Article

Wild nightingales flexibly match whistle pitch in real time

Graphical abstract

Highlights

- Nightingales in the wild perform whistle song-type matching

- Nightingales flexibly match the whistle pitch over a wide range of frequencies

- Pitch matching is more precise for fast whistle responses

- This behavior is conserved across seasons suggesting a stable neural circuit
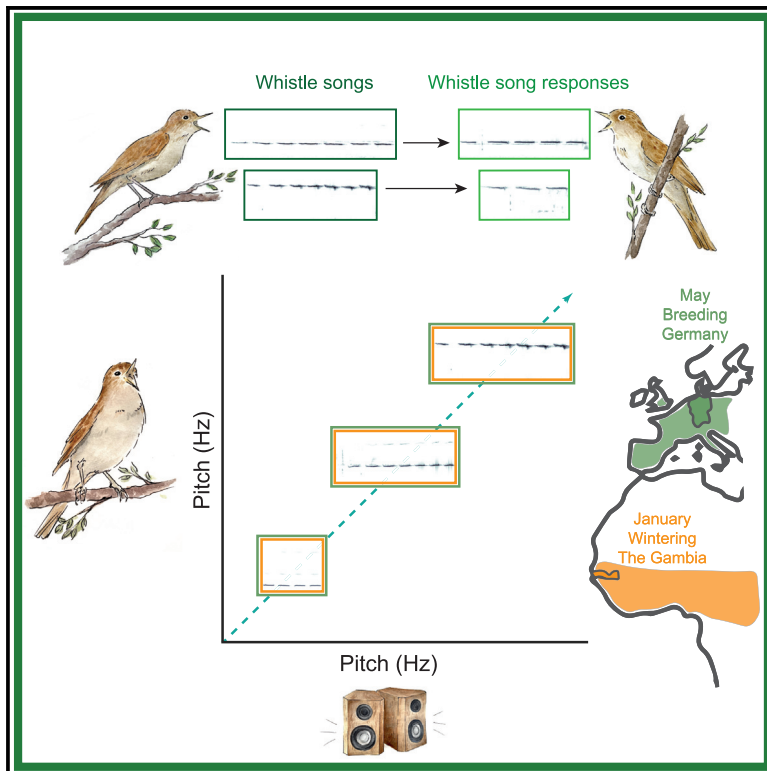
Authors

Giacomo Costalunga,
Carolina Sánchez Carpena,
Susanne Seltmann,
Jonathan I. Benichov, Daniela Vallentin

Correspondence

daniela.vallentin@bi.mpg.de

In brief

Costalunga et al. find that nightingales match the pitch of conspecific whistle songs and synthetic whistle song playbacks across a wide continuum of frequencies. Pitch matching is more precise when they produce immediate responses and is conserved year-round, independent of seasonal changes in physiology that affect other aspects of singing behavior.

CellPress

Article

Wild nightingales flexibly match whistle pitch in real time

Giacomo Costalunga,[1] Carolina Sánchez Carpena,[1] Susanne Seltmann,[1] Jonathan I. Benichov,[1] and Daniela Vallentin[1,2,3,*]

[1]Neural Circuits for Vocal Communication Research Group, Max Planck Institute for Biological Intelligence, Eberhard-Gwinner-Str., Seewiesen 82319, Germany
[2]Twitter: @LabVallentin
[3]Lead contact
*Correspondence: daniela.vallentin@bi.mpg.de
https://doi.org/10.1016/j.cub.2023.06.044

SUMMARY

Interactive vocal communication, similar to a human conversation, requires flexible and real-time changes to vocal output in relation to preceding auditory stimuli. These vocal adjustments are essential to ensuring both the suitable timing and content of the interaction. Precise timing of dyadic vocal exchanges has been investigated in a variety of species, including humans. In contrast, the ability of non-human animals to accurately adjust specific spectral features of vocalization extemporaneously in response to incoming auditory information is less well studied. One spectral feature of acoustic signals is the fundamental frequency, which we perceive as pitch. Many animal species can discriminate between sound frequencies, but real-time detection and reproduction of an arbitrary pitch have only been observed in humans. Here, we show that nightingales in the wild can match the pitch of whistle songs while singing in response to conspecifics or pitch-controlled whistle playbacks. Nightingales matched whistles across their entire pitch production range indicating that they can flexibly tune their vocal output along a wide continuum. Prompt whistle pitch matches were more precise than delayed ones, suggesting the direct mapping of auditory information onto a motor command to achieve online vocal replication of a heard pitch. Although nightingales' songs follow annual cycles of crystallization and deterioration depending on breeding status, the observed pitch-matching behavior is present year-round, suggesting a stable neural circuit independent of seasonal changes in physiology. Our findings represent the first case of non-human instantaneous vocal imitation of pitch, highlighting a promising model for understanding sensorimotor transformation within an interactive context.

INTRODUCTION

Real-time adjustments to vocal output based on auditory input are crucial for interactive vocal communication,[1,2] as seen in humans and various other species.[2-8] Although animals can discriminate sounds based on frequency distribution,[9-11] the ability to dynamically adjust specific spectral features, such as pitch, in response to incoming auditory information has primarily been observed in humans.[12] Some songbirds are known to perform song-type matching, that is, listening to and repeating the songs of conspecifics if those songs are part of their own existing repertoire[13] as an aggressive signal to address conspecifics.[14,15] Among them, the common nightingale (*Luscinia megarhynchos*) has been praised for centuries for its extraordinary singing behavior[16-18] (Figure S1A; Video S1). These birds have a seasonal repertoire[19,20] of 150–200 different songs[21,22] of which one acoustically distinct type is called "whistle songs." These songs contain relatively unmodulated tonal "whistle syllables" (Figures 1A, 1B, and S2A; and Audio S1) and are used in long-range counter-singing duels for territorial defense and mate attraction.[23-25] To quantitatively investigate whether nightingales perform real-time identification and reproduction of the

pitch of heard whistle syllables, we performed song recordings of pairs of naturally interacting male nightingales during their mating season in Germany and playback experiments both during their mating season in Germany and in nightingales' wintering quarters in The Gambia.

RESULTS

Nightingales in the wild perform song-type matching of whistle songs during counter-singing duels

We recorded 10 interacting pairs of male nightingales singing during their mating season between April and May in Brandenburg, Germany (Figures 1A, S1B, and S1C). Nightingales sang frequency-unmodulated whistle songs (14.71% ± 3.98% of all songs, 776 out of 5,261 songs from 20 birds) (Figures 1A, 1B, and S2A–S2G). Individual whistle songs were composed of 1–33 whistle syllables (Figures 1B and S2C) of similar durations and pitch[23] (Figures 1B, 1C, and S2A–S2G). The pitch of these vocalizations spanned from 1,185 to 9,169 Hz (5,190 whistle syllables from 776 whistle songs) (Figure 1D). Due to the non-normal distribution of whistle pitches (one-sample Kolmogorov-Smirnov test for normal distribution, p < 0.001),
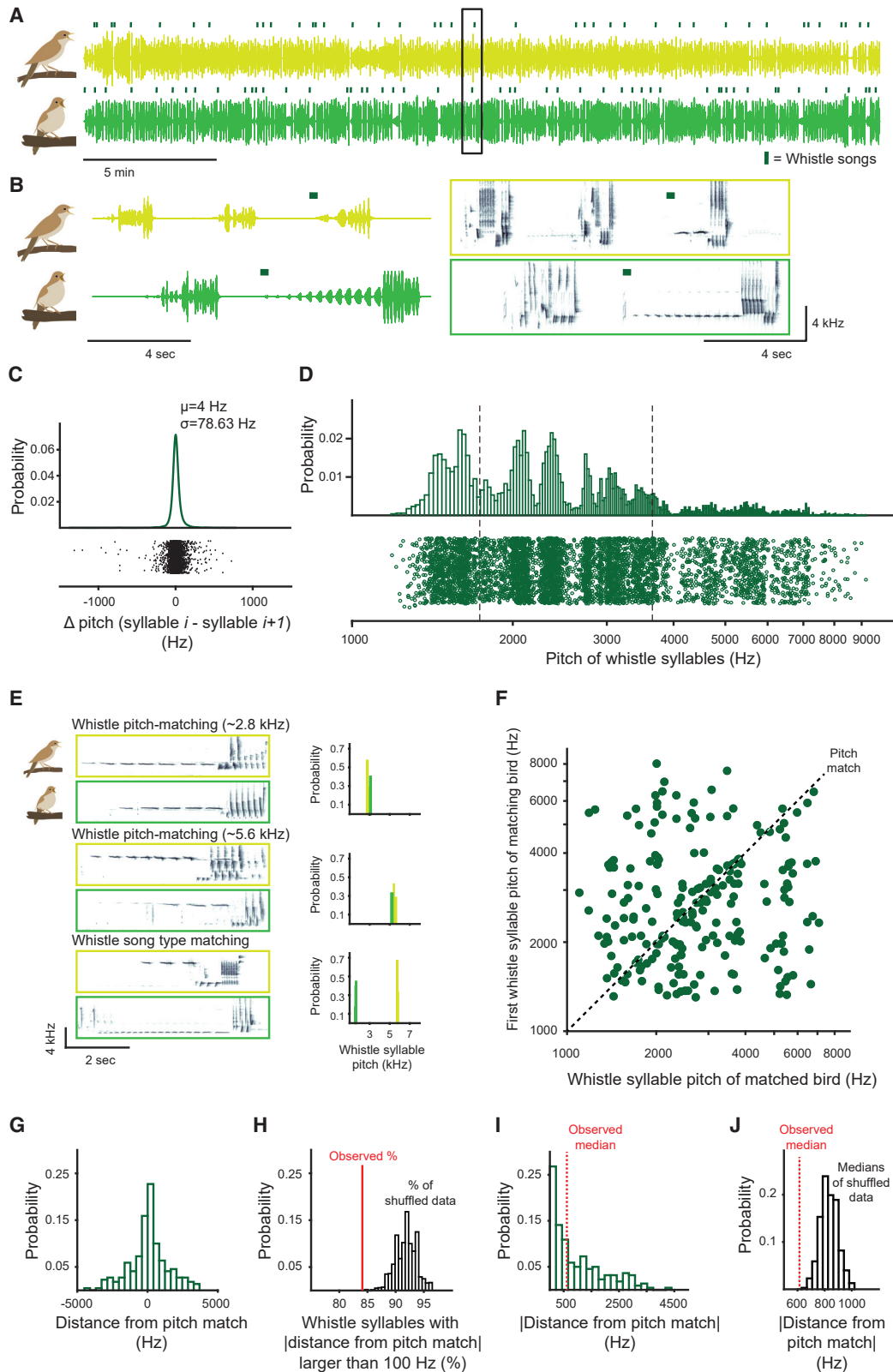
Figure 1. Nightingales in the wild perform song-type matching of whistle songs during counter-singing duels

(A) Two simultaneously recorded oscillograms from an example counter-singing duel between two nightingales in the wild. The onsets of whistle songs are indicated with dark green tick marks. The black box highlights a vocal interaction involving whistle songs.

(legend continued on next page)

we decided to perform k-means clustering to define 10 distinct clusters of whistle syllables (see STAR Methods; Figure S2E). The centroids of these clusters followed a logarithmic periodicity, with the centroids of higher pitch whistle clusters being further apart from each other compared with those of lower pitches (Figures S2F and S2G). Each bird sang whistle songs belonging to at least 8 clusters (9.19 ± 0.51 clusters, Figure S2H). In humans, the perception of pitch is logarithmic,[26,27] following Weber-Fechner's law.[28,29] The observed log-normal distribution (Lilliefors test for normal distribution, p = 0.09) of produced whistle pitches complements this perceptual phenomenon. Within this distribution, 60% of all whistle syllables were in the pitch range between 1,735 and 3,652 Hz, the 20th and 80th percentile, respectively, which we defined as the nightingales' "central pitch range" (Figures 1D and S2F).

We observed that counter-singing nightingales perform song-type matching of whistle songs in about 30% of the cases in which they produce whistle songs (mean = 29.55% ± 10.57% of whistle songs used for song-type matching, Figures 1E and S2I). To determine the rate at which this behavior would occur by chance, we simulated countersinging based on the observed whistle song rates of nightingales. In the simulated datasets, whistle song-type matching occurred in only ~5% of the cases (mean simulated events = 5.10% ± 3.85% (see STAR Methods), median observed events = 28.89%, median simulated events = 5%, Wilcoxon rank sum test, p < 0.001), suggesting that nightingales specifically target their whistle songs in response to their neighbor's whistle songs. Given that whistle songs can vary in the number of whistle syllables (Figure S2C), we tested whether nightingales matched the number of whistle syllables of their opponent's whistle song. We found no significant correlation between the number of whistle syllables in the targeted whistle song and the one produced in response (Figure S2J). Additionally, nightingales did not match the total duration of whistle song syllables when responding with whistle songs (Figure S2K). Overall, the birds did not copy the song syllable sequence but rather sang different whistle songs in response to target whistle songs (93.31% ± 6.87% whistle song responses were different from target whistle songs, Figure 1E). Next, we analyzed the pitch of whistle syllables sung in response to a whistle song (20 birds, 220 whistle syllables, Audio S1, Figure 1F). We plotted the pitch of the first whistle syllable of each whistle song, against the pitch of the first whistle from the preceding rival's song and

calculated the shortest Euclidean distance between each point and a hypothetical perfect match in pitch (i.e., the identity line, "distance from pitch match") (Figure 1G). From this Euclidean distance distribution, we determined the percentage of accurate responses, defined as responses with distances from pitch match that were less than 100 Hz. The observed percentage of inaccurate responses was smaller than the expected percentages obtained from 1,000 permutations of randomly reassigned responses (Figure 1H; see STAR Methods). Similarly, the median of the observed distribution of distances from pitch match was less than the medians of the distributions of expected distances derived from the shuffled data (Figures 1I and 1J). These results suggest that naturally interacting nightingales are able to regulate the pitch of their whistle songs to match one another, corroborating previous reports of nightingales interacting with playbacks of conspecific whistle songs.[25]

Previously, it has also been reported that nightingales repeat song sequences in a periodic manner.[30] Thus, we asked if the observed rates of inaccurate responses might in part be due to individual history-dependent patterns in song production. To test this possibility, we recorded uninterrupted singing from 11 individual birds and found that whistles of the same pitch did not reoccur in the time interval required for each bird to sing 11 whistle songs (duration for the production of 11 whistle songs: 4.48–12.43 min, Figure S2O). This suggests that the potential underlying sequence of song production[30] does not contribute to pitch-matching failures. Alternatively, the variability in responses could be due to a combination of the birds' preference for singing whistle songs in the central pitch range or the birds selectively engaging in whistle pitch matching. Although indicative of a capacity for pitch matching, these observational results from singing pairs do not elucidate whether nightingales are able to perceive and flexibly reproduce a given pitch in real time.

Nightingales in the wild perform flexible whistle pitch matching over a wide range of frequencies

To test the ability of nightingales to match arbitrary pitches covering the entire range of whistle pitch production, we designed a controlled playback experiment with a high-resolution sampling of pitches. We exposed 12 wild birds to a battery of synthetic whistle songs covering the natural range of their sung whistle frequencies and recorded their responses (Figures 1D and 2A–2C). During the night, individual nightingales sing from

---

(B) Zoomed-in view of black box in (A). Top left: nightingale 1 singing three different songs; bottom left: nightingale 2 singing two different songs. Onsets of whistle songs are highlighted with dark green tick marks. See also Audio S1. Right: spectrograms of the examples shown on the left. Nightingale 2 singing a whistle song consisting of 11 whistle syllables and nightingale 1 performs song-type matching by responding with a whistle song consisting of 2 whistle syllables.

(C) Difference in pitch between whistle syllables within whistle songs (median = 0.63 Hz, Wilcoxon rank sum test, p = 0.956).

(D) Distribution of pitches of whistle syllables (2,912 ± 1,430 Hz with a median of 2,413 Hz) on a logarithmic scale. Top: histogram of whistle syllable pitch; bottom: pitches of individual whistle syllables. Dashed lines represent the central pitch range of the pitch distribution.

(E) Spectrograms of 3 examples of whistle song replies to whistle songs. In the first and second examples, the pitch of the whistle song was matched; in the third example, the whistle song response had a different pitch. See also Audio S1. Right: histograms of pitch distribution of the corresponding whistle syllables shown on the left.

(F) Pitch of matched bird's and matching bird's first whistle syllables on a log scale. Diagonal line indicates the hypothetical perfect pitch match.

(G) Probability distribution of the Euclidean distance of each whistle pitch response to the hypothetical perfect pitch match.

(H) Distribution percentages of inaccurate responses, calculated from 1,000 permutations of shuffled data. Observed % of inaccurate responses is indicated with the red line (84.09%).

(I) Probability distribution of distance from pitch match. Median is indicated with the red dashed line (614.43 Hz).

(J) Distribution of median distances from pitch match for shuffled data. Median of observed data is marked with red dashed line. See also Figures S1, S2, and S4, Table S1, Video S1, and Audio S1.
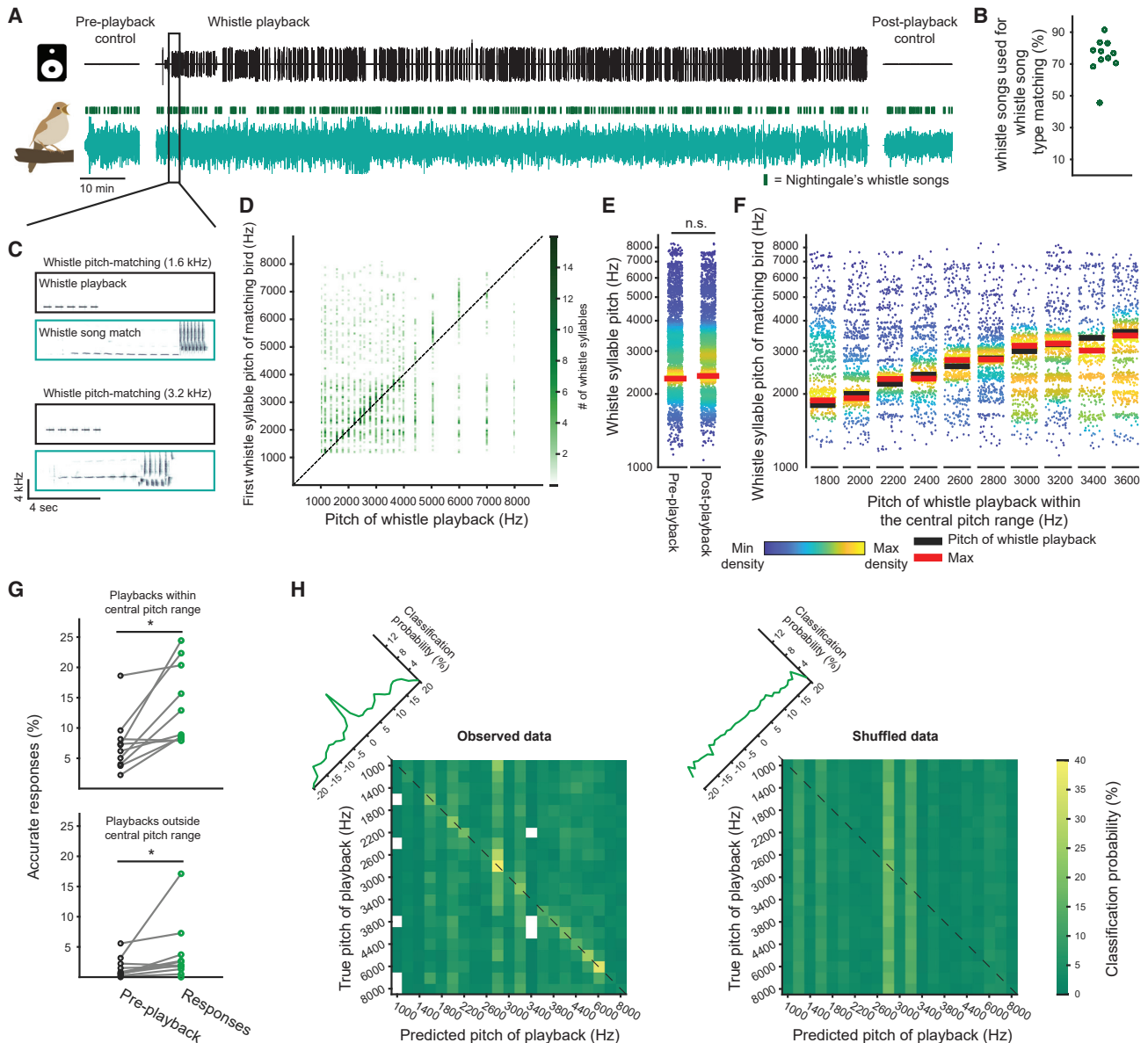
Figure 2. Nightingales in the wild perform flexible whistle pitch matching over a wide range of frequencies

(A) Oscillogram of an example whistle playback experiment with one nightingale. The playback experiment consisted of a pre-playback control period, whistle playback stimulation, and a post-playback control period (in black). During these phases, the nightingale's songs were recorded simultaneously (in green). Whistle songs of the bird are highlighted with dark green tick marks. Black box highlights examples of whistle pitch match.

(B) Percentage of whistle song used for whistle song replies to whistle playbacks in Germany.

(C) Spectrograms of two examples of whistle pitch matching to synthetic whistle playbacks of different pitch. See also Audio S2.

(D) Correlation between pitch of whistle playbacks and matching whistle syllables, color-coded by quantity (darker colors indicate more whistle syllables in a given pixel). Dashed line indicates the line of exact pitch matching.

(E) Pitch of whistle syllables sung during the pre-playback control and post-playback control phases. Note logarithmic scale of the y axes. The horizontal red bars represent the pitch at the global maximum (Max) of the whistle pitch distribution function for each playback pitch.

(F) Density scatterplot of pitches of whistle syllables sung in response to whistle playbacks within the central pitch range. The horizontal black and red bars represent the pitch of the whistle playback and the Max of the responses to the given playback, respectively.

(G) Top: percent of responses that were categorized as accurate expected by chance based on pre-playback control and observed during the playback experiments for playbacks with pitches within the central range (median accurate responses for pre-playback control = 6.69%, median accurate responses to playbacks = 10.9%, Wilcoxon signed rank test, p = 0.004). Bottom: as above, for playbacks outside the central range (median accurate responses for pre-playback control = 0.66%, median accurate responses to playbacks = 1.97%, Wilcoxon signed rank test, p = 0.006).

(H) Left: decoding accuracy based on the k-nearest neighbor classifier. Rows indicate the true pitch of playbacks that the nightingales were presented with, and columns indicate the predicted pitch. Tuning function at the diagonal represents average classification probability across all pitches. Right: as left, but for shuffled data. See also Figures S1, S3, and S4, Table S1, and Audio S2.

stable posts within their territory,[21] allowing the presentation of playback stimuli to the same bird for an entire experimental session (Figures 2A and S1D). We presented each nightingale with a total of 400 artificial whistle song stimuli with pitches spanning from 1,000 to 8,000 Hz (see STAR Methods; Figures 2A and S3A). Birds engaged with the playbacks, singing whistle songs in response to the synthetic whistle stimuli (Figures 2A and 2C; Audio S2). We first assessed how the nightingales interacted with the playback by analyzing the first whistle syllable of each whistle song response and measuring the distance from pitch match (2,538 whistle syllables, 12 birds). By comparing the Euclidean distances obtained from randomly reassigning responses to playbacks with the observed ones, we found that 11 out of 12 birds regulated the pitch of their whistle songs in a manner that approximated the pitch of artificial playbacks (Figures 2D, S3A, and S3B). We further explored whether the observed differences in whistle pitch between songs recorded during the playback period and songs recorded in a control period prior to playbacks persisted after the stimulation. We found that the pitch distribution of whistles returned to baseline in the period immediately following the playback stimulation (0–5 min after last playback, Figure 2E), indicating that nightingales specifically and rapidly adjusted their whistling behavior depending on auditory input (pre-playback control = 4,099, post-playback control = 4,367 whistle syllables from 11 birds, median pitch pre-playback control = 2,876 Hz, median pitch post-playback control = 2,897 Hz, Wilcoxon rank sum test, $p = 0.722$).

Given the predominance of whistle syllables produced within the central pitch range, we separately analyzed all whistle syllables sung in response to playbacks within and outside this interval (15,953 whistle syllables from 2,311 songs from 11 birds, Figures 1D, 2F, 2G, and S3C). We applied a kernel density function to estimate the global maximum (Max, i.e., peak) of the distributions of pitches produced in response to the different playback pitches. We found that the values of these peaks did not differ between the central and the outside pitch range (Figure S3C), indicating that nightingales mainly responded with whistles within the central pitch range even when presented with playbacks outside of it (peaks inside the central pitch range = 2,682 ± 555 Hz, median = 2,765 Hz; peaks outside central pitch range = 2,331 ± 461 Hz, median = 2,259 Hz; Wilcoxon rank sum test, $p = 0.193$). Next, we calculated the difference between these peaks and the pitches of the playbacks (Δ pitch) (Figures 2F and S3C). We found that the distribution of Δ pitches inside but not outside the central pitch range centered around zero (Δ pitch inside the central pitch range = −19 ± 165 Hz, one-sample t test, $p = 0.731$; Δ pitch outside the central pitch range = −1,615 ± 2,336 Hz, one-sample t test, $p = 0.045$) (Figure S3D). To test whether the observed |Δ pitch| could be explained by the log-normal distribution of the pitch of whistle syllables, we compared them with |Δ pitch| extracted from randomly shuffled playback pitch associated with each whistle syllable response pitch (see STAR Methods; Figures S3E and S3F). We found that differences outside of the central pitch range could be explained by the nature of the pitch distribution (median observed |Δ pitch| outside = 1,929 Hz, median expected |Δ pitch| outside = 1,675 Hz, Wilcoxon signed rank test, $p = 0.7$, Figures S3E–S3G). However, the |Δ pitch| calculated for the

central range in the observed and expected data were different (median observed |Δ pitch| inside = 97 Hz, median expected |Δ pitch| outside = 437 Hz, Wilcoxon signed rank test, $p = 0.002$, Figures S3E–S3G). This finding indicated a higher accuracy in matching the pitch of whistle playbacks within the central pitch range. Although the nightingales were less accurate when stimulated outside of the central range, the proportion of whistle syllables produced within this range increased (median observed % whistle syllables with pitch outside of central range in response to playback outside the central range = 37.31%, median observed % whistle syllables with pitch outside of central range in response to playback inside the central range = 24,78%, Wilcoxon rank sum test, $p = 0.002$, median expected from shuffled data % whistle syllables with pitch outside of central range in response to playback outside the central range = 29.6%, Wilcoxon signed rank test, $p = 0.003$, Figure S3H). To further explore how reliable the pitch matching is in response to whistle playbacks, we calculated the percentages of accurate responses to each playback (% of whistle syllables in the range of ±100 Hz around the pitch of the playback). We found a higher percentage of accurate whistle syllables in response to whistle playbacks compared with whistle syllables produced in the corresponding pitch ranges during the pre-playback control period (median 7.9% accurate responses during playback, 4.17% ± 4.45%, median 3.17% for pre-playback control, Wilcoxon signed rank test, $p < 0.001$). The increase in accurate responses was present both for playbacks inside and outside the central range (Figure 2G), demonstrating the ability of nightingales to drive their whistle pitch beyond their natural tendencies to match the pitch of stimuli near the extremes of their vocal range. Furthermore, we tested the predictability of the playback stimulus pitch given the response of the bird. To do so, we trained a k-nearest neighbor classifier (see STAR Methods) and compared the results with responses randomly shuffled across playback trials. We found an overall accuracy of classification of 14.9% for observed data and 5.2% for shuffled data (4.8% chance level with 21 possible pitch classes). For fifteen playback pitches, the classifier was able to determine the playback pitch inside and outside the central pitch range, whereas in the shuffled data (Figures 2H and S3E), the playback pitches were predicted only for four pitches inside the central range. This finding suggests that it is possible to infer the pitch of the playback stimulus from the responses exhibited by the nightingales.

Pitch matching is more precise for immediate whistle responses

During conversational interactions, vocal production is preceded by perception of the partner's vocalization and planning of the response.[2] The resulting latency of the response can be critical for accurate perception and production in humans.[1] To test whether the response latency is also important for pitch-matching precision in counter-singing nightingales, we analyzed the latencies of whistle songs in response to playbacks (Figures 3A and 3B). First, we examined the response latency of the whistle song response. We found that nightingales' response latencies distributed around two clearly separated local maxima (Figures 3A–3E). For this reason, we divided these distributions into early and late responses for each bird (mean latency of the response peak of "early" responses for all birds = 1.94 ± 0.44 s; mean latency of
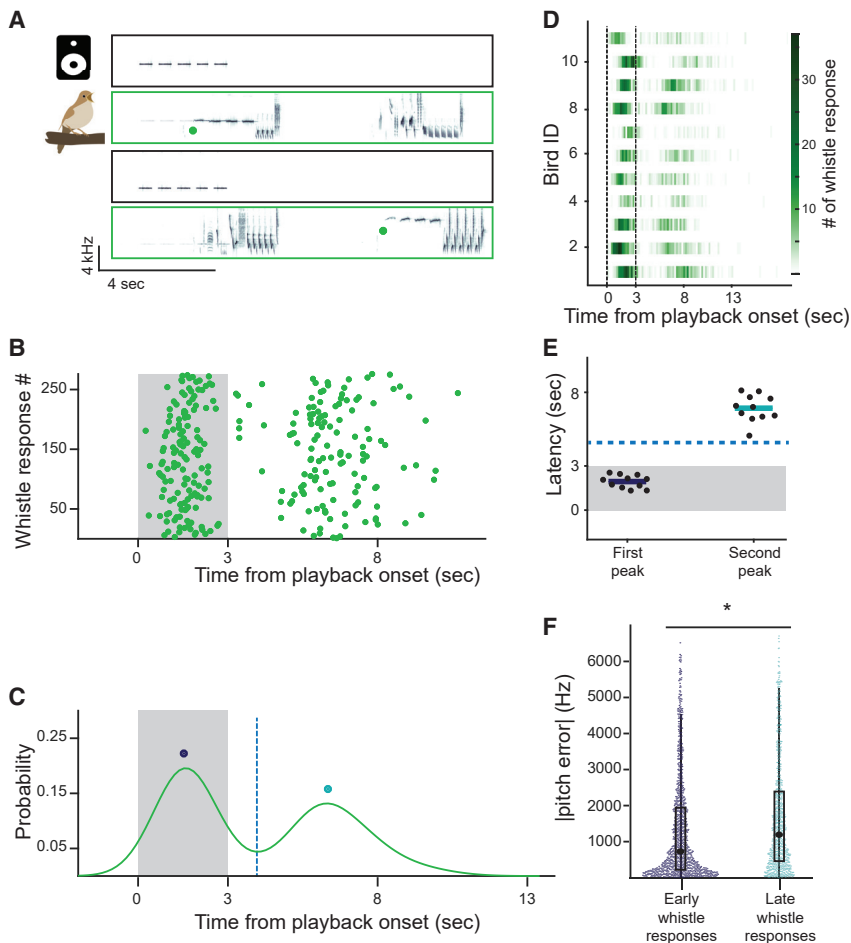
Figure 3. Pitch matching is more precise for early whistle responses

(A) Spectrograms for two examples of whistle song matching of synthetic whistle playbacks with different pitch and response latencies. Filled green dots represent the onset of the first whistle syllable of the singing bird.

(B) Onset of whistle responses from one bird, aligned to playback onset. Filled green circles represent the onset of the first whistle syllable (as in A); gray shaded area represents the duration of the whistle playback.

(C) Whistle response kernel probability distribution for the example bird in (B). Filled circles represent the first and second peaks of the distribution. Dashed blue line indicates the separation between the early and late response distribution (first trough).

(D) Heatmap of the onsets of whistle responses of 11 nightingales, color-coded by quantity (darker colors indicate more whistle syllables in a given pixel). Dashed lines represent the duration of the whistle playback.

(E) Latency of the first peak and second peak of the response probability distributions of 11 birds. Horizontal lines indicate the means. Dashed blue line indicates the mean latency of the first trough across birds.

(F) Distribution of the |pitch error|, from responses occurring before and after the trough of the response distribution of each bird (early and late responses, respectively). Black boxes represent the 25th and 75th percentiles, whiskers show the upper and lower adjacent values, and the filled black circles are the medians of the distributions (Wilcoxon rank sum test, p < 0.005). See also Figures S1 and S4 and Table S1.

the response peak of "late" responses for all birds = 6.92 ± 0.92 s, Figure S4A). The majority of the early responses (91%) had a latency shorter than the duration of the whistle playback (3 s), resulting in whistle songs that partially overlapped with the playback stimuli (latency from playback onset 1.96 ± 0.76 s, range: 0.1–4.88 s, Figure S4B). In contrast, late responses always occurred after playback offset (Figures 3A and 3E). Whether nightingales responded early or late was independent of the pitch of the target playback (Figure S4C). Then, we asked whether the difference in latency influenced the matching of the pitch of the whistle playbacks. We analyzed the absolute difference between the pitch of the playback and the pitch of the first whistle syllable (|pitch error|) and found a positive correlation with the response time (Figures S4D and S4E), indicating that more precise whistle pitch matches had shorter latencies. This was also reflected in smaller absolute error in pitch in early responses compared with the late ones (1,436 early responses, median |pitch error| = 727 Hz, 996 late responses, median |pitch error| = 1,193 Hz, Wilcoxon rank sum test, p < 0.005) (Figures 3A and 3F). When conducting the same analysis for naturally interacting pairs, we also found better pitch matching of whistles for early responses (Figures S4F–S4L). The higher accuracy of early responses might reflect a higher motivational or aggression state.[31]

Additionally, we asked whether the nightingales were able to match the target pitch of the playback from the start of the production of their first whistle syllable or whether they would adjust their pitch across successive whistle syllables in a reply to improve their matching performance over time. We analyzed the absolute difference in pitch between playback and the first or last syllable and found that the first syllable was on average closer to the target than the last (median: |first syllable(pitch) − playback(pitch)| = 892 Hz, median: |last syllable(pitch) − playback(pitch)| = 919 Hz, Wilcoxon signed rank test, p < 0.001). This finding indicated that nightingales perceive the pitch of the playback and match it directly without the need to "hunt" for the right note using auditory feedback.

Pitch-matching behavior is conserved outside the nightingale mating season
Nightingales' singing behavior follows annual cycles consisting of multiple phases related to season and breeding status.[19] The full songs are sung at night during the breeding season in Europe, from April to late June, followed by a period during which the birds stop singing.[19] In the wintering season in Africa, nightingales sing during the day and produce mostly "sub-song" (i.e., highly variable juvenile song during development) or "plastic
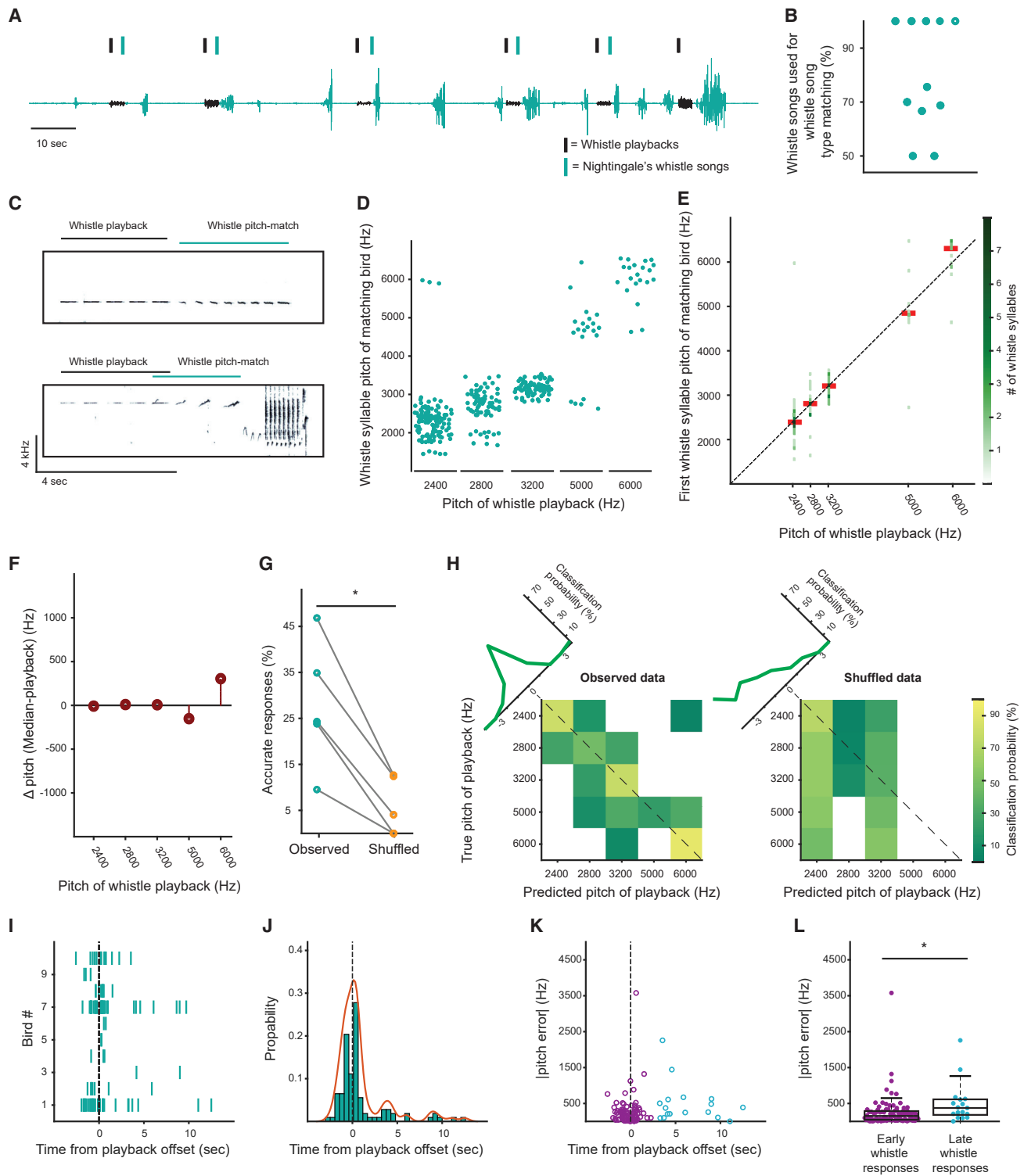
Figure 4. Pitch matching is conserved outside of the breeding season

(A) Oscillogram of an example whistle playback experiment in The Gambia with one nightingale. Whistle playbacks are highlighted in black in the oscillogram. Vertical tick marks indicate onsets of whistle playbacks (black) and of whistle responses (green).

(B) Percentage of whistle songs used for whistle song replies to whistle playbacks in The Gambia.

(C) Spectrograms for two examples of whistle pitch matching of synthetic whistle playbacks with different pitches, outside the breeding season.

(D) Pitch of whistle syllables sung in response to artificial whistle playbacks. Individual data points are plotted as green circles.

(legend continued on next page)

songs" that are thought to serve a practice function before reaching the crystallized songs that are produced during breeding season.[20] Although multiple studies have investigated nocturnal singing interactions during breeding season,[23–25] it is not known whether and how nightingales engage in song interactions in their wintering quarters. We asked whether real-time sensorimotor conversion of an arbitrary pitch is restricted to their mating season or exists year-round. We recorded 10 wild nightingales singing in January in Kartong, The Gambia (Figures S1B and S1E), presented them with whistle playbacks, and extracted the peak frequency of the whistle songs sung in response (Figures 4A–4D). Due to the reduced and sporadic singing behavior of nightingales outside breeding season,[20] we limited our whistle playbacks to pitches selected based on high predictability from the behavior observed in playback experiments during breeding season (2,800, 3,200, 5,000, and 6,000 Hz, Figure 2G). Additionally, we included one playback pitch (2,400 Hz) that was poorly predicted by our classifier (Figure 2G). We found a positive correlation between the pitch of the playback and the pitch of the first whistle syllable of each whistle song response (n = 105 whistle syllables from 10 birds, Pearson's correlation, rho = 0.913, p < 0.001) (Figure 4E). Further, we found that the median pitch values of the first whistle syllables of whistle songs sung in response to each playback were not different from the pitch of the playbacks (Figure 4F). Accurate responses were more prevalent than expected by chance (Figure 4G). When applying the k-nearest neighbor classification, we found that in total, 63.68% of the playback pitches were classified correctly, whereas the classification only yielded 22.74% for the shuffled data (20% correct classification corresponds to chance level, Figure 4H). Furthermore, we analyzed the timing of whistle responses, and as observed in nightingales during the breeding season, nightingales in The Gambia also performed better whistle pitch matches when responding earlier (Figures 4I–4L).

These findings demonstrate that the pitch-matching behavior is not limited to the breeding season, a period when only crystallized song is produced, but that the vocal flexibility required for whistle matching is maintained independently of seasonal changes observed in general song production. The presence of whistle matching outside breeding season suggests that this behavior does not exclusively play a role in courtship interactions but rather is important for maintaining territories or social communication more broadly.

DISCUSSION

Our results show that nightingales can rapidly reproduce the pitch of a wide range of tonal sounds, a behavior that was previously thought to be unique to humans, as exemplified by musicians and speakers of tonal languages. This is different from other types of song-type matching in songbirds, when the target song is copied in its entirety[14,15] rather than focusing on one acoustic parameter that can be adjusted across its entire spectrum. The perception and subsequent motor adjustments required for production in the frequency domain are not limited to song-matching birds and have also been described for animal species performing jamming avoidance.[32–35] However, pitch matching requires fine-tuned motor control over vocal production to replicate the specific peak frequency of the stimulus, whereas jamming avoidance can be achieved by suppressing vocal production at a target peak frequency.

Our observation that faster and temporally overlapping whistle responses tend to be more accurate suggests that pitch matching in nightingales involves the real-time translation of an auditory signal into a graded motor command rather than the memorization and auditory-evoked retrieval of a learned vocal-motor sequence, as presumed in other forms of vocal matching.[36] The preserved ability to perform whistle pitch matching outside of the breeding season suggests both a broader social function of the matching behavior and the presence of a core brain mechanism that is independent of seasonal physiological changes related to song production. Fast and flexible auditory-vocal transformations could be carried out if a bird had previously established a sufficiently precise internal inverse model, linking sensory experiences to corresponding motor antecedents.[37] Real-time pitch matching in nightingales would therefore represent an exceptional naturally occurring case of such an inverse model "in action."

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability

(E) Correlation between pitch of whistle playbacks and first matching whistle syllables, color-coded by quantity (darker colors indicate more whistle syllables in a given pixel). Dashed line indicates the line of exact pitch matching. Red bars represent the medians of responses for each playback.

(F) Stem plot of pitch difference between whistle playback and medians of the responses to the given playback (Δ pitch [median-playback] = 32 ± 167 Hz, one-sample t test, p = 0.69).

(G) Observed and expected accurate response rates during the playback experiment conducted outside the breeding season (mean accurate responses [observed] = 27.88%, mean accurate responses [expected] = 5.27%, paired t test, p = 0.01).

(H) Left: decoding accuracy based on the k-nearest neighbor classifier. Rows indicate the true pitch of playback that the nightingales were presented with in The Gambia, and columns represent the predicted pitch. Tuning function at the diagonal represents average classification probability across all pitches. Right: as left, but for shuffled data.

(I) Whistle response latencies of individual birds aligned to whistle playback offset.

(J) Probability distribution of response latencies for all birds aligned to playback offset.

(K) Absolute pitch error depending on latency from matched whistle song offset. Color code indicates early (purple) and late (blue) responses

(L) Distribution of |pitch error|, from early and late responses (median |pitch error| early responses = 152 Hz, median |pitch error| late responses = 374 Hz, Wilcoxon rank sum test, p = 0.008). See also Figure S1, Table S1, and Audio S2.

Current Biology
Article

AUTHOR CONTRIBUTIONS

DECLARATION OF INTERESTS

REFERENCES

1. Levinson, S.C., and Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. Front. Psychol. 6, 731.

2. Levinson, S.C. (2016). Turn-taking in human communication – origins and implications for language processing. Trends Cogn. Sci. 20, 6–14.

3. Banerjee, A., and Vallentin, D. (2022). Convergent behavioral strategies and neural computations during vocal turn-taking across diverse species. Curr. Opin. Neurobiol. 73, 102529.

4. Benichov, J.I., Benezra, S.E., Vallentin, D., Globerson, E., Long, M.A., and Tchernichovski, O. (2016). The forebrain song system mediates predictive call timing in female and male zebra finches. Curr. Biol. 26, 309–318.

5. Benichov, J.I., and Vallentin, D. (2020). Inhibition within a premotor circuit controls the timing of vocal turn-taking in zebra finches. Nat. Commun. 11, 221.

6. Pika, S., Wilkinson, R., Kendrick, K.H., and Vernes, S.C. (2018). Taking turns: bridging the gap between human and animal communication. Proc. Biol. Sci. 285, 20180598.

7. Castellucci, G.A., Guenther, F.H., and Long, M.A. (2022). A theoretical framework for human and nonhuman vocal interaction. Annu. Rev. Neurosci. 45, 295–316.

8. Castellucci, G.A., Kovach, C.K., Howard, M.A., Greenlee, J.D.W., and Long, M.A. (2022). A speech planning network for interactive language use. Nature 602, 117–122.

9. Weisman, R.G., Mewhort, D.J.K., Hoeschele, M., and Sturdy, C.B. (2012). New perspectives on absolute pitch in birds and mammals. In The Oxford Handbook of Comparative Cognition, E.A. Wassermann, and T.R. Zentall, eds. (Oxford University Press), pp. 67–79.

10. Bregman, M.R., Patel, A.D., and Gentner, T.Q. (2016). Songbirds use spectral shape, not pitch, for sound pattern recognition. Proc. Natl. Acad. Sci. USA 113, 1666–1671.

11. Brosch, M., Selezneva, E., Bucks, C., and Scheich, H. (2004). Macaque monkeys discriminate pitch relationships. Cognition 91, 259–272.

12. Zatorre, R.J. (2003). Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. Nat. Neurosci. 6, 692–695.

13. King, S.L., and McGregor, P.K. (2016). Vocal matching: the what, the why and the how. Biol. Lett. 12, 20160666.

14. Krebs, J.R., Ashcroft, R., and Van Orsdol, K.V. (1981). Song matching in the great tit Parus major L. Anim. Behav. 29, 918–923.

15. Searcy, A.W., Akçay, C., Nowicki, S., and Beecher, M.D. (2014). Aggressive signaling in song sparrows and other songbirds. In Advances in the Study of Behavior, 46, M. Naguib, S.D. Healy, J. Mitani, L. Simmons, H.J. Brockmann, T. Roper, and L. Barrett, eds. (Academic Press), pp. 89–125.

16. Shakespeare, W. (1609). Sonnets (London).

17. Homer, and Lattimore, R. (1967). The Odyssey of Homer (New York).

18. Keats, J. (1820). Ode to a Nightingale. In The Oxford Companion to English Literature, D. Birch, ed. (Oxford University Press).

19. Hultsch, H., and Todt, D. (2008). Comparative aspects of song learning. In Neuroscience of Birdsong, H.P. Zeigler, and P. Marler, eds. (Cambridge University Press), pp. 204–216.

20. Kipper, S., Sellar, P., and Barlow, C.R. (2017). A comparison of the diurnal song of the Common Nightingale (Luscinia megarhynchos) between the non-breeding season in The Gambia, West Africa and the breeding season in Europe. J. Ornithol. 158, 223–231.

21. Hultsch, H., and Todt, D. (1981). Repertoire sharing and song-post distance in nightingales (Luscinia megarhynchos B.). Behav. Ecol. Sociobiol. 8, 183–188.

22. Jäckel, D., Mortega, K.G., Brockmeyer, U., Lehmann, G.U.C., and Voigt-Heucke, S.L. (2020). Unravelling the stability of nightingale song over time and space using open, citizen science and shared data. Front. Ecol. Evol. 10, 778610.

23. Hultsch, H., and Todt, D. (1996). Rules of parameter variation in homotype series of birdsong can indicate a 'sollwert' significance. Behav. Processes 38, 175–182.

24. Bartsch, C., Hultsch, H., Scharff, C., and Kipper, S. (2016). What is the whistle all about? A study on whistle songs, related male characteristics,

and female song preferences in common nightingales. J. Ornithol. *157*, 49–60.

25. Naguib, M., Mundry, R., Hultsch, H., and Todt, D. (2002). Responses to playback of whistle songs and normal songs in male nightingales: effects of song category, whistle pitch, and distance. Behav. Ecol. Sociobiol. *52*, 216–223.

26. Levitin, D.J., and Rogers, S.E. (2005). Absolute pitch: perception, coding, and controversies. Trends Cogn. Sci. *9*, 26–33.

27. Yost, W.A. (2009). Pitch perception. Atten. Percept. Psychophys. *71*, 1701–1715.

28. Fechner, G.T. (1860). Elemente der Psychophysik, Zweiter Teil (Breitkopf and Hartel).

29. Weber, E.H. (1850). Der Tastsinn und das Gemeingefühl (Wilhelm Engelmann).

30. Todt, D. (1971). Equivalent and convalent vocal reactions of a nightingale (Luscinia megarhynchos L.) with an extremely consistent song pattern. Z. Vergl. Physiol. *71*, 262–285.

31. Todt, D., and Naguib, M. (2000). Vocal interactions in birds: the use of song as a model in communication. In Advances in the Study of Behavior, *29*, P. Slater, J. Rosenblatt, C. Snowdon, and T. Roper, eds. (Academic Press), pp. 247–296.

32. Bullock, T.H., Hamstra, R.H., and Scheich, H. (1972). The jamming avoidance response of high frequency electric fish. J. Comp. Physiol. *77*, 1–22.

33. Ulanovsky, N., Fenton, M.B., Tsoar, A., and Korine, C. (2004). Dynamics of jamming avoidance in echolocating bats. Proc. Biol. Sci. *271*, 1467–1475.

34. Gillam, E.H., Ulanovsky, N., and McCracken, G.F. (2007). Rapid jamming avoidance in biosonar. Proc. Biol. Sci. *274*, 651–660.

35. Jones, T.K., Allen, K.M., and Moss, C.F. (2021). Communication with self, friends and foes in active-sensing animals. J. Exp. Biol. *224*, jeb242637.

36. Janik, V.M. (2000). Whistle matching in wild bottlenose dolphins (Tursiops truncatus). Science *289*, 1355–1357.

37. Giret, N., Kornfeld, J., Ganguli, S., and Hahnloser, R.H.R. (2014). Evidence for a causal inverse model in an avian cortico-basal ganglia circuit. Proc. Natl. Acad. Sci. USA *111*, 6063–6068.

38. Amrhein, V., Korner, P., and Naguib, M. (2002). Nocturnal and diurnal singing activity in the nightingale: correlations with mating status and breeding cycle. Anim. Behav. *64*, 939–944.

39. Brumm, H. (2004). The impact of environmental noise on song amplitude in a territorial bird. J. Anim. Ecol. *73*, 434–440.

40. Brumm, H., and Todt, D. (2004). Male–male vocal interactions and the adjustment of song amplitude in a territorial bird. Anim. Behav. *67*, 281–286.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Code and datasets for analysis	Github	https://github.com/vallentinlab/NGwhistles
Experimental models: Organisms/strains		
Nightingale (*Luscinia megarynchos*)	Wild populations	N/A
Software and algorithms		
Matlab	MathWorks	https://www.mathworks.com/products/matlab.html
Avisoft SASLab	R. Specht, Berlin, Germany	Pro 5.2
Raven	Cornell Lab of Ornithology, Ithaca, USA	Lite 2
Audacity	https://www.audacityteam.org/)	v.2.4.2
Other		
battery-driven pre-amplifier	Roland, Japan	Duo-Capture EX
directional parabolic microphone	Telinga, Sweden	Stereo MK3
Speaker	JBL, Harman International Industries, USA	JBL
sound pressure level meter SPL	Voltcraft, Germany	SL-100
portable recorder	TEAC Corporation, USA	Tascam DR-40X

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Daniela Vallentin (daniela.vallentin@bi.mpg.de).

Materials availability
This study did not generate new unique reagents.

Data and code availability
The dataset and codes generated during this study and any additional information required to reanalyze the data reported in this paper are available here: https://github.com/vallentinlab/NGwhistles

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Common nightingales (*Luscinia megarhynchos*) are migratory songbirds that breed in the Western Palearctic and winter in sub-Saharan Africa. For experiments during breeding season, we studied a wild population of nightingales in semi-urban areas of Teltow-Fläming in the southwestern part of Brandenburg, Germany during mating season (April - May) in 2020 and 2021 (Figure S1). Unmated males during mating season are highly territorial and sing at night from designated singing posts within their territories, allowing identification of individuals over multiple days[21,38] (Video S1). All recordings were performed in accordance with the local authorities (Landesamt für Umwelt – Land Brandenburg LFU-N4 4730/14+5#181132/2021).

For experiments outside breeding season, we studied a wild population of wintering nightingales in the Kartong area in The Gambia in January 2023 (Figure S1). All recordings were performed in accordance with the local authorities (AHB 8/54/01 (90)).

METHOD DETAILS

Generation of artificial playbacks
A 3 second white noise signal was generated using the randn function in Matlab (MathWorks, Inc., Natick, MA, U.S.A.). For each whistle stimulus this signal was bandpass filtered with a bandpass spanning 60 Hz around the designated pitch frequency of either 1000 to 4000 Hz (in linear steps of 200 Hz) or 4400 Hz or 5000 Hz to 8000 Hz (in linear steps of 1000 Hz). Then four silent gaps of 200 ms were

introduced after every 440 ms of sound resulting in 5 repetitions of a whistle syllable. To avoid clipping artifacts the onset and offset of each whistle was gradually faded in or out respectively in a 1ms window.

Audio recordings in Germany

Audio recordings (16-bit precision at 44.1 kHz sampling rate) were made from a total of 33 wild male nightingales during the mating season of 2020 and 2021. Nocturnal recording sessions were conducted between 11pm and 4am CET+1, when nightingales performed dyadic counter-singing. Each nightingale was recorded with a directional parabolic microphone equipped with a windshield (Stereo MK3, Telinga, Sweden), connected to a battery-driven pre-amplifier (Roland Duo-Capture EX, Roland, Japan) and a laptop computer. Signals were acquired using Audacity v.2.4.2 (https://www.audacityteam.org/) and encoded as stereo.-wav files.

Paired recordings
Both nightingales within a pair of interacting birds (10 pairs, 2 from 2020 and 8 from 2021, in total 20 birds) were recorded simultaneously. Microphones were placed 5-15 meters away from the birds singing from two locations within their territory, spaced 30-80 meters apart.

Song playback experiment
For playback experiments with 13 birds, a speaker (JBL, Harman International Industries, USA) was connected to the pre-amplifier and placed at ∼10 meters from the singing bird and whistle playbacks were broadcasted at a sound pressure level of ∼80 dB (A) (dB re. 20 μPa measured with a sound pressure level meter SPL (SL-100, Voltcraft, Germany) to simulate a bird singing at ∼95 dB (A) and located ∼56 meters apart.[39,40]

$$Sound\ pressure\ (D2) = Sound\ pressure\ (D1) - 20 \cdot Log\ 10(D2/D1)$$

with

    Sound pressure (D1) = Known sound pressure level at the SPL held 1m from the speaker.
    Sound pressure (D2) = Unknown sound pressure level perceived by the nightingale
    D1 = Distance from the speaker and the SPL.
    D2 = Distance from speaker to the nightingale.

Each nightingale was recorded during two sessions conducted during two successive nights. A period of 13-39 min (25.14±6.73 min) was recorded while the bird was singing without playback stimulation (pre-playback control). The playback phase consisted of 200 whistle song playbacks which were manually triggered depending on the singing behavior of the nightingale. The stimuli were presented in a pseudo-randomized order (which was the same across animals). Playbacks were started during the silent gaps in between songs to avoid active overlapping with the singing bird. The playbacks were only presented after at least two consecutive non-whistle songs of the nightingale. After playback stimulation a period of 15-43 min (24.89±6.79 min) was recorded while the bird continued to sing (post-playback control).

Audio recordings in The Gambia

Audio recordings (16-bit precision at 44.1 kHz sampling rate) were made from a total of 10 wild male nightingales wintering in The Gambia in January 2023. Nightingales were located by presenting nightingale song playbacks to evoke song responses during transects between 7-11am and 4-6pm. A speaker (JBL, Harman International Industries, USA) was used to broadcast whistle playbacks and the responses of the singing birds were recorded using a portable recorder (Tascam DR-40X, TEAC Corporation, USA). The playbacks were only presented in the silence gaps between songs and after at least two consecutive non-whistle songs of the singing nightingale.

Audio data processing and analysis

Audio recordings from Germany were processed and analyzed using Audacity, Avisoft SASLab Pro 5.2 (R. Specht, Berlin, Germany) and Matlab.

For pair recordings, 16-33 min (28.14±5.87 min) of continuous recordings without contaminating background sounds were selected for data analysis.

For all recordings, stereo files were divided in mono files and a high pass filtered (frequency=1000 Hz, roll-off=6 dB, High-Pass Filter build-in function of Audacity) and noise reduced to remove environmental noise (noise reduction=12dB, sensitivity=6.00, frequency smoothing=3, noise reduction build-in function of Audacity). Sonograms were examined in Audacity for visual scoring of whistle song type matches. A song was considered a whistle song type match when the following criteria were met (1) being a whistle song (containing frequency unmodulated whistle syllables) and (2) being an immediate response to the matched song (i.e. the first or second song sung by the matching bird within 15 seconds after the termination of the matched song).

Avisoft was used for semi-automatic segmentation of songs, using the built-in functions Create section labels from waveform events with parameters adjusted case-by-case. Onset and offset of detected songs were visually inspected and manually corrected, if needed. Individual whistle syllables were manually segmented from whistle songs.

Since audio recordings from The Gambia were obtained as mono tracks (including playback and nightingale responses), they were pre-processed using Raven (Cornell Lab of Ornithology, Ithaca, USA) for manual extraction of acoustic features. Pitch of segmented whistle syllables was extracted using the pitch function estimating the fundamental frequency over the course of one whistle in Matlab.

QUANTIFICATION AND STATISTICAL ANALYSIS

We used Matlab for all data analysis and statistics. All values are reported as mean ± standard deviation if not noted otherwise. Statistical tests applied are mentioned in the text.

Generating shuffled data
For pair recordings: the observed pitches of the first whistle syllables of the matching bird were randomly assigned to an observed pitch of the first whistle syllable of the matched bird. To extract the respective parameters (% inaccurate responses and medians of | distance from pitch match|) the shuffling was repeated 1000 times.

For playback experiment: the observed pitches of whistle syllables of the whistle matches were randomly assigned to the playback pitches. The number of whistle match occurrences was preserved for each playback pitch.

Clustering of whistle syllables
Whistle syllables were clustered using the kmeans cluster function. To identify the optimal number of clusters (k), we validated the consistency within clusters of data using the function silhouette. We extracted mean values for 3 to 30 potential clusters and determined the one with the highest mean to be the optimal (k=10).

Whistle syllable pitch distribution of birds matching playbacks
The global maximum (Max) of the whistle pitch distributions were extracted as the maximum of findpeaks of the ksdensity function with specified parameters 'Support', 'positive'.

Distribution of early and late responses
Peaks of early and late responses were identified from the ksdensity probability function with the findpeaks function; the first trough with the islocalmin function.

Expected song type matching with whistle songs
An interaction with 300 songs randomly interspersed with 20% whistle songs was simulated and the percentage of whistle song type matches was identified as described above. This procedure was repeated 1000 times and the mean and median of the resulting distribution represent the expected percentage of whistle song type matches based on the observed number of whistle songs (see Figure S2B).

Classification of song types
To investigate whether nightingales copied the full song syllable sequence, whistle songs in response to whistle songs were visually scored and considered to be the same song type if all syllables were shared and followed the same order in the corresponding pairs of songs.

Distance from pitch match
The median of all Euclidean distances between each first whistle response syllable and the stimulus pitch (either the neighboring nightingale or the playback) was calculated. To determine whether this median is significantly different from chance, we independently shuffled all observed whistles syllables and the stimuli whistles and calculated the Euclidean distance of the resulting response pitches to the newly assigned stimulus pitches. Then we extracted the median from the resulting distribution. We repeated this procedure a thousand times which resulted in a distribution of 1000 expected medians. Then we compared the observed median with the expected ones and identified it as significantly different if the observed median was less than 95% of the expected medians.

Decoding classifier
To predict the pitch of the playback stimulus based on the whistle pitches produced by the nightingales, we trained a supervised machine-learning k-nearest neighbor classifier (Classification Learner, Matlab 2020b) to predict the playback classes from the responses of the birds (2-fold cross-validation, 100 neighbors for breeding season data, 10 neighbors for non-breeding season data, Euclidian distance).