complex Acoustic Signals*, ed. T. H. Bullock (New York: Springer), 17–35.

Marler, P. (1998). "Animal communication and human language," in *The Origin and Diversification of Language,* eds N. D. Jablonski and L. E. Aiello (San Francisco, CA: California Academy of Sciences), 1–19.

Mcgregor, P. K., Krebs, J. R., and Perrins, C. M. (1981). Song Repertoires and Lifetime Reproductive Success in the Great Tit (Parus major). *Am. Nat.* 118, 149–159.

Melendez, K. V., Jones, D. L., and Feng, A. S. (2006). Classification of communication signals of the little brown bat. *J. Acoust. Soc. Am.* 120, 1095–1102. doi: 10.1121/1.2211488

Mitani, J. C., and Marler, P. (1989). A Phonological Analysis of Male Gibbon Singing Behavior. *Behaviour* 109, 20–45.

Prat, Y., Taub, M., and Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Sci. Rep.* 6, 1–10. doi: 10.1038/srep39419

R Core Team (2019). *R: a Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Russ, J. M., and Racey, P. A. (2007). Species-specificity and individual variation in the song of male Nathusius' pipistrelles (*Pipistrellus nathusii*). *Behav. Ecol. Sociobiol.* 61, 669–677. doi: 10.1007/s00265-006-0295-9

Scharff, C., and Petri, J. (2011). Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philos. Trans. R. Soc. B Biol. Sci.* 366, 2124–2140. doi: 10.1098/rstb.2011.0001

Smotherman, M., Knörnschild, M., Smarsh, G., and Bohn, K. (2016). The origins and diversity of bat songs. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 202, 535–554. doi: 10.1007/s00359-016-1105-0

Toth, C. A. (2016). *Lek Breeding in the Lesser Short-Tailed bat (Mystacina tuberculata): Male Courtship, Female Selection, and the Determinants of Reproductive Strategies*. Ph.D. thesis. Auckland: The University of Auckland.

Toth, C. A., and Parsons, S. (2018). The high-output singing displays of a lekking bat encode information on body size and individual identity. *Behav. Ecol. Sociobiol.* 72:102. doi: 10.1007/s00265-018-2496-4

Toth, C. A., Dennis, T. E., Pattemore, D. E., and Parsons, S. (2015). Females as mobile resources: communal roosts promote the adoption of lek breeding in a temperate bat. *Behav. Ecol.* 26, 1156–1163. doi: 10.1093/beheco/arv070

Toth, C. A., Santure, A. W., Holwell, G. I., Pattemore, D. E., and Parsons, S. (2018). Courtship behaviour and display-site sharing appears conditional on body size in a lekking bat. *Anim. Behav.* 136, 13–19. doi: 10.1016/j.anbehav.2017.12.007

Vernes, S. C. (2017). What bats have to say about speech and language. *Psychon. Bull. Rev.* 24, 111–117. doi: 10.3758/s13423-016-1060-3

Vernes, S. C., and Wilkinson, G. S. (2020). Behaviour, biology and evolution of vocal learning in bats. *Philos. Trans. R. Soc. B Biol. Sci.* 375:20190061. doi: 10.1098/rstb.2019.0061

Wiley, R. H. (2018). "Design Features of Language," in *Encyclopedia of Evolutionary Psychological Science*, eds T. K. Shackelford and V. A. Weekes-Shackelford (New York, NY: Springer International Publishing), e1–e13. doi: 10.1007/978-3-319-16999-6_3837-1