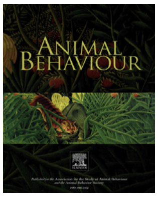
Animal Behaviour 112 (2016) 285–300

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav



# Waltzing *Taeniopygia*: integration of courtship song and dance in the domesticated Australian zebra finch

Robert Ullrich, Philipp Norton, Constance Scharff [*]

*Department of Animal Behaviour, Freie Universität Berlin, Berlin, Germany*

ARTICLE INFO

*Article history:*
Received 11 May 2015
Initial acceptance 1 July 2015
Final acceptance 13 October 2015
Available online 29 January 2016
MS. number: 15-00394R

*Keywords:*
choreography
communication
courtship
dance
multimodal signals
song
zebra finch

Multimodal signalling can improve or maximize information exchange. A challenge is to show that two independent signals, such as vocalizations and visual displays, are deliberately coordinated. Male zebra finches, *Taeniopygia guttata*, signal visually and acoustically during courtship, performing a stereotyped dance while singing. The male approaches the female hopping in a zig-zag pattern, turning his body axis, and wiping his beak repeatedly on or above the perch. The only previous quantitative study of song and dance choreography in zebra finches revealed that the distribution of all movements during song was not strongly patterned across birds but very similar in fathers and sons. This raises the possibility that particular movements may follow a choreography. Here we report that three operationally defined dance movements, 'beak wipe' (BA), 'turn-around' (TA) and 'hop', occurred with different frequencies and speed during singing than during silence. BW, TA and hops clustered significantly at the start and end of song bouts and were arranged in a nonrandom fashion. In addition, BW, but not TA, were performed faster during song than nonsong. Finally, hops coincided significantly more often than expected by chance with particular notes. Together, these results suggest that male zebra finches integrate their song and dance during courtship. This may help females to identify courting males in a noisy environment and evaluate the intensity and quality of the courtship performance. Our results underscore that the choreography of movement gestures with learned vocalizations, such as hand gestures accompanying speech, is a further parallel between human and avian signalling. They invite future investigations into the underlying neural mechanisms and consequences for mate choice.

Communication is essential for organizing the social interactions of zebra finches, *Taeniopygia guttata*, that live in large flocks (Immelmann, 1969). Both sexes use several types of calls but only males sing songs (Slater, Eales, & Clayton, 1988; Zann, 1996). Males direct their song towards females during courtship but also sing so-called 'undirected song' in nonreproductive contexts, e.g. when alone or in the company of males (Dunn & Zann, 1996; Immelmann, 1959; Jarvis, Scharff, Grossman, Ramos, & Nottebohm, 1998; Morris, 1954; Sossinka & Böhner, 1980). Males can rapidly switch between directed and undirected songs (Caryl, 1981), which differ bioacoustically and in the underlying neural processes (Cooper & Goller, 2006; Jarvis et al., 1998; Kao & Brainard, 2006; Sossinka & Böhner, 1980; Woolley & Doupe, 2008). Courtship song is accompanied by a characteristic dance during which the male advances towards the female turning 180° with each hop, twisting his head and tail towards the female while continuing to sing. When facing the female, he also rhythmically swings his head from side to side while singing (Barclay, Harding, & Waterman, 1992; Immelmann, 1959; Kunkel, 1959; Morris, 1954; Zann, 1996, pp. 170–174; Fig. 1).

When animals use multiple signalling modes to attract a mate, those signals need not necessarily be coordinated with each other in their fine temporal pattern (e.g. butterfly coloration and pheromones for mate attraction). However, in some animals it has been shown that courtship can involve two signalling modes that are integrated into a coherent performance, providing the basis for an integrated percept in the receiver. For instance, jumping spiders, *Habronattus pugillis*, coordinate visual and seismic signals using limbs and body appendages (Elias, 2006). Golden-collared manakins, *Manacus vitellinus*, and club-winged manakins, *Machaeropterus deliciosus*, integrate acoustic wing-snaps with visual displays during courtship (Barske, Schlinger, Wikelski, & Fusani, 2011; Bostwick, 2000). Barbary doves, *Streptopelia risoria*, perform visual bows accompanied by acoustic bowing-calls during

---

[*] Correspondence: C. Scharff, Freie Universität Berlin, Takustraße 6, 14195 Berlin, Germany.
*E-mail address:* constance.scharff@fu-berlin.de (C. Scharff).

http://dx.doi.org/10.1016/j.anbehav.2015.11.012

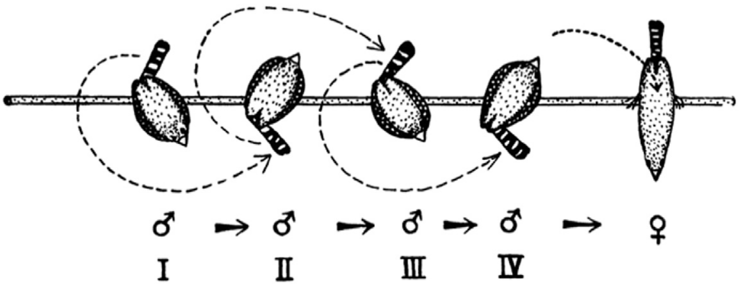


**Figure 1.** Schematic representation of a male zebra finch courtship dance facing a female bird. Reprinted with permission from Morris (1954, p. 286).

courtship. The onset of both signals is not temporally synchronized, but the intervals between bowing and calling are rhythmically correlated, leading to an integrated complex signal (Fusani, Hutchison, & Hutchison, 1997). Male superb lyre-birds, *Menura novaehollandiae*, integrate acoustic and movement signals voluntarily and within a predictable pattern during courtship (Dalziell et al., 2013). Nonvocal bill-clicking patterns and specific song note sequences in Java sparrows, *Lonchura oryzivora*, are closely integrated. They occur for instance more frequently with the beginning of song, which might be an outcome of cultural transmission (Soma & Mori, 2015). In various species, female recipients are known to be sensitive to the multimodality of these signals. For instance, females of the estrildid family have a complex dance in species in which males also have a complex dance (Soma & Garamszegi, 2015). Also, wing-spread displays in cowbirds, *Molothrus ater*, accompanying song (Cooper & Goller, 2004) elicit longer lasting copulation solicitation displays from females than presentation of song or wing-spread displays alone (O'Loghlen & Rothstein, 2010). Audiovisual playback experiments revealed that the female's sexual behaviour varies with the intensity of the male's body movements (O'Loghlen & Rothstein, 2012). Interestingly, wing movements and song were integrated even in males that had been reared without visual or auditory input from male tutors (Hoepfner & Goller, 2013). Female golden-collared manakins are influenced by speed and frequency of the displays by courting males (Barske et al., 2011). Female túngara frogs, *Physalaemus pustulosus*, prefer a robotic frog that correctly coordinates the visual signal of the vocal sac inflation with the produced calls (Taylor, Klein, Stein, & Ryan, 2011). In humans, the integration of gesture and speech is intimately linked to learning and creating language (Goldin-Meadow & Alibali, 2013). These examples stress the additive effect of two signalling channels, and the relevance of temporal integration of the two signals with each other. In contrast to these examples, in other birds visual and vocal displays are not synchronized. For instance, long-tailed manakins, *Chiroxiphia linearis*, and starlings, *Sturnus vulgaris*, perform visual and acoustic signals in a parallel, but not in an integrated fashion (Böhner & Veit, 1993; Lukianchuk & Doucet, 2014).

In male zebra finches it is not clear whether the dance movements accompanying song are synchronized to the song. The only study that examined whether song and dance are coupled in specific ways found no significant association of stereotypic body movements with acoustic features of song (Williams, 2001).

Here we operationally defined two dance elements, beak wipes and turn-arounds, characteristic of high-intensity courtship (Barclay et al., 1992), as well as a third dance element, hopping behaviour, and analysed their relationship to vocalizations in videotapes. We hypothesized that these movements were associated with song more so than when birds were silent. We further speculated that movements were associated with specific positions in the song motifs.

## METHODS

### *Subjects*

A total of 20 captive-bred male zebra finches (9–45 months old) participated in three variations of the experimental set-up, seven in 2010 (experiment A), four in 2011 (B) and 10 in 2013 (C). One male was tested twice (2010, 2011) and his data were averaged. Before experiment A male birds were housed in groups of seven males for several months. Birds were transferred from large aviaries into smaller group cages 1 week before experiments B and C. Experiment B was performed in May outdoors, and before, during and after testing, birds were also kept outdoors. Indoor conditions were kept at 25 ± 3 °C and 12:12 h light:dark cycle. All subjects had access to seed, water, grit and cuttlebone ad libitum.

### *Recording*

An adult male was introduced to the experimental set-up containing one female that he could hear but not see. After at least 2 h the visual separation was removed for 5 min and the male could see the female. Males that performed courtship dances and directed their song to females during this time were audiorecorded and videotaped. During 2 days at least three video sessions per bird, separated by 20 min or more, took place. All recorded songs were directed to the female, accompanied by the typical courtship-associated movements and body posture. We varied the experimental set-up during the three experiments with respect to location, cage size and video/audio equipment which ensured that our findings were not affected by the physical constraints of a particular experimental set-up (Fig. 2, Table A1). Because zebra finches in the wild court on branches rather than on the floor, we equipped our cages with a perch in contrast to a former study (Williams, 2001). Video and audio streams were digitized and stored on hard disks or on SD cards.

### *Behaviour Definitions, Audio and Video Analysis*

Audio files (22.05 kHz, 16 bit resolution) were converted into sound spectrograms using Avisoft-SASLab Pro 4.38 software (Avisoft Bioacoustics, Berlin, Germany; settings: FFT_256 points, Hamming window, overlap 50%). Video recordings were analysed frame by frame using Noldus Observer 9 XT (Noldus Information Technology, Wageningen, The Netherlands) and dance-associated

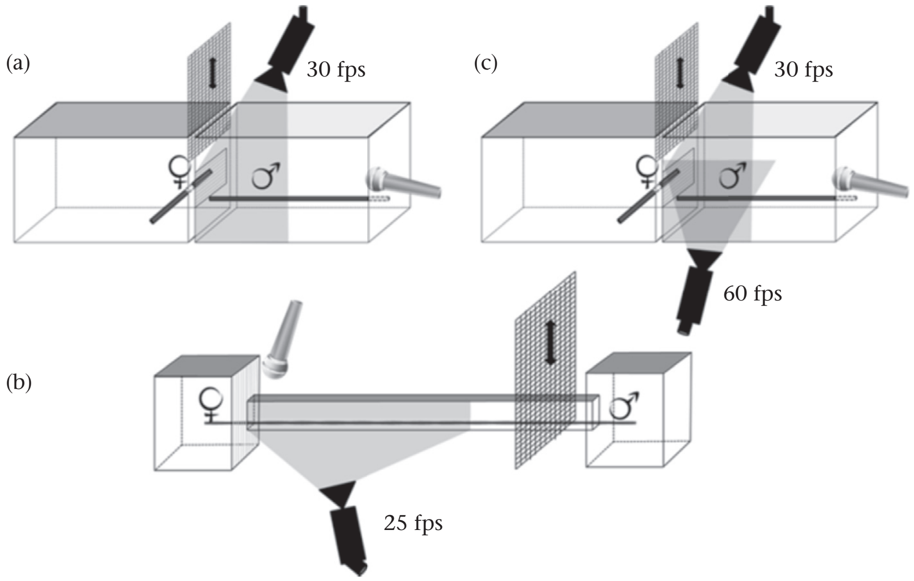


**Figure 2.** Experimental set-up for (a) experiment A, (b) experiment B and (c) experiment C. For experiments A and C, males and females occupied two separate plastic-walled cages (51 × 30 cm and 40 cm high), separated by a removable opaque divider. The top of the cage and one of the long sides consisted of transparent plastic material. Videos were recorded from above and additionally from the front (experiment C); audio was recorded from the side through an opening of the cage. For experiment B, cages (40 × 30 cm and 40 cm high) were connected by a 1.5 m long tunnel (150 × 20 cm and 30 cm high) containing a single long perch. The front of the tunnel consisted of transparent plastic material, the other sides of metal wire mesh. Video was taped from the front, audio from above. The male entered the tunnel at the start of the 5 min test period after an opaque divider was removed. The male and female heard each other in all set-ups, but could only see each other through a 17 × 17 cm wire mesh (experiment A, B) or transparent plastic window (experiment C) during the 5 min test period. Birds could approach each other to a vicinity of 5 cm but not interact physically. During indoor experiments (A, C) the cage was illuminated by four fluorescent lamps (18 W), experiments outdoors (B) were performed under ambient light.

movements (see Fig. A1) were scored using the Observer software.

Adult zebra finch song begins with several repetitions of a single introductory note, followed by a set of dissimilar notes. The song notes are uttered in a stereotyped sequential order that constitute a ‘motif’ (Fig. 3). We defined a song note as a continuous, morphologically discrete trace on a sound spectrogram. Motifs that are sung in close succession result in bouts with pauses usually lasting less than 0.5 s. Pauses longer than 2 s were considered the end of both a bout and a motif (Sossinka & Böhner, 1980). The number of motifs analysed per bird ranged from 93 to 324 (mean ± SD = 178.1 ± 75.62). Between songs, birds utter a number of different calls that are temporally less stereotypically delivered than song.

Based on the audio track we divided experiments into ‘song segments’ and ‘nonsong segments’ (Fig. 3c). We defined as ‘song segment’ the period comprising the first 2 s before the first motif note, the motif and the 2 s following the last motif note. The rest of the experimental 5 min were scored as ‘nonsong segment’ (Fig. 3a). Within the ‘song segment’ we distinguished three subdivisions (Fig. 3b and c): (1) start (of a song bout: 2 s before the onset of the first introductory note to the end of the first motif note, or of a within-bout motif: from the onset of occasional introductory notes to the end of the first motif note); (2) centre (including all other notes from the end of the first motif note to the onset of the last); (3) end (the final segment of the motif from the onset of the last (not necessarily canonical) note and up to 2 s following it, less than 2 s if followed closely by another motif). Half of the intervals between two motifs within a bout were considered part of the ‘start’ segment and the other half to be part of the ‘end’ segment. Because ‘start’, ‘centre’ and ‘end’ differed in duration we report the number of movements per second (Fig. 3b and c).

Among the various courtship dance-associated movements, such as the inflation of the gular sac, stereotypic head movements, upright and lower body posture, we focused on three visually salient ones with large head and body displacements: beak wipes (BW), turn-arounds (TA) and hops (Fig. A1, Supplementary Videos S1 and S2). Morris (1957, p. 4) wrote that a beak wipe ‘consists of (a) twist body (b) lower and rotate head (c) scrape [beak] (d) raise head and (e) twist body’. Beak wipes vary in completeness, from slight nods or bows to double wiping of the beak (Morris, 1957, p. 9; Zann, 1996, p. 170). Although Morris distinguished ‘displacement beak wipes’ (incomplete; shorter duration) from ‘autochthonous’ (complete; longer duration) ones, he does not score them as two qualitatively different movements, but sees a close relationship between them (Morris, 1954, pp. 307–311). We followed the judgement when we compared the durations of beak wipes during song and nonsong and consequently pooled all manifestations of the movement. Morris also described turn-arounds: ‘As the male advances towards the female down the branch, it swings its body from side to side, turning first to the left and then to the right, changing the position of its feet as it does so’ (Morris, 1954, p. 285). During this manoeuvre the male twists his head and tail towards the female (Barclay et al., 1992). Analogous to gesture analysis in human communication (McNeil, 1992, p. 131), we subdivided the BW movement into a preparation phase (lowering the head), stroke (scraping the beak on the perch or wiping above, respectively) and poststroke phase (lifting the head) and quantified occurrences and durations. Likewise, TA movements were divided into preparation (beginning of the turn), stroke (both feet displaced from the perch) and post stroke phase (completion of the turn). Hop movements were scored for the 10 birds of set-up C as events, counting the video frame when both feet were in the air above the perch as the stroke of the hop and used this as the time point for hop quantification (Fig. A1). Video scorings were made

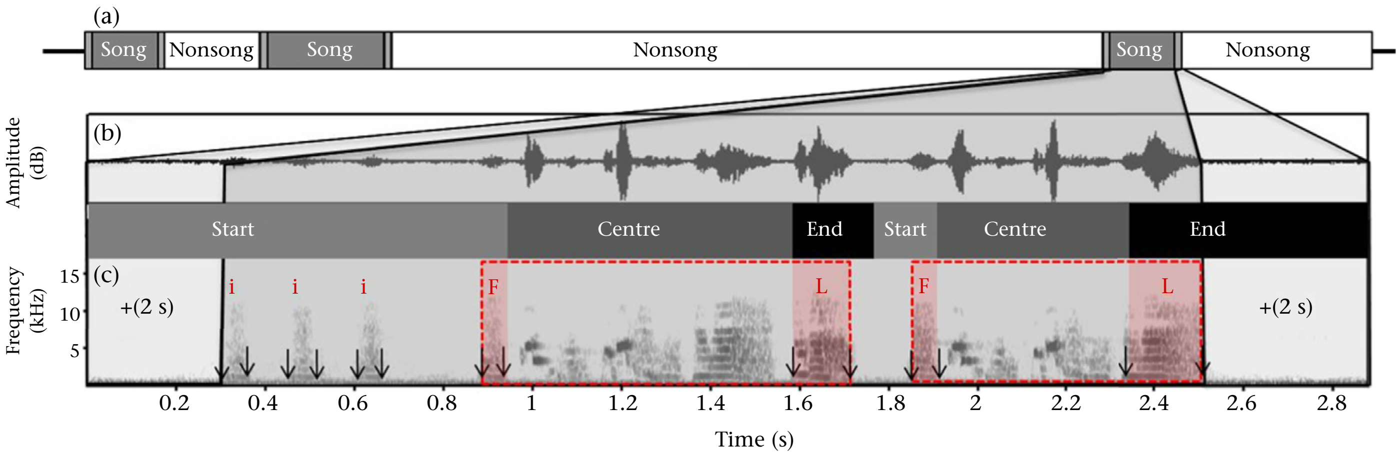


**Figure 3.** Schematic depiction of audio scoring. (a) The 5 min that males were recorded included song (dark grey boxes) and nonsong (white boxes) segments. (b) Waveform and (c) spectrogram of one bout, from male 3641. We defined song segments as the time span comprising song and the 2 s immediately before and after (indicated by light grey bars in (a)). A song was defined as lasting from the start of the first introductory note (i), containing one or more motifs (outlined by red hatched lines in (c)) to 2 s after the last note of a succession of motifs. Motifs are a stereotyped, recurring sequence of notes; first (F) and last (L) notes of motifs are indicated by red overlay. Consecutive motifs were considered to belong to the same bout if they were separated by less than 2 s of silence or call notes (c). The dark grey bar between (b) and (c) illustrates the definition of 'start', 'centre' and 'end' of songs. The 'start' of a bout was defined as lasting from 2 s before the first introductory note to the end of the first motif note, F (light grey bar). For consecutive motifs the 'start' included half of the interval between successive motifs and the duration of the last note (L) of the preceding motif. The 'end' of a bout or motif was respectively defined as the last note (L) and half of the succeeding interval before the end of the first note of the following motif. We extracted the time stamps (indicated by downward pointing arrows in (c)) of the beginning and end of introductory notes (i), first notes (F) and last notes (L), permitting calculation of note duration and of the duration of 'start', 'centre' and 'end' segments.

blind with respect to song behaviour, e.g. without listening to the sound track.

To analyse whether hopping coincided with utterances, we videorecorded with a medium high-speed camera (60 fps) from the front of the cage during experiment C ($N = 10$). We measured how closely in time hops occurred to introductory notes (i), the first (F) and the last note (L) of a motif. We also included any calls between bouts and calls during the 'nonsong' segments as well. Hops taking place more than 1 s before or after a call were not considered. Hops occurring between notes were assigned to the note that was closer in time. 'Call-associated' hops fell within a 225 ms time window around the time stamp of the call. We chose this time window because many hops occurred immediately before, during or after a call (see Fig. 8c in the Results). The average call length was determined empirically to be 75 ms (see Appendix).

To assess whether a bird's hop and call might coincide by chance, we distributed both events randomly and compared the coincidence of hop and call by chance with the actual observation using Monte Carlo simulation. To examine the sequential order of BW, TA and hops during song we transcribed strings of the three movements following each other with intervals shorter than 2 s in all birds from experiment C ($N=10$).

Statistical analysis and figures were prepared using R (version 2.14.0; R Development Core Team, 2011) plus R packages 'nlme' (Pinheiro, Bates, DebRoy, Sarkar, & R Development Core Team R, 2011), 'multcomp' (Hothorn, Bretz, & Westfall, 2008), 'lme4' (Bates, Maechler, & Bolker, 2012) and 'car' (Fox & Weisberg, 2011). For analysing the allocation of BW, TA and hopping in the surrounding of a bout we ran a linear mixed model (LMM). We tested whether the averaged occurrences of body movements per second (dependent variable) differed between locations within bouts (i.e. 'start', 'centre', 'end'). As a fixed effect we entered the location into the model and as a random effect we added an intercept for the birds used in the experiment as well as a by-bird random slope for the effect of location. We used Akaike's information criterion (AIC) for model comparisons (i.e. linear regression, random slope or intercept model, or ANCOVA with or without interaction) and used the model with the lowest AIC value (i.e. LMM random slope). LMMs were fitted using the 'lmer' function in the 'lme4' library. The analysis was followed by Tukey's multiple comparisons post hoc test. For the remaining analysis we made use of paired *t* tests. Correlation analyses were performed by use of simple linear regression models, which also allowed us to extract the coefficient of determination ($R^2$). We used a chi-square test of goodness-of-fit for the overall difference between observed and expected transition ratios. Differences between observed and expected ratios for individual transition types between movements were tested by LMM with a random intercept for birds. *P* values were obtained by likelihood ratio tests. In cases of multiple testing on the same data, false discovery rate was controlled using the Benjamini–Hochberg procedure. Results reported as significant assume a false discovery rate of 0.05. Data were tested for normal distribution using the Shapiro-Wilk test and cube-root transformed (see Appendix) beforehand if the initial data were not normally distributed.

### *Ethical Note*

The study with domesticated zebra finches conforms to the ASAB/ABS guidelines for the use of animals in research as well as to the legal requirements of the German LAGESO board (permit ZH 144). During the study, stress was minimized and all birds were cared for and treated appropriately in accordance with the German laws for animal experiments. Birds used for the recording showed normal behaviour and did not appear to suffer from their occasional removal from the group. After the study, birds were used either for other experiments or for breeding purposes.

## RESULTS

Courtship song and dance was elicited by allowing visual and acoustic but no physical contact between a female and the focal male. The behaviour of the male was audio and video recorded for 5 min. For the video analysis we divided beak wipe (BW) and turn-around movements (TA) into a preparation phase, stroke and poststroke phase (Fig. A1). Hop movements were scored as point events (Fig. A1). To investigate whether hops, BW and TA were coordinated with song, we quantified occurrences and timing of these movements both during song and nonsong segments (Fig. 3).

### *Body Movements were Associated with Song*

All three body movements occurred significantly more often per second during song than nonsong (Fig. 4a–c). Interestingly, birds initiated and ended song motifs significantly more often with BW, TA and hops than in the 'centre' of motifs (Fig. 4d–f; for definition of motif divisions, see Methods, Fig. 3). Also, BW and hops, but not TA, were more numerous at the 'start' than at the 'end' of motifs (Fig. 4d–f). These findings were consistent with coordination between song and dance.

### *During Song, BW, but not TA, were Performed Faster*

If body movements were coordinated with song they might be performed at different durations during singing and nonsinging. To assess this, we compared the durations of BW and TA in the two conditions. Indeed, BW were performed roughly 20% (111 ms) faster during singing, whereas the duration of TA did not differ (Fig. 5). Detailed frame-by-frame analysis revealed that the shorter BW duration during song was due to a faster stroke phase of the BW, while preparation and poststroke phases lasted comparable amounts of time during song and during silence. This suggests that birds can modulate the duration of BW during song, facilitating integration of song and body movements.

### *BW and TA were Correlated with Amount of Song*

To further pursue whether song and body movements are coordinated, we analysed individual birds. Indeed, the more motifs a male sang the more BW and TA he performed (Fig. 6a and b). In contrast, the number of hops was not correlated with the amount of singing (Fig. 6c), even though hops occurred significantly more often during song than nonsong (Fig. 4c). Comparing the number of movements to the number of bouts sung by each bird yielded very similar results (Pearson product moment correlation: BW versus bout: $t_{18} = 2.45$, $P = 0.025$; TA versus bout: $t_{18} = 2.99$, $P = 0.0077$; hop versus bout: $t_8 = -0.098$, $P = 0.92$; Fig. A2). Thus, BW and TA appear to be more directly associated with song than hops.

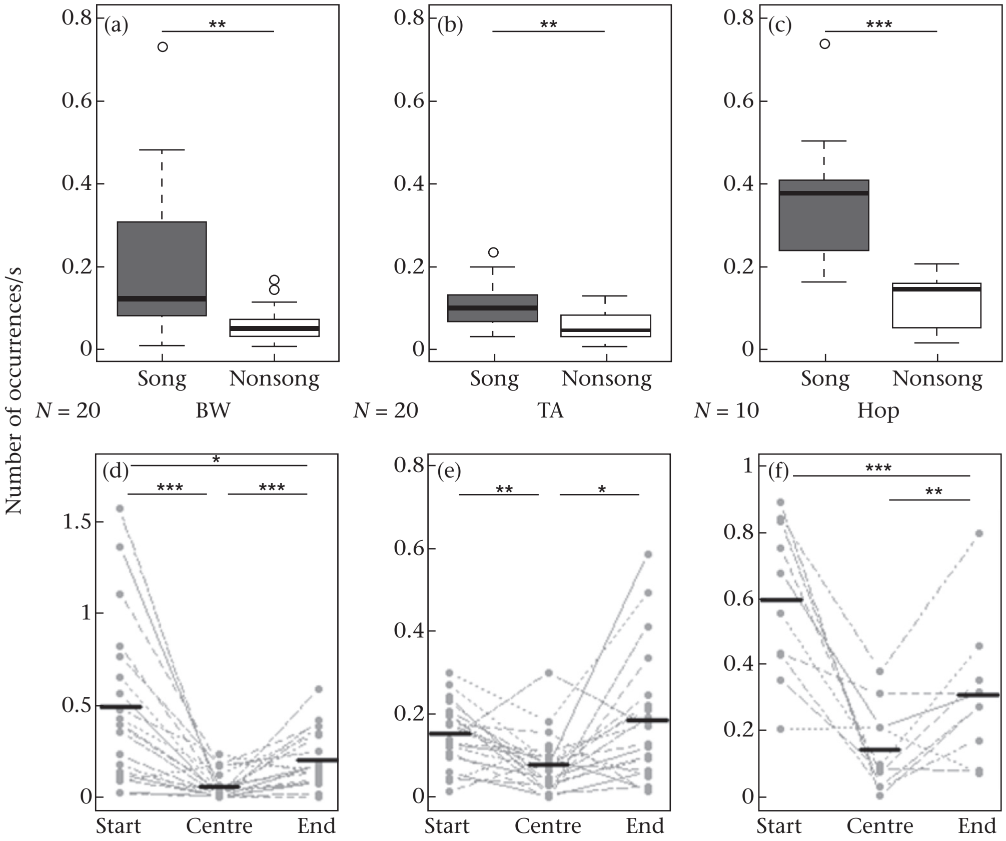


**Figure 4.** (a–c) Number of body movements/s during song and nonsong segments. Box plots show median (black, horizontal line), 25% and 75% quartiles (box segment below or above median), outermost values within the range of 1.5 times the respective quartiles (whiskers) and outliers (points). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Paired $t$ test: (a) BW song versus nonsong: $t_{19} = 4.55$, $P < 0.01$; (b) TA song versus nonsong: $t_{19} = 4.2$, $P < 0.01$; (c) hops song versus nonsong: $t_9 = 5.14$, $P < 0.001$. Number of (d) BW, (e) TA and (f) hops during the 'start', 'centre' and 'end' of song motifs. Grey points represent the birds' average of occurrences in all videos. Grey lines connect data of the same individual. Black horizontal bars depict the group's average. (d) Linear mixed model (LMM): 'centre' versus 'start': estimate = −0.43, SE = 0.058, $z = -7.43$, $P < 0.001$; 'end' versus 'start': estimate = −0.14, SE = 0.049, $z = -2.94$, $P = 0.009$; 'end' versus 'centre': estimate = 0.29, SE = 0.058, $z = 5.01$, $P < 0.001$. (e) LMM: 'centre' versus 'start': estimate = −0.14, SE = 0.042, $z = -3.24$, $P = 0.003$; 'end' versus 'start': estimate = 0.012, SE = 0.039, $z = 0.31$, $P = 0.94$; 'end' versus 'centre': estimate = 0.15, SE = 0.05, $z = 2.71$, $P = 0.018$. (f) LMM: 'end' versus 'centre': estimate = 0.16, SE = 0.06, $z = 2.78$, $P = 0.014$; 'centre' versus 'start': estimate = −0.39, SE = 0.08, $z = -4.96$, $P < 0.001$; 'end' versus 'start': estimate = −0.24, SE = 0.05, $z = -4.52$, $P < 0.001$.

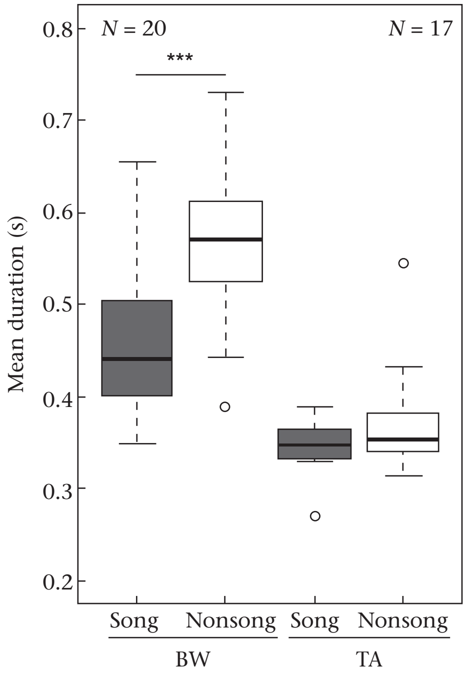


**Figure 5.** Duration of BW and TA movements during song and nonsong. For each bird the average duration of BW ($N$=20 birds) and TA ($N$=17 birds) were calculated. ***$P < 0.001$. Paired $t$ test: BW: $t_{19} = -6.15$, $P < 0.001$; TA: $t_{16} = -1.33$, $P = 0.20$). For box plot specifications see legend to Fig. 4.

### *Hop Rates During Song Correlated with those Outside Song*

Individuals varied considerably in frequency of BW, TA and hops during song. We wondered whether an individual's propensity to perform more or less song-associated movements was related to how often these movements occurred when they were not singing. Interestingly, the numbers of hops (Fig. A3c) during song and nonsong were significantly correlated whereas this was not the case for BW (Fig. A3a) and TA (Fig. A3b).

### *During Song, Numbers of BW and TA were Tightly Related*

Given that all three movements occurred more frequently with song than outside the song context (Fig. 4a–c), we checked whether all three movements are part of an integrated song choreography. Consistent with this notion, during song but not outside the song context the number of BW and TA were more related to each other than TA were with hops or BW with hops (Fig. 7a–c; Pearson product moment correlation: song: TA versus BW: $t_{18} = 2.04$, $P = 0.056$; TA versus hop: $t_8 = 0.16$, $P = 0.88$; BW versus hop: $t_8 = -0.17$, $P = 0.87$; nonsong: BW versus TA: $t_{18} = -0.08$, $P = 0.93$; hop versus TA: $t_8 = 1.16$, $P = 0.28$; 9. hop versus BW: $t_8 = 0.38$, $P = 0.71$). This finding, together with the results shown in Fig. 6, raised the possibility that BW and TA are a more tightly integrated part of the song choreography than hops.

### *Hops Coincided with Particular Notes*

To determine whether hops were in fact less coordinated with song than BW and TA, we quantified when hops occurred in relationship to song using frame by frame video analysis. Because hopping behaviour is often accompanied by calls (Moorman & Bolhuis, 2013, p. 112; Zann, 1996, p. 197), we measured the interval between a hop and the introductory, the first and the last motif notes. We found that most hops coincided exactly or closely in time with the utterance of a note or call. This was particularly striking for the introductory notes (Fig. 8a). Note, however, that hops and vocalizations were not obligatorily coupled.

Might the striking coincidence of hops and introductory notes stem from physical constraints of hopping, such as a minimum interval between hops mandated by movement mechanics? To answer this question, we analysed how the intervals between two hops were distributed. Only two hop intervals out of 588 were shorter than 200 ms, probably reflecting the limit of how fast two successive hops can be performed. During song, 49% of all interhop intervals ranged between 200 and 1000 ms (Fig. 8b). Interhop intervals of this range of durations were between 200 and 400 ms long. Hops of this duration were almost twice as numerous during song than during nonsong (24% versus 13%). Although birds performed a very wide range of interhop intervals during song (see rightmost columns, Fig. 8b), we observed 24% of them within the same time span as interintroductory note intervals (mean ± SD = 255 ± 215 ms). Although in both contexts hop intervals between 200 and 400 ms were most frequent, the high coincidence of hops with introductory notes therefore does not appear to be a simple by-product of a narrow range of hop interval durations mandated by physical constraints (Fig. 8a). Instead, it is consistent with the possibility that hops and notes coincide as a consequence of voluntary control.

Were hops and calls also associated when zebra finches were not singing? We analysed the duration between a hop and the call closest in time and found that 56% of all hopping behaviour clustered in a 2 s time window around calls (i.e. 1 s before plus 1 s after). Twenty per cent of all hops were directly associated with calls (Fig. 8c and see the Appendix). We ruled out the possibility that the close association of hops and calls was a function of two independent events being associated coincidentally; hops and calls were significantly more often linked in time than would be expected by chance (Fig. 8d, hop estimated versus observed: one-tailed, paired $t$ test: $t_9 = -3.14$, $P < 0.01$).

### *Behaviour Transition Ratios Differed from those Expected*

To analyse the sequence of movements we transcribed the order of transitions between BW, TA and hops following each other with intervals shorter than 2 s. Longer intervals were considered the start and end of a transition string. Fig. 9 shows the sequence diagram of movement transition probabilities across all 10 birds from experiment C during bouts. Transition probabilities that would be expected if the three movements were equally likely to follow each other were calculated based on the observed number of movements and transition strings (see Appendix for details). Observed probabilities of all transition types in their entirety differed significantly from the expected values ($\chi^2_{14} = 116.78$, $P < 0.001$). Transitions from TA to TA (LMM: $\chi^2_1 = 11.92$, $P = 0.0005$, difference $= -0.11 \pm 0.02$ SE) and from hop to TA (LMM: $\chi^2_1 = 8.64$, $P = 0.0033$, difference $= -0.07 \pm 0.02$ SE) occurred significantly less frequently than expected by chance, while transitions from hop to hop were significantly more frequent (LMM: $\chi^2_1 = 10.72$, $P = 0.001$, difference $= +0.15 \pm 0.03$ SE; see also Figs. A4 and A5).

In individual birds, transition string length varied between 2 and 17 movements. Not all possible movement combinations

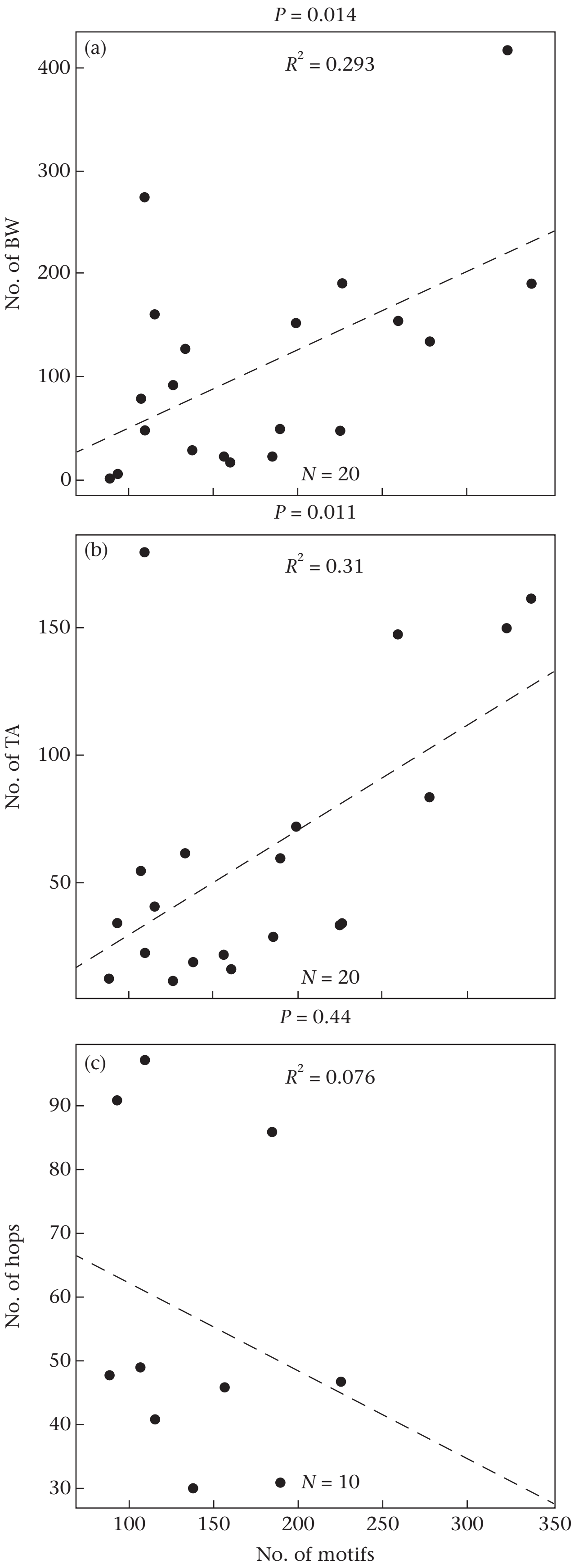


**Figure 6.** Number of (a) BW, (b) TA and (c) hops in relation to number of motifs sung. Pearson product moment correlation: BW versus motif: $t_{18} = 2.7$, $P < 0.05$; TA versus motif: $t_{18} = 2.84$, $P < 0.05$; hop versus motif: $t_8 = -0.81$, $P = 0.44$.

occurred and some were particularly frequent (Table A2). Consistent with our hypothesis that sequences of movements during song are part of a choreography, the frequency distribution of string length was skewed towards longer strings during song than during nonsong (Fig. A6).

Finally, we examined whether age (and by inference, experience) correlated with dance vigour. We analysed the number of BW, TA and hops in relation to age but did not find a systematic association, as reported before (Williams, 2001; Pearson product moment correlation: age versus BW: $t_{18} = -0.14$, $P = 0.88$; age versus TA: $t_{18} = -1.5$, $P = 0.15$; age versus hop: $t_8 = -0.84$, $P = 0.43$).

## DISCUSSION

Here we comprehensively show for the first time multiple ways in which the expression of dance is strongly, but not obligatorily, associated with the expression of song in zebra finches. Stereotypic movements, that is, BW, TA and hops, occurred significantly more during song than during silence. All three movements clustered at the start and end of motifs and hops coincided with notes, particularly the introductory but also the first and last motif notes. BW were performed faster during song than during silence but TA were not. The more birds sang, the more BW and TA occurred. Dance sequences were flexibly but not randomly arranged. Together, these results suggest a choreography of song and dance that signals courtship intensity. A previous study (Williams, 2001) did not find a strongly patterned association between song and dance movements, which might be because hops, forward, left, right, back, up and down head movements were pooled. However, Williams (2001, p. 3505) observed qualitatively that ‘the relatively infrequent dance movements appear to be initiated at a number of specific ‘hot spots’ within the song’.

Clustering of BW, TA and hops at the beginning and end of motifs may serve as ‘initiation and closure’ signals. The male could draw attention to his courtship before the relatively quiet song becomes audible to the receivers. The repeated 180° turning of the bird while singing as he approaches the female is associated with a regular change in song sound amplitude. This could also allow her to discriminate the courting male's sound pattern from the noisy environment of the vocalizing flock mates and provide an example of ‘spatial release from masking’ (Bee & Micheyl, 2008). Gestures such as BW and TA at the beginning and end of motifs might also focus the female's attention on the motif structure and thus aid in the evaluation of song stereotypy, a feature relevant for female choice (Riebel, 2009). When analysing the location of movements within the motif, we did not distinguish the very first and last from the other motifs of a bout. However, the number of movements a bird performed correlated with both the motif and bout count to a very similar degree, suggesting that the placement of body movements is not primarily tied to the beginning or end of a song bout.

Courtship intensity in zebra finch males is influenced by the social background (Ruploh, Bischof, & Engelhardt, 2012), the size of the experimental cage, the experimental procedure itself (Immelmann, 1959, p. 447), dopamine turnover in brain areas controlling the motor patterning of song (Rauceo et al., 2008) and the female's reaction (Zann, 1996, pp. 171–173). As we used different cage sizes and experimental set-ups our results encompass this variability, yet particular movements were consistently and significantly associated with courtship song. A vigorously courting male sings more motifs per unit time (Riebel, 2009) and we show that the number of song motifs in individual males also strongly predicts their BW and TA activity. Moreover,

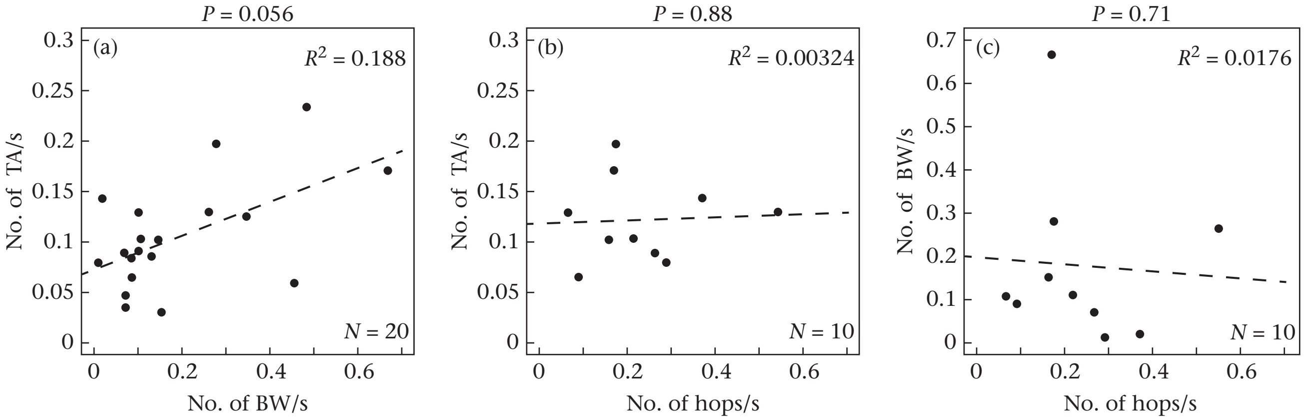


**Figure 7.** Correlations between (a) number of TA and BW, (b) number of TA and hops and (c) number of BW and hops. Hops performed as part of a TA were removed prior to TA analysis. See text for statistics.

TA were tightly related with BW during singing but not outside the song context. Together, these results are consistent with the notion that TA and BW indicate intensity of courtship in a linear fashion.

Interindividual differences were not only observed in courtship intensity, but also existed in overall activity levels. For instance, the number of hops performed during the time when birds were not singing was predictive of the number of hops during song.

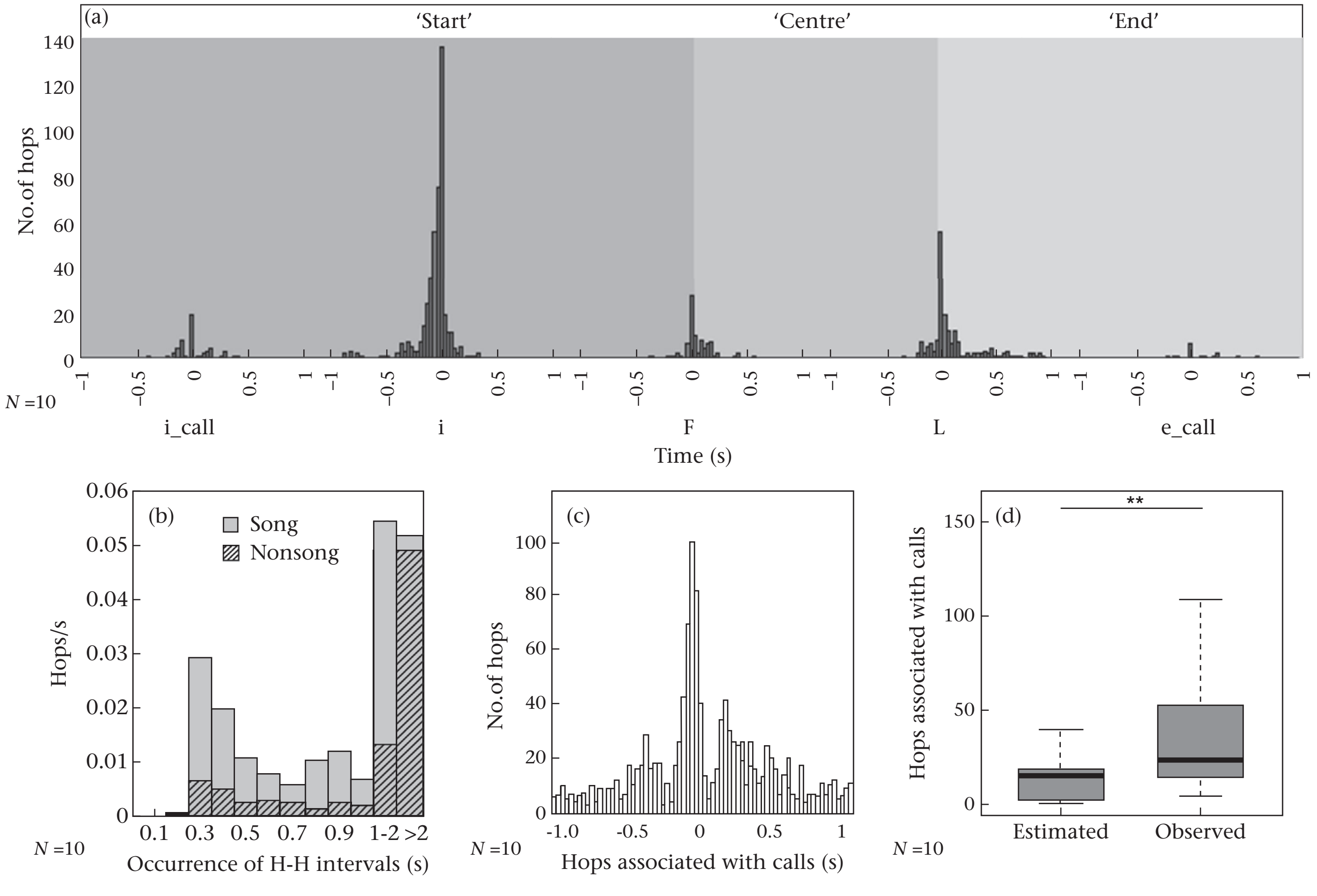
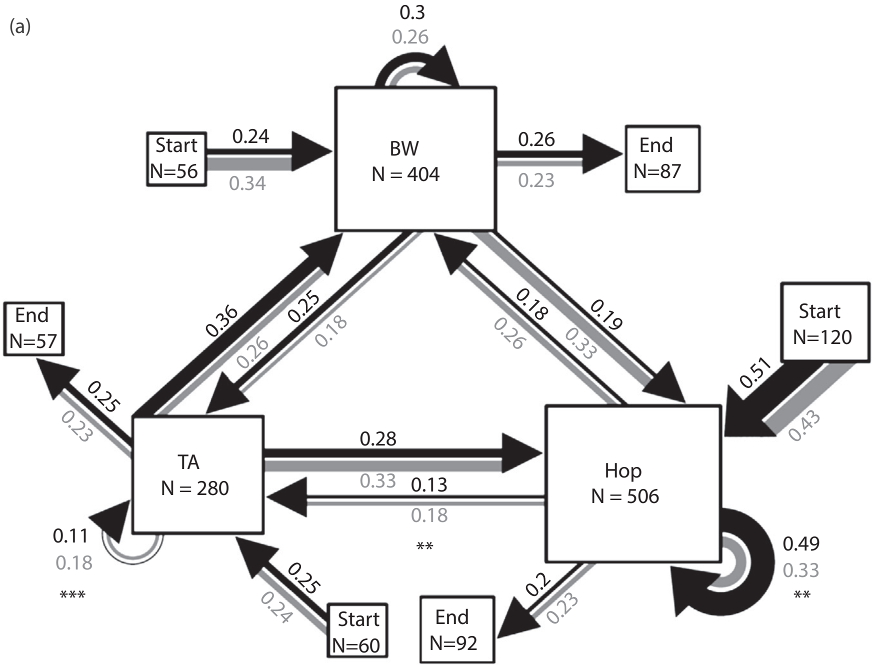


**Figure 8.** (a) Frequency histogram depicting the number of hops associated with song notes. We scored a hop as coincident with a song note when the 16 ms video frame that showed both feet displaced from the perch during a hop fell within the duration of a note (0 on *x*-axis). For hops occurring before (minus numbers on *x*-axis) or after (positive numbers on *x*-axis) a note, the duration between the hop and the closest note are depicted in 30 ms bins. The larger song segments ('start', 'centre' 'end', see Figs. 2 and 3) are highlighted by different grey shadings. (b) Histogram depicting the distribution of intervals between hops (H–H intervals) performed during song (shaded grey) and nonsong (striped). H–H song versus H–H nonsong: paired *t* test: $t_9 = 3.6$, $P < 0.01$. The remaining intervals are pooled in the two rightmost columns. (c) Histogram illustrating the association between calls (plotted as time 0) and hops occurring outside the song context. (d) Based on the number and duration of hops and calls the estimated coincidence of these two events was calculated. See text for statistics. For box plot specifications see legend of Fig. 4.



(b) Start – BW – TA – BW – End

Start – hop – hop – hop – hop – TA – End

Start – BW – TA – BW – TA – hop – TA – End

**Figure 9.** (a) Sequence diagram of BW, TA, and hops, as well as start and end points of dance, during song of 10 birds. Observed transition probabilities (black) are indicated by arrows. Their size corresponds to the numerical values indicated next to the arrow. Values refer to the fraction of occurrence of a transition in relation to all transitions originating at the same node. Grey arrows and numbers indicate the calculated values expected if the three movements were equally likely to follow each other, given the observed number of movements and transition strings. The box size is scaled to the number of movements, starts and ends, respectively ($N$ = inside the boxes). Asterisks mark transitions that differed significantly in observed versus expected probabilities: $^{**}P < 0.01$; $^{***}P < 0.001$. (b) Three representative sequences of movements during song are shown. For more examples of transition strings and frequency of occurrence see Table A2.

However, this was not the case for BW and TA, indicating that both movements seem to be more specifically related to courtship activity.

Together, our results provide quantitative evidence that the zebra finch courtship dance is coordinated temporally with courtship song. As such it could constitute an honest multimodal signal that allows females to evaluate male fitness during mate choice. Song production depends on both peripheral components (including the lungs, air sacs, respiratory muscles, the vocal organ (syrinx) and the upper vocal tract) and central neural control by respiratory-vocal circuits (Schmidt & Wild, 2014; Suthers, Goller, & Pytte, 1999). In fact, a previous study on zebra finches demonstrated that during song, oxygen consumption increases compared to nonsong (Franz & Goller, 2003). This could explain the observed absence of bodily movements within motifs as a result of physical constraints. However, the study happened under the exclusion of dance, which limits its results for interpretation of the current results. Because dance movements are similarly subject to peripheral and central motor constraints and have to be coordinated with song, multimodal signalling could provide information about the developmental history or current physical fitness of the male (Byers, Hebets, & Podos, 2010; Cotton, Small, & Pomiankowski, 2006; Fuxjager, Longpre, Chew, Fusani, & Schlinger, 2013). In fact, developmental nutritional stress does affect song learning in various songbird species (Brumm, Zollinger, & Slater, 2009; Buchanan, Leitner, Spencer, Goldsmith, & Catchpole, 2004; Buchanan, Spencer, Goldsmith, & Catchpole, 2003; Nowicki, Searcy, & Peters, 2002; Zann & Cash, 2007), but to date dance or other motor behaviours have not been investigated in this context. In humans, developmental disturbances are also linked to deficits in integrating speech with gestures (Iverson, 2010).

The current study provides fertile ground to explore the neural substrates underlying courtship dance and its integration with song. The fact that the song nucleus LMAN is required for the acoustic differences between undirected and directed song but not needed for courtship dancing emphasizes that song and dance are independently controlled, but integrated, motor behaviours (Kao & Brainard, 2006). Candidate regions for the control of the dance movements are the motor regions in close proximity to the song control nuclei (Feenders et al., 2008). In male golden-collared manakins, famous for their elaborate courtship dances, volume differences in brain areas of males and nondancing females point towards the arcopallium and hippocampus as candidate control regions (Schlinger, Barske, Day, Fusani, & Fuxjager, 2013).

Further studies in this subject are also promising because one prominent theory of language evolution posits a gestural origin (Arbib, Liebal, & Pika, 2008; Armstrong, Stokoe, & Wilcox, 1995). This notion is supported by observations that language and action systems in the brain overlap (Willems, Ozyürek, & Hagoort, 2007). Thus investigating the neural interconnection of vocalization and gesture in birds could extend the numerous parallels between human speech and birdsong (Bolhuis, Okanoya, & Scharff, 2010; Ohms, Gill, Van Heijningen, Beckers, & ten Cate, 2010). This is particularly interesting, because human gestures are a universal feature of communication and tightly timed with speech (McNeil, 1992, p. 1). Moreover, our finding that the stroke of a hop (e.g. when both feet are in the air) coincides precisely with, or just before, the introductory notes of the song is reminiscent of the fact that in people the stroke of a gesture occurs shortly before or directly on ‘the prosodic stress peak of the accompanying spoken utterance’ (McNeil, 1992, p. 131). Because gestures in human language are so closely associated with semantics it would also be interesting to extend our kind of study to other behavioural contexts, such as parental-offspring interactions (Goldstein, King, & West, 2003), and other areas of social learning (Katsnelson, Motro, Feldman, & Lotem, 2008) with an eye on the differential ‘meaning’ of same-sound/different-gesture combinations (O'Loghlen & Rothstein, 2010).

In summary, our findings provide a comprehensive analysis of gestural signalling in male zebra finches while courting. Our results invite follow-up studies into the consequences and mechanisms of multimodal signalling in songbirds.

## Acknowledgments

We thank F. Nottebohm, H. Hultsch and H. Williams and two anonymous referees for critical comments on the manuscript. We are indebted to M. Weiss for statistical support. For assistance with the video recordings, we thank M. Kloß and M. Gessinger. Also thanks to excellent bird care by Janett Birkenfeld. We acknowledge the support of a consortium grant of the German Federal Ministry of Education and Research (BMBF; ‘Varying Tunes’, 01 GQ 0963) as well as a grant from SFB 665: Developmental Disturbances in the Nervous System.

## Supplementary Material

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.anbehav.2015.11.012.

## References

Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate vocalization, gesture, and the evolution of human language. *Current Anthropology, 49*(6), 1053–1076. http://dx.doi.org/10.1086/593015.

Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language*. Cambridge, NY: Cambridge University Press.

Barclay, S. R., Harding, C. F., & Waterman, S. (1992). Correlations between catecholamine levels and sexual behavior in male zebra finches. *Pharmacology, Biochemistry, and Behavior, 41*(1), 195–201.

Barske, J., Schlinger, B. A., Wikelski, M., & Fusani, L. (2011 April). Female choice for male motor skills. *Proceedings of the Royal Society B: Biological Sciences, 278*, 3523–3528. http://dx.doi.org/10.1098/rspb.2011.0382. rspb20110382.

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. Retrieved from http://cran.r-project.org/package=lme4.

Bee, M. A., & Micheyl, C. (2008). The ‘Cocktail Party Problem’: what is it? How can it be solved? and why should animal behaviorists study it? *Journal of Comparative Psychology, 122*(3), 235–251. http://dx.doi.org/10.1037/0735-7036.122.3.235.

Böhner, J., & Veit, F. (1993). Gesangsstruktur und Formen der Flügelbewegung beim Star (*Sturnus vulgaris*). *Journal of Ornithology, 134*(3), 309–315.

Bolhuis, J. J., Okanoya, K., & Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nature Reviews. Neuroscience, 11*(11), 747–759. http://dx.doi.org/10.1038/nrn2931.

Bostwick, K. S. (2000). Display behaviors, mechanical sounds, and evolutionary relationships of the club-winged manakin (*Machaeropterus deliciosus*). *Auk, 117*(2), 465–478. http://dx.doi.org/10.1642/0004-8038(2000)117[0465:DBMSAE]2.0.CO;2.

Brumm, H., Zollinger, S., & Slater, P. (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. *Behavioral Ecology and Sociobiology, 63*(9), 1387–1395. http://dx.doi.org/10.1007/s00265-009-0749-y.

Buchanan, K. L., Leitner, S., Spencer, K. A., Goldsmith, A. R., & Catchpole, C. K. (2004). Developmental stress selectively affects the song control nucleus HVC in the zebra finch. *Proceedings of the Royal Society B: Biological Sciences, 271*(1555), 2381–2386. http://dx.doi.org/10.1098/rspb.2004.2874.

Buchanan, K. L., Spencer, K. A., Goldsmith, A. R., & Catchpole, C. K. (2003). Song as an honest signal of past developmental stress in the European starling (*Sturnus vulgaris*). *Proceedings of the Royal Society B: Biological Sciences, 270*(1520), 1149–1156. http://dx.doi.org/10.1098/rspb.2003.2330.

Byers, J., Hebets, E., & Podos, J. (2010). Female mate choice based upon male motor performance. *Animal Behaviour, 79*(4), 771–778. http://dx.doi.org/10.1016/j.anbehav.2010.01.009.

Caryl, P. (1981). The relationship between the motivation of directed and undirected song in the zebra finch. *Zeitschrift für Tierpsychologie, 57*, 37–50.

Cooper, B. G., & Goller, F. (2004). Multimodal signals: enhancement and constraint of song motor patterns by visual display. *Science (New York, N.Y.), 303*(5657), 544–546. http://dx.doi.org/10.1126/science.1091099.

Cooper, B. G., & Goller, F. (2006). Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *Journal of Neurophysiology, 95*(6), 3798–3809. http://dx.doi.org/10.1152/jn.01123.2005.

Cotton, S., Small, J., & Pomiankowski, A. (2006). Sexual selection and condition-dependent mate preferences. *Current Biology: CB, 16*(17), R755–R765. http://dx.doi.org/10.1016/j.cub.2006.08.022.

Dalziell, A., Peters, R., Cockburn, A., Dorland, A. D., Maisey, A. C., & Magrath, R. D. (2013). Dance choreography is coordinated with song repertoire in a complex avian display. *Current Biology, 23*(12), 1132–1135. http://dx.doi.org/10.1016/j.cub.2013.05.018.

Dunn, A. M., & Zann, R. A. (1996). Undirected song in wild zebra finch flocks. *Ethology, 102*(4), 529–539. http://dx.doi.org/10.1111/j.1439-0310.1996.tb01145.x.

Elias, D. O. (2006). Female preference for complex/novel signals in a spider. *Behavioral Ecology, 17*(5), 765–771. http://dx.doi.org/10.1093/beheco/arl005.

Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., et al. (2008). Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. *PLoS One, 3*(3), e1768. http://dx.doi.org/10.1371/journal.pone.0001768.

Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression*. Thousand Oaks, CA: Sage. Retrieved from https://cran.r-project.org/web/packages/car/index.html.

Franz, M., & Goller, F. (2003). Respiratory patterns and oxygen consumption in singing zebra finches. *The Journal of Experimental Biology, 206*(Pt 6), 967–978. http://dx.doi.org/10.1242/jeb.00196.

Fusani, L., Hutchison, R. E., & Hutchison, J. B. (1997). Vocal-postural co-ordination of a sexually dimorphic display in a monomorphic species: the Barbary dove. *Behaviour, 134*(5), 321–335.

Fuxjager, M. J., Longpre, K. M., Chew, J. G., Fusani, L., & Schlinger, B. A. (2013). Peripheral androgen receptors sustain the acrobatics and fine motor skill of elaborate male courtship. *Endocrinology, 154*(9), 3168–3177. http://dx.doi.org/10.1210/en.2013-1302.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology, 64*(c), 257–283. http://dx.doi.org/10.1146/annurev-psych-113011-143802.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America, 100*(13), 8030–8035. http://dx.doi.org/10.1073/pnas.1332441100.

Hoepfner, A. R., & Goller, F. (2013). Atypical song reveals spontaneously developing coordination between multi-modal signals in brown-headed cowbirds (*Molothrus ater*). *PLoS One, 8*(6), 1–8. http://dx.doi.org/10.1371/journal.pone.0065525.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346–363.

Immelmann, K. (1959). Experimentelle Untersuchungen über die biologische Bedeutung artspezifischer Merkmale beim Zebrafinken (*Taeniopygia castanotis* Gould). *Zoologisches Jahrbuch Abteilung Für Systematik, 86*, 435–592.

Immelmann, K. (1969). *Der Zebrafink (*Taeniopygia guttata*)*. Stuttgart, Germany: Franckh'sche Verlagshandlung.

Iverson, J. M. (2010). Multimodality in infancy: vocal-motor and speech-gesture coordinations in typical and atypical development. *Enfance; Psychologie, Pedagogie, Neuropsychiatrie, Sociologie, 2010*(3), 257–274. http://dx.doi.org/10.4074/S0013754510003046.

Jarvis, E. D., Scharff, C., Grossman, M. R., Ramos, J. A., & Nottebohm, F. (1998). For whom the bird sings: context-dependent gene expression. *Neuron, 21*(4), 775–788. http://dx.doi.org/10.1016/S0896-6273(00)80594-2.

Kao, M. H., & Brainard, M. S. (2006). Lesions of an avian basal ganglia circuit prevent context-dependent changes to song variability. *Journal of Neurophysiology, 96*(3), 1441–1455. http://dx.doi.org/10.1152/jn.01138.2005.

Katsnelson, E., Motro, U., Feldman, M. W., & Lotem, A. (2008). Early experience affects producer–scrounger foraging tendencies in the house sparrow. *Animal Behaviour, 75*(4), 1465–1472. http://dx.doi.org/10.1016/j.anbehav.2007.09.020.

Kunkel, P. (1959). Zum Verhalten einiger Prachtfinken (Estrilidinae). *Zeitschrift für Tierpsychologie, 16*(3), 302–350.

Lukianchuk, K. C., & Doucet, S. M. (2014). Cooperative courtship display in long-tailed manakins *Chiroxiphia linearis*: predictors of courtship success revealed through full characterization of display. *Journal of Ornithology, 155*(3), 729–743. http://dx.doi.org/10.1007/s10336-014-1059-3.

McNeil, D. (1992). *Hand and mind. What gestures reveal about thought. What gestures reveal about*. Chicago, IL: The University of Chicago Press.

Moorman, S., & Bolhuis, J. J. (2013). Behavioral similarities between birdsong and spoken language. In J. J. Bolhuis, & M. Everaert (Eds.), *Birdsong, speech, and* *language: exploring the evolution of mind and brain* (pp. 111–123). Cambridge, MA: MIT Press.

Morris, D. (1954). The reproductive behaviour of the Zebra Finch (*Poephila guttata*), with special reference to pseudofemale behaviour and displacement activities. *Behaviour, 6*(4), 271–322.

Morris, D. (1957). 'Typical Intensity' and its relation to the problem of ritualisation. *Behaviour, 11*(1), 1–12.

Nowicki, S., Searcy, W. A., & Peters, S. (2002). Brain development, song learning and mate choice in birds: a review and experimental test of the 'nutritional stress hypothesis'. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology, 188*(11–12), 1003–1014. http://dx.doi.org/10.1007/s00359-002-0361-3.

Ohms, V. R., Gill, A., Van Heijningen, C. A., Beckers, G. J., & ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society B: Biological Sciences, 277*(1684), 1003–1009. http://dx.doi.org/10.1098/rspb.2009.1788.

O'Loghlen, A. L., & Rothstein, S. I. (2010). Multimodal signalling in a songbird: male audiovisual displays vary significantly by social context in brown-headed cowbirds. *Animal Behaviour, 79*(6), 1285–1292. http://dx.doi.org/10.1016/j.anbehav.2010.03.001.

O'Loghlen, A. L., & Rothstein, S. I. (2012). When less is best: female brown-headed cowbirds prefer less intense male displays. *PLoS One, 7*(5), 1–8. http://dx.doi.org/10.1371/journal.pone.0036130.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Development Core Team, R. (2011). *nlme: Linear and nonlinear mixed effects models*. Retrieved from http://cran.r-project.org/package=nlme.

R Development Core Team, R. (2011). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/.

Rauceo, S., Harding, C. F., Maldonado, A., Gaysinkaya, L., Tulloch, I., & Rodriguez, E. (2008). Dopaminergic modulation of reproductive behavior and activity in male zebra finches. *Behavioural Brain Research, 187*(1), 133–139. http://dx.doi.org/10.1016/j.bbr.2007.09.003.

Riebel, K. (2009). Song and female mate choice in zebra finches: a review. *Advances in the Study of Behavior, 40*, 197–238. http://dx.doi.org/10.1016/S0065-3454(09)40006-8.

Ruploh, T., Bischof, H.-J., & Engelhardt, N. (2012). Adolescent social environment shapes sexual and aggressive behaviour of adult male zebra finches (*Taeniopygia guttata*). *Behavioral Ecology and Sociobiology, 67*(2), 175–184. http://dx.doi.org/10.1007/s00265-012-1436-y.

Schlinger, B. A., Barske, J., Day, L., Fusani, L., & Fuxjager, M. J. (2013). Hormones and the neuromuscular control of courtship in the golden-collared manakin (*Manacus vitellinus*). *Frontiers in Neuroendocrinology, 34*(3), 143–156. http://dx.doi.org/10.1016/j.yfrne.2013.04.001.

Schmidt, M. F., & Wild, J. M. (2014). The respiratory-vocal system of songbirds: anatomy, physiology, and neural control. *Progress in Brain Research, 212*, 297–335. http://dx.doi.org/10.1016/B978-0-444-63488-7.00015-X.

Slater, P. J. B., Eales, L. A., & Clayton, N. S. (1988). Song learning in zebra finches progress and prospects. In *Advances in the study of behavior* (Vol. 18, pp. 1–34). London, U.K.: Academic Press.

Soma, M., & Garamszegi, L. Z. (2015). Evolution of courtship display in Estrildid finches: dance in relation to female song and plumage ornamentation. *Frontiers in Ecology and Evolution, 3*(February), 1–11. http://dx.doi.org/10.3389/fevo.2015.00004.

Soma, M., & Mori, C. (2015). The songbird as a percussionist: Syntactic Rules for non-vocal sound and song production in java sparrows. *PLoS One, 10*(5), e0124876. http://dx.doi.org/10.1371/journal.pone.0124876.

Sossinka, R., & Böhner, J. (1980). Song types in the zebra finch *Poephila guttata castanotis*. *Zeitschrift für Tierpsychologie, 53*(2), 123–132.

Suthers, R. A., Goller, F., & Pytte, C. (1999). The neuromuscular control of birdsong. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 354*(1385), 927–939. http://dx.doi.org/10.1098/rstb.1999.0444.

Taylor, R. C., Klein, B. A., Stein, J., & Ryan, M. J. (2011). Multimodal signal variation in space and time: how important is matching a signal with its signaler? *The Journal of Experimental Biology, 214*(Pt 5), 815–820. http://dx.doi.org/10.1242/jeb.043638.

Willems, R. M., Ozyürek, A., & Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cerebral Cortex (New York, N.Y.), 17*(10), 2322–2333. http://dx.doi.org/10.1093/cercor/bhl141.

Williams, H. (2001). Choreography of song, dance and beak movements in the zebra finch (Taeniopygia guttata). *The Journal of Experimental Biology, 204*(20), 3497–3506.

Woolley, S. C., & Doupe, A. J. (2008). Social context–induced song variation affects female behavior and gene expression. *PLoS Biology, 6*(3), e62. http://dx.doi.org/10.1371/journal.pbio.0060062.

Zann, R. A. (1996). *The zebra finch: A synthesis of field and laboratory studies*. Oxford, U.K.: Oxford University Press.

Zann, R. A., & Cash, E. (2007). Developmental stress impairs song complexity but not learning accuracy in non-domesticated zebra finches (*Taeniopygia guttata*). *Behavioral Ecology and Sociobiology, 62*(3), 391–400. http://dx.doi.org/10.1007/s00265-007-0467-2.

# APPENDIX

## *Methods*

**Table A1**

Camera and recording specifications for experiment A, B, C

| Experiment | Camera | Recording software | Microphone |
|---|---|---|---|
| A<br>Indoor<br>2010 | Panasonic WV-CP500/G; 30 fps; 720×576 pixel resolution; perspective: top view | Noldus Recorder | Sennheiser ME80/K3U+preamplifier vivanco MA222 |
| B<br>Outdoor<br>2011 | Canon LEGRIA FS200 Camcorder; 25 fps; 720×576 pixel resolution; perspective: side view | | Sennheiser ME80/K3U |
| C<br>Indoor<br>2013 | Microsoft LifeCam VX-700 (USB); 30 fps; 640×480 pixel resolution; perspective: top view.<br>Imaging Source FireWire CCD Color DFK 21AF04; 60 fps; 640 × 480 pixel resolution; perspective: side view. | Noldus Media Recorder 2 | Sennheiser ME80/K3U+Conrad USB sound card 7.1 |

### *Behaviour definitions, audio and video analysis*

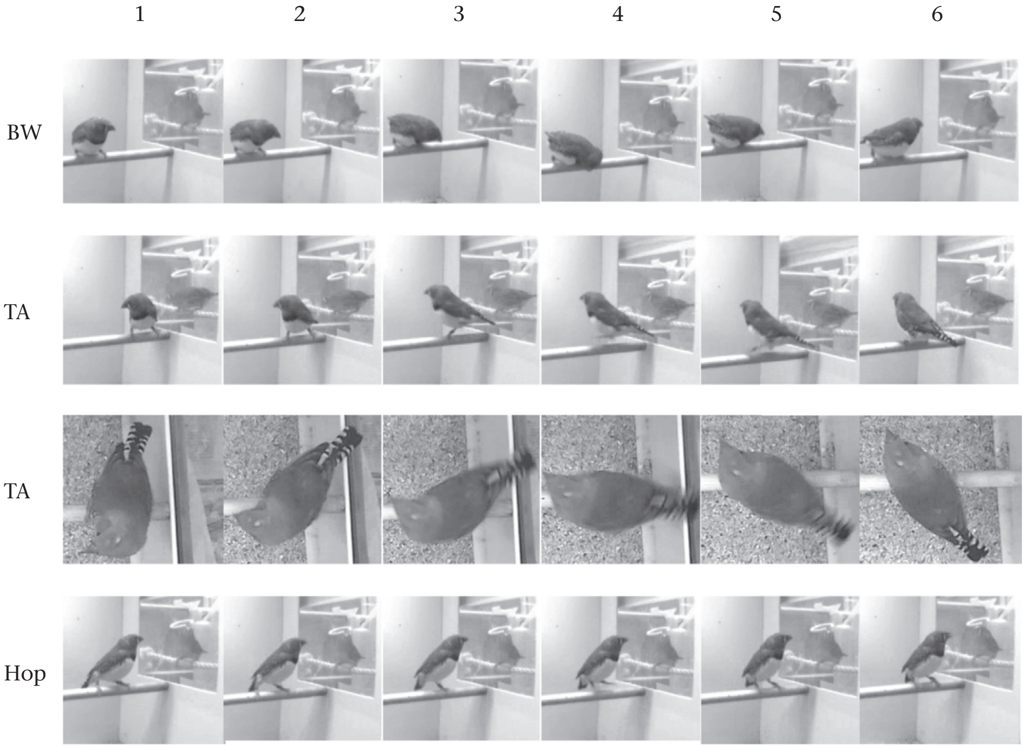


**Figure A1.** Still frames from videotaped material (30 fps) of male 3641 illustrating key moments of the three song-associated body movements: BW, TA (filmed from the front and above) and hops. The female is visible behind the window in the far right. The top row illustrates six of 19 positions in a BW sequence that lasted 570 ms in total. Frame 1: inception of the movement; frames 2 and 3: lowering of the head; frame 4: 'stroke', which depicts the lowest head position; frame 5: head lifting; 6: completion of poststroke phase. The time stamps of frames 1 and 6 were used to mark the beginning and end of the BW. The time stamp of frame 4 served to mark the BW itself as a point event. The second and third rows illustrate six of 21 frames in a TA sequence lasting 630 ms in total. Frame 1: inception of the movement; frames 2 to 5: the body axis is turning towards the right; frame 4 depicts the stroke of a TA, when the tail crosses the perch and both feet are displaced from the perch; frame 6 depicts the completion of the TA. The time stamps of frames 1 and 6 were used to mark the beginning and end of the TA. The stroke was used to mark the TA as a point event. The stroke of a TA was also scored as a hop. The bottom row shows six of 10 frames in a hop sequence, covering a total of 300 ms. Only the stroke of the sequence, when both feet were off the perch, was scored for hops (frame 4).

For one experiment ($N$=10) we included time stamps for the 2 s time window before the onset of the introductory note, for the introductory note itself and for the onset and end of the first motif note. However, when analysing the frequency of the movement between the single segments, we could not find any significant difference. We therefore decided to merge the mentioned segments to an overall segment called 'start'. We did the same for both of the other categories and ended up with three categories: 'start', 'centre', 'end'.

### *Hopping and calling associated by chance?*

We focused only on calls and hops that occurred before and after songs, not during songs (i.e. excluding song itself and 2 s before and after). To do so we measured for each bird the average call duration of 10 calls per 5 min video produced during 'nonsong', in four to 10 videos, resulting in 40–100 calls per bird. The average call length was determined empirically to be 75 ms (SD ± 26 ms). Hop duration was determined by measuring three randomly chosen hops from one bird using a series of video still images which yielded a duration of 300–333 ms. To avoid overestimation of hops and calls coinciding, we shortened the duration of hopping to 225 ms to receive more conservative results for both calculations: the real observation and the following Monte Carlo simulation. We thus counted hops and calls as 'overlapping' when they occurred within a 300 ms time window (225 ms estimated hop duration + 75 ms call duration). For each bird we used the actual number of hops and calls and distributed them randomly over time for 1000 times. Note that on one side every turn-around included a hop, but for the specific analyis of hopping movements we excluded these. We counted the number of times hopping and calling coincided as a result of the random distribution. A mean value from all 1000 randomizations was used to compare these results to the observed values.

### *Calculation of the expected transition ratios during song*

To determine whether movement sequences differed from chance, we analysed the transition frequencies between movements and then calculated the transition ratios that would be expected if the movement sequences were in random order (given the observed total numbers of movements and observed transition strings).

Because we only considered transition probabilities between two events (movements or start/end of a string) independent of other preceding events, the probability of a transition between these two events depends only on the relative frequency of the second event. Thus (1) the expected probabilities for the transitions in our sequence analysis from start to TA (or to BW, or to hop) were calculated by dividing the total number of TA (or BW or hops, respectively) by the sum of all three movements, yielding the relative frequency of the respective movement (Table A3, column 1). (2) To calculate the expected probability that a TA (or BW or hop) ends a transition string, we had to take string length into account, since longer strings result in fewer endings. We thus obtained the probability of any movement ending a transition string by dividing one by the average observed string length (1/4.33 = 0.23). Put differently, this value is the proportion of transitions from one movement to the end of a string in relation to all transitions (both from one movement to another movement and from a movement to an end; Table A3, fourth row). (3) All movements not terminating a string have a probability of transitioning to a TA (or BW, or hop) equal to the relative frequency of the respective movement (see 1). Therefore we calculated the expected probabilities of the transitions from one movement to another as the product of the relative frequency of the second movement and 1 minus the probability of the first movement ending a string (1–0.23 = 0.77; Table A3, columns 2–4).

### *Statistical Analysis: data transformation*

We chose the cube root to transform non-normal data because applying the square root did not lead to a normal distribution in some cases and log transform was not suitable due to the presence of several zero values in some data sets.

## *Results*

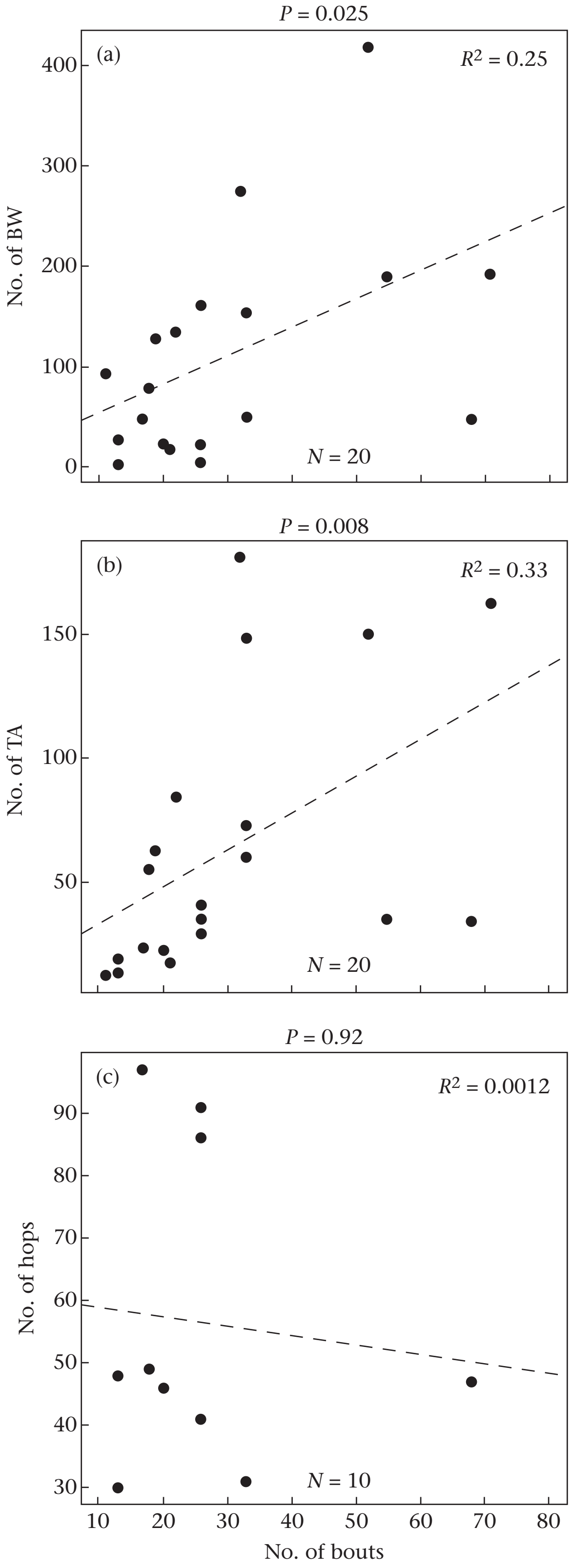


**Figure A2.** Number of (a) BW, (b) TA and (c) hops in relation to number of bouts sung. See text for statistics.

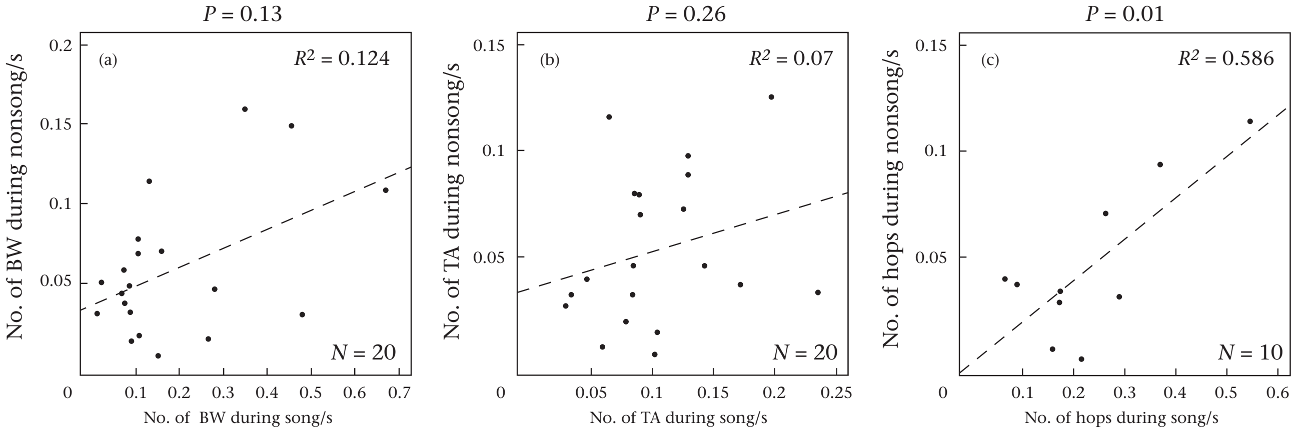


**Figure A3.** Correlations of body movements between song and nonsong: (a) BW, (b) TA and (c) hops. Pearson product moment correlation: BW song versus nonsong: $t_{18} = 1.6$, $P < 0.13$; TA song versus nonsong: $t_{18} = 1.16$, $P = 0.26$; hop song versus nonsong: $t_8 = 3.36$, $P < 0.001$.

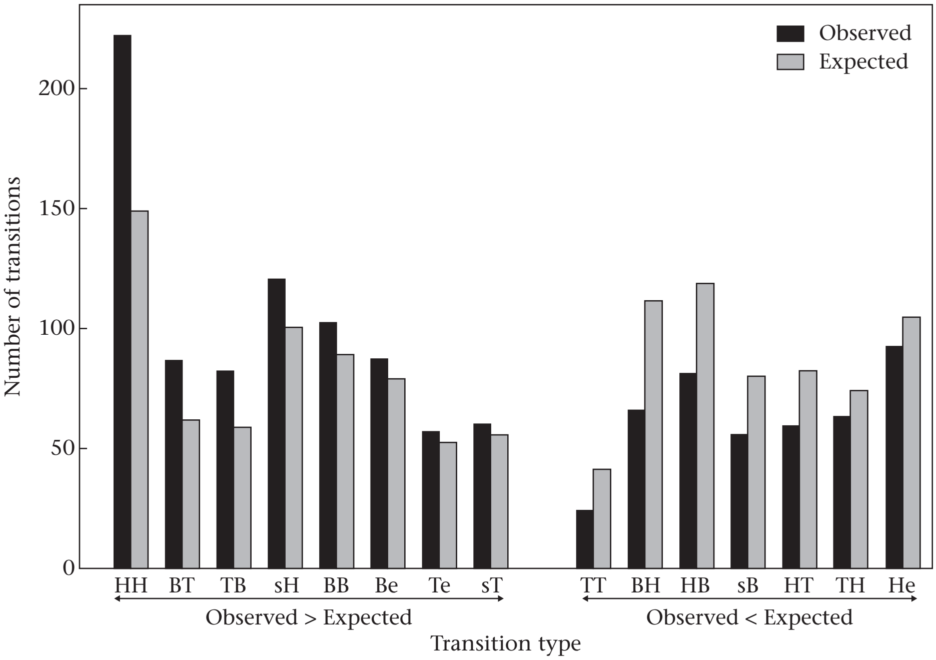


**Figure A4.** Differences between observed and expected transition frequencies during song. Observed: absolute number of transitions observed during song of 10 birds (black); expected: number expected if the three movements were equally likely to follow each other, given the observed number of all movements and transition strings (grey). The bars are ordered by relative difference between observed and expected number, and grouped as O > E (left) and O < E (right). B: beak wipe; H: hop; T: turn-around; s: start of a transition string; e: end of a transition string.

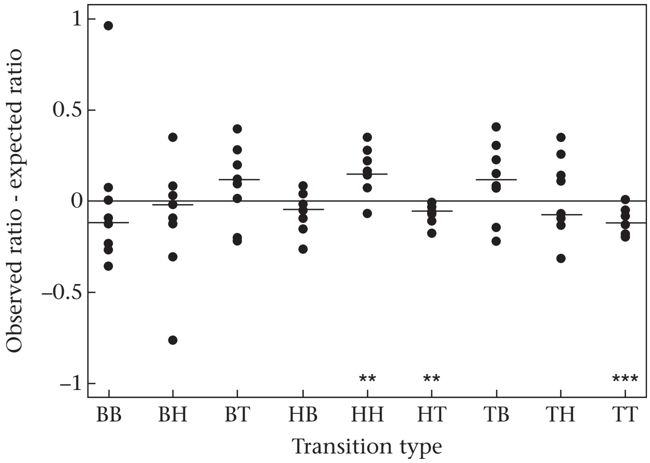


**Figure A5.** Expected transition ratios subtracted from observed transition ratios during song. Transition probabilities expected if the three movements were equally likely to follow each other, subtracted from the observed ratios of 10 birds (black dots) and medians (horizontal lines) for the nine transition types. B: beak wipe; H: hop; T: turn-around. Asterisks mark the transition types where observed ratios differed significantly from expected ones: $^{**}P < 0.01$; $^{***}P < 0.001$. See text for statistics.

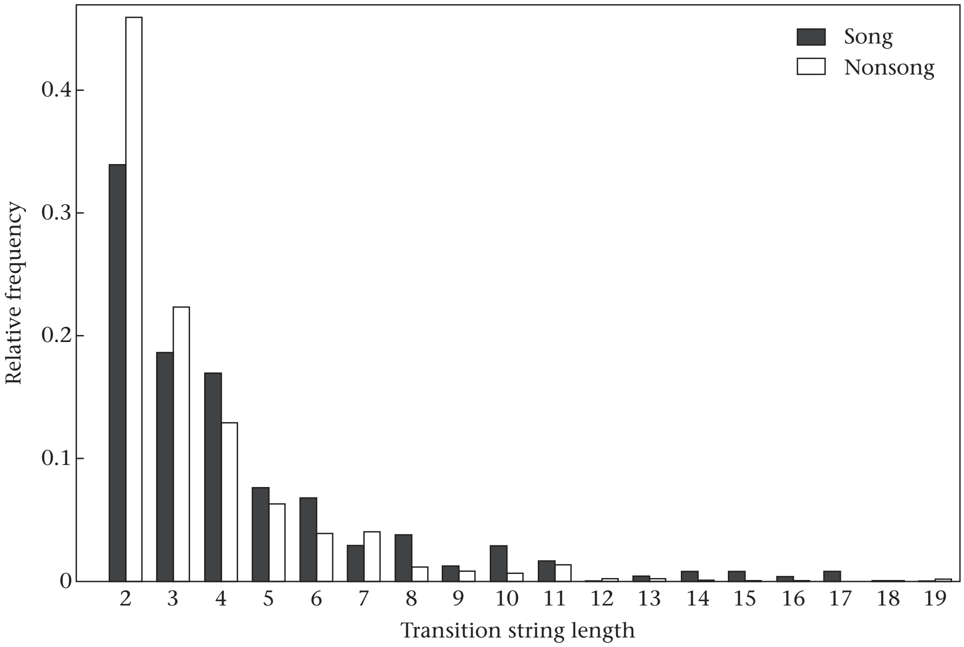


**Figure A6.** Transition string length during song (dark bars) and nonsong (light bars). Values in the histogram are given as frequencies relative to the total number of transitions during song and nonsong, respectively. A transition length of 2 represents a transition between two movements, 3 = three movements, etc.

**Table A2**
List of the most frequently observed movement sequences

| Transition string | Frequency of occurrence as a string or as part of a longer string | Frequency of occurrence as a string | String length | Format indicating incorporated movements |
|---|---|---|---|---|
| BTB | 42 | 4 | 3 | BW & TA |
| **HHB** | 42 | 0 | 3 | *Hop & TA* |
| **HBH** | 26 | 0 | 3 | **BW & hop** |
| *THH* | 24 | 4 | 3 | All three |
| TBT | 22 | 2 | 3 | |
| TBB | 22 | 1 | 3 | |
| *HHT* | 21 | 3 | 3 | |
| **BHH** | 20 | 0 | 3 | |
| **HHHB** | 20 | 2 | 4 | |
| **HHBH** | 15 | 0 | 4 | |
| BTBB | 14 | 1 | 4 | |
| *HHHT* | 13 | 2 | 4 | |
| TBBB | 12 | 2 | 4 | |
| *THHH* | 12 | 2 | 4 | |
| BTBT | 12 | 1 | 4 | |
| **HHBB** | 11 | 1 | 4 | |
| *HHHHT* | 10 | 3 | 5 | |
| **HHHHB** | 8 | 0 | 5 | |
| **HHHBB** | 7 | 1 | 5 | |
| TBBBB | 7 | 0 | 5 | |
| *THHHH* | 6 | 2 | 5 | |
| *HTHHH* | 6 | 1 | 5 | |
| BTBBB | 6 | 0 | 5 | |
| BBTBB | 6 | 0 | 5 | |
| TBBBBB | 6 | 1 | 6 | |
| BTBBBB | 5 | 1 | 6 | |
| *HHHHTH* | 5 | 1 | 6 | |
| **HHHHHB** | 4 | 1 | 6 | |
| *HHHHHT* | 4 | 1 | 6 | |
| **HHHBHB** | 4 | 0 | 6 | |
| HHBBTB | 4 | 0 | 6 | |
| **HBHHBH** | 4 | 0 | 6 | |
| BTBBBBB | 4 | 0 | 7 | |
| **HBHHBHH** | 4 | 0 | 7 | |
| *HHHTHHH* | 4 | 0 | 7 | |
| **HHHHBHB** | 3 | 1 | 7 | |
| TBBBBBB | 3 | 0 | 7 | |
| BBTBBBB | 3 | 0 | 7 | |
| HHBBTBB | 3 | 0 | 7 | |
| *HTHHHHH* | 3 | 0 | 7 | |

Strings are ordered by their frequency of appearing either as a whole transition string or as part of a longer transition string during song of 10 birds. Included in this table are the eight most frequent strings that contain at least two types of movements, for string lengths of three to seven movements each. Roman/bold/italics/underlining indicate which of the movements are included in each string. B: beak wipe; H: hop; T: turn-around.

**Table A3**
Calculation of expected transition probabilities

| Probability of transition | From start | From TA | From BW | From hop |
|---|---|---|---|---|
| To TA | nTA/nALL | 0.77×(nTA/nALL) | 0.77×(nTA/nALL) | 0.77×(nTA/nALL) |
| To BW | nBW/nALL | 0.77×(nBW/nALL) | 0.77×(nBW/nALL) | 0.77×(nBW/nALL) |
| To hop | nHop/nALL | 0.77×(nHop/nALL) | 0.77×(nHop/nALL) | 0.77×(nHop/nALL) |
| To end | – | 0.23 | 0.23 | 0.23 |
| Sum | 1 | 1 | 1 | 1 |

nTA, nBW and nHop are the total number of turn-arounds, beak wipes and hops as part of transitions, respectively. nALL = nTA + nBW + nHop. Note that the probabilities for all transitions starting at a specific event sum to 1.