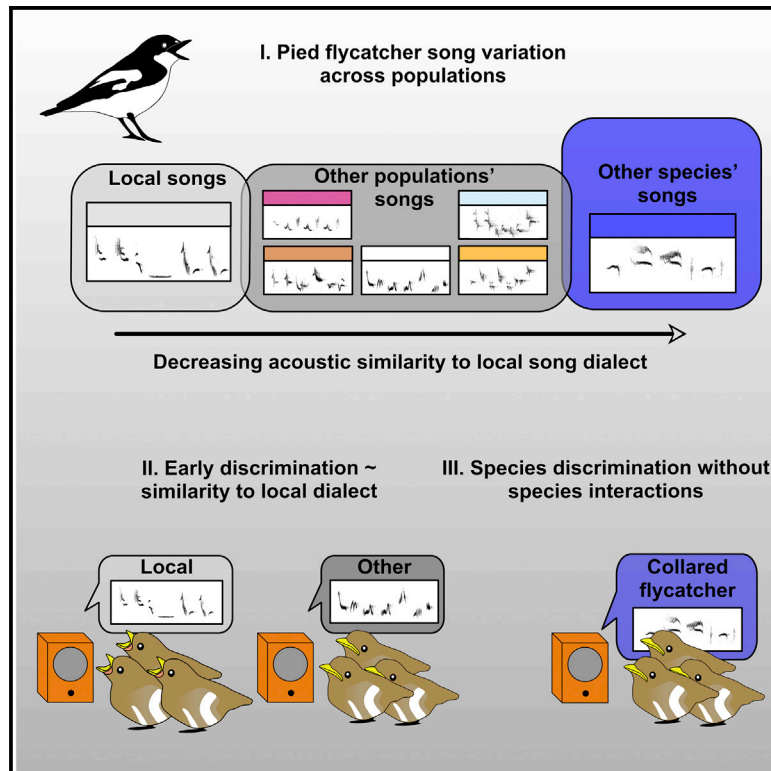# Species-specific song responses emerge as a by-product of tuning to the local dialect

Graphical abstract

Authors

David Wheatcroft, Louis Bliard,
Myriam El Harouchi, ...,
Jesús Martínez-Padilla,
Marcel E. Visser, Anna Qvarnström

Correspondence

david.wheatcroft@zoologi.su.se

In brief

Wheatcroft et al. show that nestling pied flycatchers preferentially beg in response to songs from their own population's dialect. Nestlings beg weakly to songs from other populations and from collared flycatchers, a closely related species. These findings suggest that cultural evolution drives the emergence of pre-mating reproductive barriers.

Highlights

- Begging in nestling pied flycatchers is tuned to their own population's song dialect

- Nestling begging is weaker to the songs of other populations and species

- Dialect tuning may promote the emergence of pre-mating reproductive barriers

CellPress

Current Biology

Report

Species-specific song responses emerge as a by-product of tuning to the local dialect

David Wheatcroft,[1,2,15,16,*] Louis Bliard,[2,3] Myriam El Harouchi,[4] David López-Idiáquez,[5,6] Tiia Kärkkäinen,[7] Fanny-Linn H. Kraft,[1] Jaime Muriel,[8] Samyuktha Rajan,[1] Tomas Tuvillo,[9] Malcolm D. Burgess,[4,10,11] Alejandro Cantarero,[7,12,13] Toni Laaksonen,[7] Jesús Martínez-Padilla,[14] Marcel E. Visser,[9] and Anna Qvarnström[2]
[1]Ethology, Department of Zoology, Stockholm University, Svante Arrhenius väg, Stockholm 106 91, Sweden
[2]Animal Ecology, Department of Ecology and Genetics, Uppsala University, Norbyvägen, Uppsala 752 36, Sweden
[3]Department of Evolutionary Biology and Environment Studies, Zurich University, Winterthurerstrasse, Zurich 8057, Switzerland
[4]Piedfly.Net, Yarner Wood, Newton Abbot TQ13 9LJ, UK
[5]CEFE, CNRS, University of Montpellier, EPHE, IRD, route de Mende, Montpellier 34293, France
[6]Department of Plant Biology and Ecology, University of the Basque Country UPV/EHU, Barrio Sarriena, Leioa 48940, Spain
[7]Department of Biology, University of Turku, Vesilinnantie, Turku 20014, Finland
[8]Instituto de Investigación en Recursos Cinegéticos (IREC-CSIC, UCLM, JCCM), Ronda de Toledo, Ciudad Real 13005, Spain
[9]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg, Wageningen 6708 PB, the Netherlands
[10]RSPB Centre for Conservation Science, Potton Road, Sandy SG19 2DL, UK
[11]Centre for Research in Animal Behaviour, University of Exeter, Perry Road, Exeter EX4 4QG, UK
[12]Department of Evolutionary Ecology, Museo Nacional de Ciencias Naturales (CSIC), C. de José Gutiérrez Abascal, Madrid 28006, Spain
[13]Department of Physiology, Veterinary School, Complutense University of Madrid, Avenida Puerta de Hierro, Madrid 28040, Spain
[14]Instituto Pirenaico de Ecología (CSIC), Avda. Nuestra Señora de la Victoria, Jaca 22700, Spain
[15]Twitter: @wheatcroftlab
[16]Lead contact
*Correspondence: david.wheatcroft@zoologi.su.se
https://doi.org/10.1016/j.cub.2022.09.063

SUMMARY

Oscine birds preferentially respond to certain sounds over others from an early age, which focuses subsequent learning onto sexually relevant songs.[1–3] Songs vary both across species and, due to cultural evolution, among populations of the same species. As a result, early song responses are expected to be shaped by selection both to avoid the fitness costs of cross-species learning[4] and to promote learning of population-typical songs.[5] These sources of selection are not mutually exclusive but can result in distinct geographic patterns of song responses in juvenile birds: if the risks of interspecific mating are the main driver of early song discrimination, then discrimination should be strongest where closely related species co-occur.[4] In contrast, if early discrimination primarily facilitates learning local songs, then it should be tuned to songs typical of the local dialect.[5–7] Here, we experimentally assess the drivers of song discrimination in nestling pied flycatchers (*Ficedula hypoleuca*). We first demonstrate that early discrimination against the songs of the closely related collared flycatcher (*F. albicollis*) is not strongly affected by co-occurrence. Second, across six European populations, we show that nestlings' early song responses are tuned to their local song dialect and that responses to the songs of collared flycatchers are similarly weak as to those of other conspecific dialects. Taken together, these findings provide clear experimental support for the hypothesis that cultural evolution, in conjunction with associated learning predispositions, drives the emergence of pre-mating reproductive barriers.

RESULTS AND DISCUSSION

Species discrimination without species interactions

In pied flycatchers, song is produced by males on their arrival to the breeding grounds, during territorial interactions and pair formation,[8] and plays an important role in female mate choice.[9,10] Pied flycatchers co-occur with closely related collared flycatchers (*F. albicollis*) in central Europe as well as the Baltic islands of Gotland and Öland.[11] The two species are highly similar in their morphology, ecology, and plumage, and they occasionally

hybridize.[11] Hybrids are completely sterile,[12] implying large costs to cross-species song learning.[13] Where they co-occur, pied flycatcher nestlings discriminate against the songs of collared flycatchers, an ability that arises independently of the early environment.[14] If early song discrimination is primarily influenced by selection to avoid cross-species mating interactions, nestlings should more strongly discriminate against collared flycatcher songs where the two species co-occur (Figure 1C).

Since songs vary within and across populations of both species,[15] we analyzed songs from 168 individual pied and collared
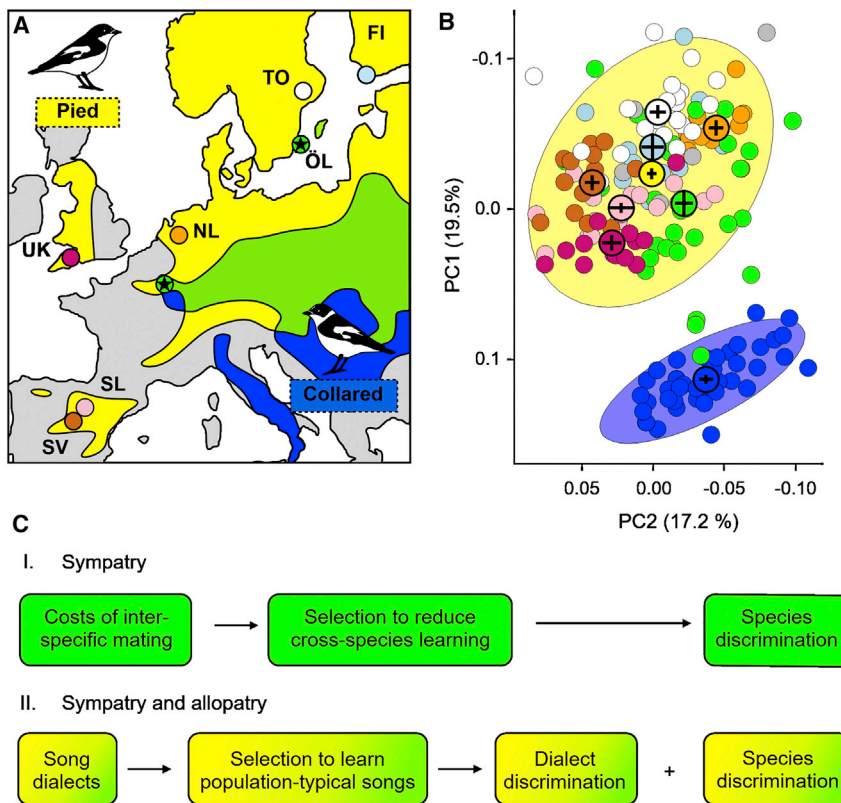
Figure 1. Study populations, song variation, and hypotheses

(A) Study sites: England (UK), Spain-Valsaín (SV), Spain-La Hiruela (SL), the Netherlands (NL), Öland (ÖL), Tovetorp (TO), and Finland (FI). Yellow denotes breeding range of pied flycatchers, blue denotes breeding range of collared flycatchers, and green denotes where both species co-occur in sympatry. Starred green circles denote sympatric locations used to estimate each population's distance to sympatry.

(B) Species and population variation in song. PC1 and PC2 scores for song phrases, averaged within individual males (see Figure S1A). Colors refer to populations in (A). Gray denotes recordings obtained from additional populations. Large circles denote population means. Whiskers give SE around each population mean. Ellipses contain 95% of the individuals for each species. Large yellow and blue circles denote species means.

(C) Overview of hypotheses explaining species discrimination in nestling songbirds. (I) The costs of interspecific mating in sympatry drive selection to reduce cross-species learning, which, in turn, leads to early species discrimination. (II) Selection to learn population-typical songs promotes early dialect discrimination. Due to larger song differences between species than among dialects, this leads to the pattern of species discrimination even in allopatric populations. Colors of boxes denote whether the hypothesis predicts species discrimination in sympatry (green) or both allopatry and sympatry (mixed).

flycatcher males in order to obtain a measure of song variation independent from species category (Figure S1A). Pied flycatcher song phrases consist of 2–2.5 s long sequences of around 9 smaller units, termed syllables, that are often repeated within and across song phrases.[16,17] We took multiple acoustic measures of the syllables in each individual's song and found that PC1, a single principal component summarizing acoustic variation in song phrases, correctly classified 100% of the collared flycatcher singers (N = 41 individuals, mean PC1 score ± SD, 0.11 ± 0.02) and 97.6% of the pied flycatcher singers (N = 127, −0.02 ± 0.04) to the species level (Figure 1B; https://doi.org/10.17632/v49k6v4fw3.1). Given this independent support for classifying songs at the species level, we considered playback species as a categorical variable when explaining song responses. We found strong evidence that pied flycatcher nestlings discriminate in favor of pied flycatcher songs. In an extensive playback experiment across six populations (Figure S1B; Video S1), nestlings produce more begging calls in response to playbacks of pied flycatcher songs (estimated coefficient on pied flycatcher playbacks: 0.33 ± 0.10 SE, likelihood-ratio test: $X_1^2 = 10.73$, p = 0.001; Figure 2; Table S1; https://doi.org/10.17632/v49k6v4fw3.1). We found little evidence for an effect of age on species discrimination (likelihood-ratio test: $X_1^2 = 0.11$, p = 0.74), and discrimination was similar across the tested nestling age categories (Table S1).

Next, we evaluated the effect of co-occurrence on early species discrimination by comparing nestling responses in allopatry with those previously collected in sympatry on Öland. We found no evidence of an effect of co-occurrence on species

discrimination (likelihood-ratio test: $X_1^2 = 0.48$, p = 0.49). Post hoc tests suggest stronger responses to pied flycatcher playbacks in both allopatry and sympatry with collared flycatchers (Figure 2; Table S2). Although typical pied flycatcher natal dispersal distances are between 5–20 km,[18–20] there are extreme reports of up to 660 km.[21] We thus considered that populations even outside of current zones of contact might be or have been in the recent past exposed to species interactions. Therefore, we assessed the impact of geographic distance to sympatry on nestling song responses. We found little evidence for a positive effect of proximity to sympatry on species discrimination (likelihood-ratio test: $X_1^2 = 0.60$, p = 0.44): discrimination was similar at varying distances to sympatry (Table S2).

These results run counter to the idea that early discrimination against collared flycatcher songs arises primarily due to selection against cross-species learning. Instead, our results demonstrate that the ability to discriminate collared from pied flycatcher songs arises independently of co-occurrence. Although species interactions are hypothesized to impact responses to sexual traits in general,[22] we have shown here that "species discrimination" in nestling birds is a pattern that has arisen in the apparent absence of them (i.e., "endogenously").[23] However, species interactions may potentially influence discrimination both directly and indirectly in other ways. First, experience can modify the strength of responses to other species' sexual signals.[24] On Öland, where both pied and collared flycatchers co-occur, adult female pied flycatchers express stronger preferences for pied flycatcher songs compared
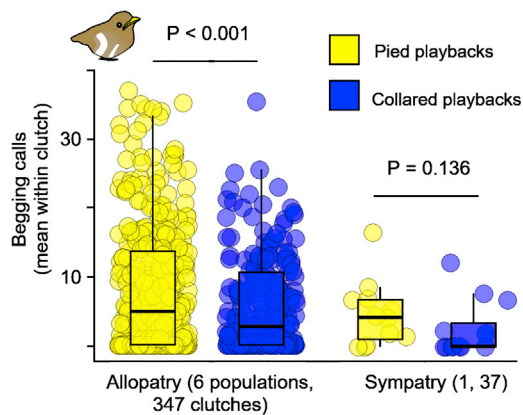
Figure 2. Song discrimination in allopatry and sympatry

Nestling responses in allopatry (left) and sympatry (right) to playbacks of collared and pied flycatcher songs (see Figure S1B). Clutches that produced no begging calls during any playback were removed for plotting purposes. Results are averaged across nestlings within nest. Boxplots give median, quartiles, and quartiles ± 1.5 × IQR. p values are from post hoc tests (see Tables S1 and S2).

with in allopatry,[10] and collared flycatcher males have more species-specific song responses compared with juveniles[25] (see also Hudson et al.[26]), suggesting that experience gained throughout development may strengthen the song discrimination we observe in nestling birds. Second, discrimination against the songs of collared flycatchers could have evolved in pied flycatcher nestlings as a by-product of selection against learning the songs of other species that co-occur with pied flycatchers throughout their breeding range, such as the great tit.[8] Even in this scenario, we have shown that interactions with collared flycatchers in particular are not required for the emergence of discrimination against their songs.[23]

Song discrimination tuned to the local dialect

Adult birds are thought to gain social and sexual benefits by producing and preferring song types that are common in their population,[6,7] which drive discrimination toward locally typical songs.[27,28] However, social transmission of songs promotes cultural evolution within populations,[29,30] leading to the prediction that early song responses should themselves evolve to track ongoing song evolution.[27] Co-evolution between songs and associated learning predispositions might therefore promote discrimination against the songs of non-co-occurring species and even among populations of the same species.[26]

We evaluated the hypothesis that early song discrimination arises due to selection to learn locally common songs by testing the prediction that nestlings respond strongest to songs typical of their own population (Figure 1C). Song acoustic features vary statistically across pied flycatcher populations. The majority (82.6%) of song phrases could be correctly assigned to population in a linear discriminant analysis (Figure 1B; https://doi.org/10.17632/v49k6v4fw3.1), providing clear support for the existence of song dialects. Next, we assessed whether nestling song responses are tuned to their own population's song dialect. We related nestling begging responses to a given playback based on the acoustic distance between the playback and the

nestlings' local dialect ("local-similar," "local-dissimilar," "collared-similar"; Figure 3A). We found a clear effect of similarity to the local dialect on responses (likelihood-ratio test: $X_2^2 = 6.74$, $p = 0.034$), with post hoc tests indicating that nestlings respond strongest to local-similar playbacks and weakest to local-dissimilar and collared-similar playbacks (Table S3). In addition, there was a clear effect of age on these responses (likelihood-ratio test: $X_2^2 = 9.93$, $p = 0.007$). Although there was no clear effect of age on discrimination between local-similar and collared-similar playbacks, discrimination between local-similar and local-dissimilar songs increased with nestling age (Table S3), with a clear difference between ages 11 and 12, the age categories with the largest sample sizes (interaction contrast: $p = 0.004$; Figures 3B and 3C). To account for the possibility that increased responses to local-similar songs with age are a consequence of the large difference in sample sizes of nestlings aged 11 and 12, we bootstrapped our results using subsampling. Models run on 99 of 100 sub-sampled datasets showed significantly stronger discrimination between local-similar and local-dissimilar songs at age 12 compared with age 11 (interaction contrasts: $p < 0.015$; https://doi.org/10.17632/v49k6v4fw3.1), implying that differences in song responses across age are not due to variation in sample size.

A long-standing debate has centered around the role of song dialects in the emergence of incipient reproductive barriers.[31] Some recent studies have suggested that dialects may influence patterns of mating in mixed populations. In white-crowned sparrows, dialect divergence is in part linked with genetic divergence, and song acts as a potential behavioral barrier to reproduction between subspecies.[32] Moreover, juvenile sparrows preferentially learn their own dialect,[33] suggesting direct links between divergence in songs and early discrimination. We evaluated the association between dialect divergence, genetic divergence, and geographic distance using previously published pairwise F_{st} measures for five populations included in our study.[34] We found support for a model including both genetic and geographic distance ($r^2 = 0.60$, $F = 5.18$, $p = 0.01$), which demonstrated a clear positive association between dialect and genetic divergence ($t = 3.11$, $p = 0.010$) and a trend for a negative association with geographic distance ($t = -2.17$, $p = 0.088$). Since we show that early nestling responses to songs are tuned to their local dialect, these results are consistent with the idea that genetic variation among populations might directly impact both song dialect and early discrimination, thereby promoting the emergence of pre-mating barriers among flycatcher populations.[27] Although perceptual predispositions expressed early in life can be significantly altered by experience,[28,33] weak responses to songs outside the typical range of the local dialect, irrespective of species, imply that dialect-specific responses are likely to promote incipient pre-mating reproductive barriers between the collared and pied flycatcher prior to secondary contact.

Previous work on collared flycatchers identified the earliest metabolic responses to songs around day 4, while species discrimination did not develop until around day 9.[35] This is consistent with what is known about auditory development in the pied flycatcher. Electrophysiology experiments on nestling birds demonstrated responses in the auditory cortex to tone frequencies typical of pied flycatcher vocalizations between
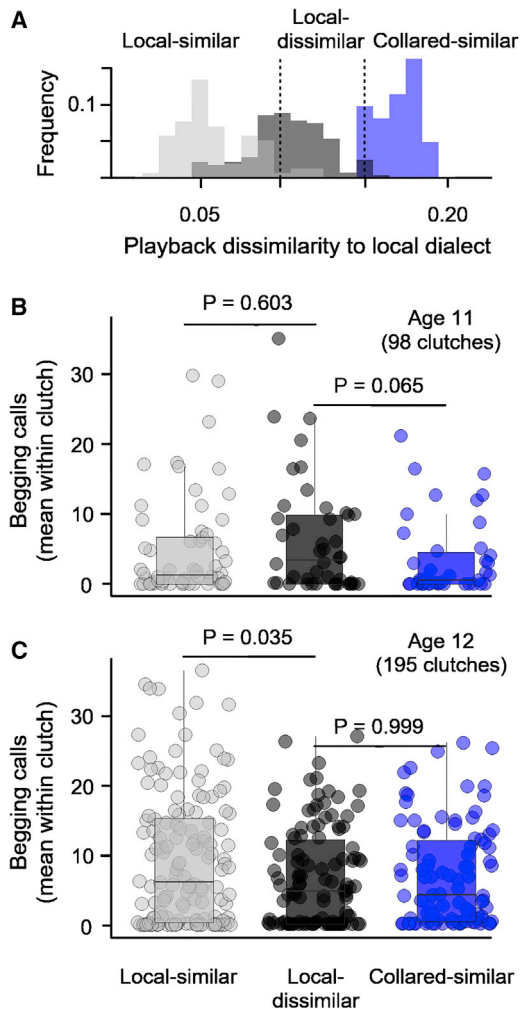
A

B

C

Figure 3. Playback categories and nestling responses to local-similar song playbacks

(A) Euclidean distance between playbacks and local dialect centroids. Light gray, local songs; dark gray, foreign songs; blue, collared flycatcher songs. Playback categories denoted by dashed lines. Local-similar category contains 95% of local dialect songs, while collared-similar category contains 90% of collared flycatcher songs.

(B) Responses of 11-day-old nestlings to the different playback categories.

(C) Responses of 12-day-old nestlings to the different playback categories. Clutches that produced no begging calls during any playback were removed for plotting purposes. Responses are averaged across nestlings within clutches. Boxplots give median, quartiles, and quartiles ± 1.5 × IQR. p values are from post hoc tests (see Figure S1 and Table S3).

days 4 and 5,[36,37] whereas specific responses to natural pied flycatcher vocalizations did not arise until day 11.[38] Thus, the nestling auditory system first attains the ability to perceive songs by day 4, whereas discrimination among behaviorally relevant sounds develops between days 9 and 11. Our findings that song discrimination increases with nestling age are consistent with these earlier studies. The apparent earlier expression of species level (by day 11) compared with dialect-level discrimination (by day 12) implies that, like other songbirds, the pied flycatcher auditory system achieves increasingly specific responsiveness with age.[39] Next, we explored whether

increasing specificity is likely to be a result of early exposure to conspecific sounds.

No evidence for an effect of song experience on song responses

The responses of nestling songbirds to the songs of heterospecifics have been shown to be either independent of exposure to the other species' songs[40] or independent of the early environment in general.[14] The degree to which dialect-specific responses depend on previous experience is less clear. Nestling golden-crowned sparrows express stronger responses to their local dialect,[26] but experiments on the closely related white-crowned sparrow suggest that dialect-specific responses depend on previous experience.[33] Thus, nestling song responses to playbacks similar to the local dialect might arise due to learning from early song exposure from surrounding males, including the social father.

We used three tests to evaluate whether nestling song responses are affected by early exposure to conspecific songs. First, in a single population, Tovetorp, we estimated early song exposure using automated recordings at 28 nest boxes during the nestling period (https://doi.org/10.17632/v49k6v4fw3.1). We found weak evidence for an effect of song exposure on nestling song responses (likelihood-ratio test: $X_2^2 = 3.61$, p = 0.165). Contrary to the prediction that song exposure to local songs should increase discrimination, post hoc tests suggested that nestling responses to local-similar songs decrease with increasing song exposure (Table S4). Second, at 24 of these nest boxes, we measured the acoustic features of songs present in these recordings, presumed to be from the nestlings' social father (https://doi.org/10.17632/v49k6v4fw3.1). Contrary to the prediction of early learning promoting song responses, we found evidence that nestlings respond weaker to playbacks similar to the songs of their social father (estimated coefficient on dissimilarity to play back: 5.13 ± 2.82 SE, likelihood-ratio test: $X_1^2 = 3.19$, p = 0.074; Table S4). Finally, we tested whether the similarity of the social father's song to the local dialect influences responses to local-similar songs. We found little evidence for an effect of the similarity of the father's song to the local dialect (likelihood-ratio test: $X_2^2 = 0.52$, p = 0.771). Again, contrary to the prediction of early learning promoting song responses, post hoc tests suggested trends for nestling responses to local-similar songs to decrease with the similarity of the father's song to the local dialect (Table S4).

We found no evidence that nestling responses to local dialect songs are increased by early song experience. In contrast, our results suggest that nestlings respond less to playbacks similar to their social fathers' songs. The context and function of male singing during the nestling stage are unclear, but one possibility is that males sometimes produce song before feeding visits,[41] leading nestlings to naturally produce begging calls when hearing their social father's song. Males that sing at abnormally high rates during the nestling phase—for example, due to seeking out extra-pair copulations[8]—may decouple song production from feeding and thereby drive habituation of the begging response in highly exposed nestlings. This does not imply that nestlings do not learn songs from their father.[41] Rather, it suggests that early responses to dialect variation are unlikely to be explained through early learning. Without experimental manipulation of early experience, we remain cautious in our conclusions.

However, no matter the mechanism driving it, nestlings express early discrimination in favor of their own dialect shortly before leaving the nest and well before song learning begins in pied flycatchers.[8]

Conclusion

We show that pied flycatcher nestlings respond most strongly to songs with acoustic qualities typical of their own local dialect, an ability that emerges just before fledging. We demonstrate that co-occurrence with the collared flycatcher has little effect on early song responses in pied flycatchers, suggesting that the risks of cross-species mating have, at most, a minor and secondary influence on the evolution of early song discrimination. Instead, we show that nestling birds express reduced responses to songs that are dissimilar from their local dialect, irrespective of from which species they derive. This implies that species discrimination is a pattern resulting from acoustic dissimilarity of songs, rather than the costs of species interactions. Although the precise developmental mechanisms leading to associations between songs and early discrimination within populations remain to be clarified, the negligible effects of early experience on nestling song responses suggest innate development. Our results therefore support the long-held hypothesis that cultural evolution can drive the emergence of song learning predispositions, which hastens the divergence in mate recognition systems between closely related species.[42] The tremendous species richness of vocal learning birds[43] may thus be partially explained as a side-effect of cultural evolution.

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
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Song recordings and playbacks
  - Song analysis
  - Playback recordings
  - Playback experiments
  - Behavior analysis
  - Estimating song exposure
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.09.063.

AUTHOR CONTRIBUTIONS

D.W. and A.Q. designed the experiment. D.W. and L.B. analyzed data. All authors shared in field research and writing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Shizuka, D. (2014). Early song discrimination by nestling sparrows in the wild. Anim. Behav. *92*, 19–24.

2. Wheatcroft, D., and Qvarnström, A. (2015). A blueprint for vocal learning: auditory predispositions from brains to genomes. Biol. Lett. *11*, 20150155.

3. Nelson, D.A., and Marler, P. (1993). Innate recognition of song in white-crowned sparrows: a role in selective vocal learning? Anim. Behav. *46*, 806–808.

4. Olofsson, H., Frame, A.M., and Servedio, M.R. (2011). Can reinforcement occur with a learned trait? Evolution *65*, 1992–2003.

5. Lachlan, R.F., and Servedio, M.R. (2004). Song learning accelerates allopatric speciation. Evolution *58*, 2049–2063.

6. Searcy, W.A., Nowicki, S., Hughes, M., and Peters, S. (2002). Geographic song discrimination in relation to dispersal distance in song sparrows. Am. Nat. *159*, 221–230.

7. Lachlan, R.F., Anderson, R.C., Peters, S., Searcy, W.A., and Nowicki, S. (2014). Typical versions of learned swamp sparrow song types are more effective signals than are less typical versions. Proc. Biol. Sci. *281*, 20140252.

8. Lundberg, A., and Alatalo, R.V. (1992). The Pied Flycatcher (T & AD Poyser).

9. Lampe, H.M., and Sætre, G.-P. (1995). Female pied flycatchers prefer males with larger song repertoires. Proc. R. Soc. Lond. B *262*, 163–167.

10. Wheatcroft, D.W., and Qvarnström, A. (2017). Reproductive character displacement of female, but not male song discrimination in an avian hybrid zone. Evolution *71*, 1776–1786.

11. Qvarnström, A., Rice, A.M., and Ellegren, H. (2010). Speciation in Ficedula flycatchers. Philos. Trans. R. Soc. Lond. B Biol. Sci. *365*, 1841–1852.

12. Ålund, M., Immler, S., Rice, A.M., and Qvarnström, A. (2013). Low fertility of wild hybrid male flycatchers despite recent divergence. Biol. Lett. *9*, 20130169.

13. Qvarnström, A., Haavie, J., Saether, S.A., Eriksson, D., and Pärt, T. (2006). Song similarity predicts hybridization in flycatchers. J. Evol. Biol. *19*, 1202–1209.

14. Wheatcroft, D., and Qvarnström, A. (2017). Genetic divergence of early song discrimination between two young songbird species. Nat. Ecol. Evol. *1*, 0192.

15. Haavie, J., Borge, T., Bureš, S., Garamszegi, L.Z., Lampe, H.M., Moreno, J., Qvarnström, A., Török, J., and Saetre, G.-P. (2004). Flycatcher song in allopatry and sympatry—convergence, divergence and reinforcement. J. Evol. Biol. *17*, 227–237.

16. Gelter, H.P. (1987). Song differences between the pied flycatcher *Ficedula hypoleuca*, the collared flycatcher *F. albicollis*, and their hybrids. Ornis Scand. *18*, 205–215.

17. Espmark, Y.O., and Lampe, H.M. (1993). Variations in the song of the pied flycatcher within and between breeding seasons. Bioacoustics *5*, 33–65.

18. Paradis, E., Baillie, S.R., Sutherland, W.J., and Gregory, R.D. (1998). Patterns of natal and breeding dispersal in birds. J. Anim. Ecol. *67*, 518–536.

19. Sokolov, L. (2000). Philopatry, dispersal and population structure of passerines on the Courish Spit of the Baltic Sea. Vogelwarte *40*, 302–314.

20. Chernetsov, N., Sokolov, L.V., Kosarev, V., Leoke, D., Markovets, M., Tsvey, A., and Shapoval, A.P. (2006). Sex-related natal dispersal of pied flycatchers: how far away from home? Condor *108*, 711–717.

21. Both, C., Robinson, R.A., and van der Jeugd, H.P. (2012). Long-distance dispersal in migratory pied flycatchers Ficedula hypoleuca is relatively common between the UK and the Netherlands. J. Avian Biol. *43*, 193–197.

22. Hudson, E.J., and Price, T.D. (2014). Pervasive reinforcement and the role of sexual selection in biological speciation. J. Hered. *105* (*Suppl 1*), 821–833.

23. Mendelson, T.C., and Shaw, K.L. (2012). The (mis)concept of species recognition. Trends Ecol. Evol. *27*, 421–427.

24. Verzijden, M.N., ten Cate, C., Servedio, M.R., Kozak, G.M., Boughman, J.W., and Svensson, E.I. (2012). The impact of learning on sexual selection and speciation. Trends Ecol. Evol. *27*, 511–519.

25. Bliard, L., Qvarnström, A., and Wheatcroft, D. (2021). The role of introductory alarm calls for song discrimination in *Ficedula* flycatchers. Anim. Behav. *177*, 241–251.

26. Hudson, E.J., Hahn, M., and Shizuka, D. (2019). Nestling and adult sparrows respond differently to conspecific dialects. Behav. Ecol. *30*, 48–56.

27. Lachlan, R.F., and Feldman, M.W. (2003). Evolution of cultural communication systems: the coevolution of cultural signals and genes encoding learning preferences. J. Evol. Biol. *16*, 1084–1095.

28. ter Haar, S.M., Kaemper, W., Stam, K., Levelt, C.C., and ten Cate, C. (2014). The interplay of within-species perceptual predispositions and experience during song ontogeny in zebra finches (*Taeniopygia guttata*). Proc. Biol. Sci. *281*, 20141860.

29. Derryberry, E.P. (2007). Evolution of bird song affects signal efficacy: an experimental test using historical and current signals. Evolution *61*, 1938–1945.

30. Mennill, D.J., Doucet, S.M., Newman, A.E.M., Williams, H., Moran, I.G., Thomas, I.P., Woodworth, B.K., and Norris, D.R. (2018). Wild birds learn songs from experimental vocal tutors. Curr. Biol. *28*, 3273–3278.e4.

31. Toews, D.P.L. (2017). From song dialects to speciation in white-crowned sparrows. Mol. Ecol. *26*, 2842–2844.

32. Lipshutz, S.E., Overcast, I.A., Hickerson, M.J., Brumfield, R.T., and Derryberry, E.P. (2017). Behavioural response to song and genetic divergence in two subspecies of white-crowned sparrows (*Zonotrichia leucophrys*). Mol. Ecol. *26*, 3011–3027.

33. Nelson, D.A. (2000). A preference for own-subspecies' song guides vocal learning in a song bird. Proc. Natl. Acad. Sci. USA *97*, 13348–13353.

34. Lehtonen, P.K., Laaksonen, T., Artemyev, A.V., Belskii, E., Berg, P.R., Both, C., Buggiotti, L., Bureš, S., Burgess, M.D., Bushuev, A.V., et al. (2012). Candidate genes for colour and vision exhibit signals of selection across the pied flycatcher (*Ficedula hypoleuca*) breeding range. Heredity *108*, 431–440.

35. McFarlane, S.E., Söderberg, A., Wheatcroft, D., and Qvarnström, A. (2016). Song discrimination by nestling collared flycatchers during early development. Biol. Lett. *12*, 20160234.

36. Aleksandrov, L.I. (1997). Delay in the development of auditory sensitivity in nestlings and reorganization of feeding behavior. Neurosci. Behav. Physiol. *27*, 27–29.

37. Aleksandrov, L.I., and Dmitrieva, L.P. (1992). Development of auditory sensitivity of altricial birds: absolute thresholds of the generation of evoked potentials. Neurosci. Behav. Physiol. *22*, 132–137.

38. Aleksandrov, L.I., Korneeva, E.V., and Golubeva, T.B. (2001). Increasing selectivity of defense behavior during development of pied flycatcher nestlings. Zh. Vyssh. Nerv. Deiat. Im. I P Pavlova *51*, 110–113.

39. Woolley, S.M.N. (2012). Early experience shapes vocal neural coding and perception in songbirds. Dev. Psychobiol. *54*, 612–631.

40. Hudson, E.J., Creanza, N., and Shizuka, D. (2020). The role of nestling acoustic experience in song discrimination in a sparrow. Front. Ecol. Evol. *8*, https://doi.org/10.3389/fevo.2020.00099.

41. Labra, A., and Lampe, H.M. (2018). The songs of male pied flycatchers: exploring the legacy of the fathers. PeerJ *6*, e5397.

42. Mason, N.A., Burns, K.J., Tobias, J.A., Claramunt, S., Seddon, N., and Derryberry, E.P. (2017). Song evolution, speciation, and vocal learning in passerine birds. Evolution *71*, 786–796.

43. Edwards, S.V., Kingan, S.B., Calkins, J.D., Balakrishnan, C.N., Jennings, W.B., Swanson, W.J., and Sorenson, M.D. (2005). Speciation in birds: genes, geography, and sexual selection. Proc. Natl. Acad. Sci. USA *102* (*Suppl 1*), 6550–6557.

44. Lachlan, R.F. (2016). Luscinia: a bioacoustics analysis computer program. rflachlan.github.io/Luscinia.

45. K. Lisa Yang Center for Conservation Bioacoustics (2014). Raven Pro: Interactive Sound Analysis Software (The Cornell Laboratory of Ornithology).

46. R Core Team (2019). R: a language and environment for statistical computing (R Foundation for Statistical Computing).

47. Padgham, M., and Sumner, M.D. (2020). geodist: fast, dependency-free geodesic distance calculations. https://CRAN.R-project.org/package=geodist.

48. Stoffel, M.A., Nakagawa, S., and Schielzeth, H. (2017). rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol. Evol. *8*, 1639–1644.

49. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Soft. *67*, 1–48.

50. Wang, I.J. (2013). Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. Evolution *67*, 3403–3411.

51. Kleiber, C., and Zeileis, A. (2008). Applied Econometrics with R (Springer-Verlag).

52. Brooks, M.E., Kristensen, K., Benthem, K., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Mächler, M., and Bolker, B.M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. *9*, 378–400.

53. Lenth, R. (2019). emmeans: estimated marginal means, aka least-squares means. https://CRAN.R-project.org/package=emmeans.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data	This paper; Mendeley Data	https://doi.org/10.17632/v49k6v4fw3.1
Analyzed data	Wheatcroft and Qvarnström[14]	N/A
Analyzed data	Lehtonen et al.[34]	N/A
Software and algorithms		
Luscinia	Lachlan[44]	https://rflachlan.github.io/Luscinia/
RAVEN	K. Lisa Yang Center for Conservation Bioacoustics[45]	https://ravensoundsoftware.com/
Vertex AI	Google	https://cloud.google.com/vertex-ai/
Python code	This paper	N/A
R code	This paper; Mendeley Data	https://doi.org/10.17632/v49k6v4fw3.1
R	R Core Team[46]	https://www.R-project.org/

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, David Wheatcroft (david.wheatcroft@zoologi.su.se).

Materials availability
This study did not generate new unique reagents.

Data and code availability

- Song playback information, song recording measurements, nestling behavioral observations, and song exposure data has been deposited at Mendeley Data. DOIs are listed in the key resources table.
- All original code has been deposited to Mendeley Data. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Pied flycatchers are small, insectivorous, migratory songbirds.[8] They winter in central and western Africa, while their breeding range in Europe extends from Spain and the United Kingdom in the west to northern Scandinavia and to central Russia in the east. Song playback experiments were conducted between 2018-2021 at six distinct breeding populations along a latitudinal gradient in the western portion of the pied flycatcher's range: Valsaín (Spain_V, 40°52'N, 4°01'W, 2019 and 2020), La Hiruela (Spain_LH, 41°4'N, 3°27'W, 2018 and 2020), East Dartmoor (England, 50°36'N, 3°43'W, 2018), De Hoge Veluwe National Park (The Netherlands, 52°04'N, 5°49'E, 2018), Tovetorp (Sweden, 58°57'N, 17°089'E, 2020 and 2021), and Turku (Finland, 60°27'N, 22°15'E, 2018). In addition, song playback experiments, part of a previous study,[14] were conducted between 2013-2016 on Öland, Sweden (57°10'N, 17°00'E). Geographic distances among playbacks and populations were calculated using the R package geodist 0.0.4.[47] The closest sympatric site to Tovetorp and Finland was Öland, while the closest sympatric site to all other populations was estimated in northeastern France (48°56'N, 5°1'E).

Animal procedures were approved by Jordbruksverket (Linköpings djurförsöksetiska nämnd, Dnr 01110-2020).

METHOD DETAILS

Song recordings and playbacks
Together, song recordings from 168 different individuals were analyzed. Each recording contained songs from a single individual. Most individuals were recorded from the populations where subsequent playback experiments were conducted. Four individuals from the Netherlands, one individual from England, 7 collared flycatchers, and all of the songs from Finland were obtained from

an online repository of avian vocalizations (https://xeno-canto.org/). For all other populations, recordings were made by the authors with a condenser microphone (Sennheiser K6 module and ME66 capsule, Sennheiser Electronic GmbH & Co. KG, Wedemark, Germany) connected via XLR cable to a digital audio recorder (Tascam DR-40, TEAC Corporation, Tokyo, Japan or Fostex FR-2, Foster Electric Company, Ltd, Tokyo, Japan). For each pied flycatcher population, we obtained 11-30 recordings (median: 15), whereas we obtained 50 recordings of collared flycatcher individuals from multiple populations.

Song analysis

We analyzed between 2-34 song phrases for each of the 168 individuals for which we had song recordings (median = 9). Song responses are likely to be determined by these acoustic features, rather than species identity or population identity *per se*. In order to extract a measure of song variation independent of these categories, we analyzed and compared songs from both species and from each pied flycatcher population. Songs were measured by a single author (D.W.) using Luscinia software.[44] Each song recording was imported and song phrases were measured separately. Spectrograms were visualized using a Gaussian windowing function with the following settings: 12 kHz maximum frequency, 5 ms frame length, 220 spectrograph points, 80% spectrograph overlap, 50 dB dynamic range, 30% dereverberation, and 50 ms of dereverberation range. For each song phrase, we measured individual elements as continuous sound traces. Elements were grouped into syllables if their sound traces were overlapping in time or the gap between them was less than 30 ms.

We compared songs using Luscinia's built-in dynamic time-warping (DTW) algorithm, which finds the optimal alignment of each pair of syllables based on multiple acoustic features and uses this alignment to calculate dissimilarity.[44] Songs are subsequently compared by aligning their syllables sequentially. We performed DTW using the following acoustic feature weightings: time (5), mean frequency (1), mean frequency change (1), and normalized mean frequency (1). All other measures were weighted by their standard deviation. We used a compression factor of 0.001, a minimum element length of 25 samples, time SD weighting of 1, a maximum warp of 25%, and compared syllables by stitching elements. All other options were left at the default settings. These settings were chosen based on the recommended values and because they reliably grouped song phrases from the same individual that were subjectively determined to be highly similar.[7] Luscinia generates dissimilarity matrices for each acoustic measure and then performs non-metric multi-dimensional scaling to reduce the number of dimensions while preserving overall dissimilarity. Subsequently, Luscinia performs a principal components analysis to generate 10 independent axes of variation. The eigenvalues of each principal component relate to the proportion of total variance in dissimilarity explained. Ten principal components explained 75% of the total variance in dissimilarity among song phrases. A Kruskal stress test had a stress value of 0.025 indicating a good representation of dissimilarity among songs.

Due to the potential inaccuracy in acoustic measurements of recordings in mp3 format (i.e., those taken from https://xeno-canto.org/), we repeated all analyses excluding these recordings and found highly similar results. Thus, all recordings were included in the results presented.

The 10 principal component scores were used to compare playbacks (taken as the mean of the component songs within a playback), individuals (taken as the mean centroid of the songs within an individual), population (taken as the mean centroid of the individuals within the population), and species (taken as the mean centroid of the individuals within the population). Songs varied between species in PC1 (linear mixed models with individual nested within population as random effects; N = 2088 song phrases, N = 168 individuals, $X_1^2 = 11.58$, P < 0.001) and there was a weak trend for species-level differences in PC2 ($X_1^2 = 1.98$, P = 0.159), but we did not detect evidence for differences in the other principal components (likelihood-ratio test: $X_1^2 < 1.66$, P > 0.198). PC1 scores alone correctly classified 100% of the collared flycatcher singers (n = 41 individuals, mean PC1 score ± SD, 0.11 ± 0.02) and 97.6% of the pied flycatcher singers (n = 127, -0.02 ± 0.04) to the species level. Given this independent support for species-level song classification, we considered playback species as a categorical variable in subsequent analyses.

We found that songs varied across populations in 9 out of 10 principal components (linear mixed models with individual as random effect for PC1-PC9, likelihood-ratio test: $X_5^2 > 14.20$, P < 0.015). We calculated the Euclidean distance of each playback in multidimensional principal component space to the mean centroid of each population and used this as an estimate of a playback's similarity to songs from a given population. Unsurprisingly, pied flycatcher playbacks are closer to the centroid of their source population (n = 80 playbacks; mean Euclidean distance ± SE: 0.06 ± 0.00) than to the centroids of other populations (n = 143; 0.10 ± 0.00): 95% of all playbacks are within 0.10 of their source population's centroid, while around 50% of foreign playbacks are within 0.10 of other populations' centroids. Likewise, collared flycatcher playbacks have greater Euclidean distances to population centroids compared to either local or foreign playbacks (n = 31; 0.17 ± 0.00; Figure S1B): 90% of collared playbacks, but only 3% of pied flycatcher playbacks, are more than 0.15 from population centroids. We considered nestling responses to playback similarity to songs from their own population using Euclidean distance as a factor ("local-similar": < 0.10, "local-dissimilar": ≥ 0.10 and < 0.15, "collared-similar": ≥ 0.15).

Playback recordings

One-minute-long playback files were obtained directly from the song recordings used to compare species and populations. Individual recordings that were contained more than three minutes of song were divided up into distinct playbacks. In total, we produced 245 playbacks from 168 individuals. A single author (D.W.) removed background sounds using RAVEN v1.5[45] and constructed playback files containing between 8-12 song phrases, each separated by four seconds of background noise. Background noise was obtained either from the recording itself or from recordings from the same site, in which case the amplitude was adjusted to match the background of the song recording. At each site, we played back between 26-31 unique collared flycatcher playbacks (median 30.5), 27-75 unique foreign playbacks (38.5), and 12-36 unique local playbacks (20.5).

Playback experiments

We broadcast pied and collared flycatcher song playbacks to whole clutches of pied flycatcher nestlings (N = 347 clutches, for a total of 1722 nestlings). Each nestling was marked with a unique symbol on the top of its head using correction fluid to allow individual discrimination during video analysis and placed into an experimental wooden nest box. The nest box had an artificial nest constructed of moss and/or leaves on the bottom and was fitted with a video camera (ELP-USB-130W01MT-MG40[BW], Alipu Technology Co., Ltd, Guangdong, China) and a microphone (Go Mic, Samson Technologies, Hicksville, NY, USA) positioned to record the behavior and vocalizations of all nestlings. The camera and microphone were connected to a laptop computer for recording. A speaker was positioned roughly one meter away and facing the nest box for song playbacks (Eco Extreme, Grace Digital, Poway, CA, USA). Around half of the experiments conducted at Tovetorp used a different video recording device, as described in Wheatcroft and Qvarnström.[14]

Nestlings acclimatized to the nest box for at least 5 minutes before the experiment started. An experiment consistent of three one-minute periods of silence each followed by a one-minute song playback. The order of the song playback treatments (collared, foreign and local pied flycatcher) was alternated between experiments. The individual song playbacks making up an experiment were alternated to limit repeated pairing of the same playbacks. Throughout each experiment, audio and video of the nestling behaviors were recorded. Song playbacks were broadcast using a smartphone connected to the speaker. Playbacks were broadcast at 75 dB, measured at 3 m distance (Digital Sound Level Meter, RadioShack). After the third playback, the nestlings were weighed and returned to their home nest box. Most (n = 327 out of 347) clutches were tested once. Broods were tested on a second occasion, at least 12 hours after the first, only when no nestlings produced any behavioral response on the first trial (Tovetorp: n = 20 clutches repeated). Between 29 (Finland) and 99 (Spain_V) clutches were tested at each population (median: 38, supplemental information). In addition, 37 clutches were tested on Öland.

We recorded individual mass measurements around the time of playback for most individual nestlings (n = 1718 out of 1722 nestlings, excluding Öland). In order to include behavioral responses for those nestlings without individual mass measurements, we imputed their mass by taking the mean mass of their nestmates. For nestlings with individual mass measurements, a large proportion of the variance in individual mass was explained by the mean mass of nestlings in the clutch (adjusted r^2 in a linear model = 0.73, estimated coefficient on mean mass = 1.00 ± 0.02 SE). For nestlings from Öland (n = 186 nestlings), where we could not associate individual mass with their nestling behavioral responses, we used the mean nestling mass of each clutch. Two experiments from Tovetorp, one from Spain_LH, and seven from Öland were excluded from subsequent analyses due to missing nestling mass data.

Behavior analysis

A single observer (L.B.) watched and noted the behavioral responses of all nestlings in each video recording both during periods of silence and song playbacks. Individual nestlings were tracked by the painted symbols marked on their heads. These symbols could later be associated with ring numbers and masses recorded on the day of playback. The observer was partially blind to the experimental treatment based on initial unfamiliarity with distinguishing the songs of either species and inability to distinguish the songs by population, but could identify when songs were being played back as well as the population from which each video originated. Based on previous work in this system,[14,25] the observer recorded the following behaviors as counts during each silence and song playback period: 1) begging calls (defined as an audible low-intensity begging call, Video S1), 2) gapes (defined as opening and closing the bill completely), 3) jumping begging calls (defined as a high-intensity begging call, accompanied by jumping forward or upward), 4) position shifts (defined as the nestling moving continuously from one orientation or position to another, such that there was little overlap in its body after shifting), and 5) looking up (defined as the nestling shifting the orientation of its head and bill from perpendicular to facing the nest box opening). For begging calls and jumping begging calls, the observer referred to spectrograms of extracted audio to confirm counts when multiple nestlings were begging at the same time.

Producing begging calls and looking up were the most common behaviors (n = 540 and 375 out of 1908 nestlings produced at least one begging call or look-up, respectively, during one of the playback treatments, including the Öland population). During a given playback treatment, these two behavioral measures were positively correlated within nestlings (N = 5516 playback treatments, Spearman's rho = 0.47, P < 0.001). We used begging calls as a response variable in subsequent analyses to facilitate comparison with previous studies and because it was determined to be less subjective than looking up.[14,25]

Videos of pied flycatcher nestlings from Öland (n = 37 clutches, for a total of 186 nestlings), collected using equivalent methods to the current study, were re-analyzed by L.B. These data were originally published in Wheatcroft and Qvarnström.[14] We conducted tests of repeatability of begging call counts made by L.B. and Wheatcroft and Qvarnström.[14] Since nestlings in the Öland population were not individually marked, we summarized counts within each trial by taking the mean of nestling counts within each trial. We assessed repeatability of begging call counts across observers using rptR v0.9.2.[48] Begging call counts were highly repeatable across observers for both begging calls prior to (1000 bootstrap iterations, R = 0.88, 95% CI 0.83 – 0.92, P < 0.001) and during the playback period (1000 bootstrap iterations, R = 0.89, 95% CI 0.83 – 0.92, P < 0.001)

Estimating song exposure

Audio recordings were conducted at Tovetorp in 2020 and 2021 for around three hours in the morning (approximately two hours around civil dawn and one hour between 7:00 - 8:00AM) and we obtained multiple recordings for each nest box (12.5 median; range: 4.4 - 22.3 hours of recordings per box). A single observer (D.W.) annotated spectrograms generated from 10 two-hour-long recordings in RAVEN,[45] forming selection boxes enclosing each song phrase in time and frequency. This resulted in 1,108 song phrase

selections from 10 individual males. Standardized images of each selection (1000 x 500 pixels, 10 s x 15 kHz) were generated using custom scripts in Python and used to train (80% of images), validate (10%), and test (10%) a cloud-based object detection model (https://cloud.google.com/vertex-ai/). The model had an average precision of 0.903. Evaluation of the false negatives suggested that the model tended to miss song phrases that had just begun or were soon to end within the image. Overlapping windows were used for model predictions to increase the detection rate.

The model was downloaded and run within a docker container (https://www.docker.com/). Standardized spectrograms from each audio recording were generated (500 x 500 pixels, 5 s x 15 kHz) with a 2.5 s window overlap. We obtained model predictions for each image, using custom Python scripts to account for window overlap. Predictions were filtered with a confidence score of 0.95 and all predictions were subsequently manually validated through analysis of audio and/or spectrogram. Over all boxes, we recorded an average of around one song every 10 minutes, with substantial variation across nests (0.08 ± 0.19 mean songs per minute ± SD). We estimated song exposure as the total number of song phrases divided by the total recording time. Counts of song phrases were divided by the total recording time to obtain estimated song exposure. To obtain acoustic features of the songs to which each clutch was exposed, we extracted song phrases from audio recorded at 24 nest boxes (median: 9.5 song phrases per nest box, range: 1 - 16) and analyzed the songs in Luscinia using identical methods as above. All songs were subject to a second dynamic time warping procedure with identical methods as above, and Euclidian distances between the songs to which nestlings were exposed and playbacks were calculated using all 10 principal components (see Table S4).

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical tests were performed in R v3.6.1.[46] To evaluate differences in song principal component scores between species and among populations, we used linear mixed models, implemented using *lmer* in lme4 v1.1-21,[49] using individual singer as a random intercept and species or population as a fixed effect. Significance of interactions and main effects was evaluated using likelihood ratio tests.

We evaluated the correlation between genetic, geographic, and song divergence among populations using the function *MMRR*.[50]

To evaluate variation in begging calls in response to songs, we first tested for overdispersion of begging calls using the function *dispersiontest* in package AER v. 1.2-7[51] on fixed effect poisson generalized linear mixed models with begging calls during the playback as the dependent variable. We found evidence for significant overdispersion. Similarly, visualization of the data revealed that a large number of trials consisted of zero responses (i.e., no nestling in a given clutch produced a begging call during a given playback treatment). We then applied zero-inflated quasi-poisson regression using the R package glmmTMB v1.0.2.1.[52] In all models, we included the order of the given playback within a trial, the brood size, estimated nestling mass within a given nest (g), the number of begging calls produced by each individual nestling during the immediately preceding silent period, and the age of the nestlings. In addition, we included a random effect of the playback recording used to stimulate begging calls:

(a) $begs_{playback} \sim intercept + order + brood\_size + mass + begs_{silence} + age + (1|playback)$

In addition, all models included a zero-inflation term in each model that consisted of an intercept and a random effect of the brood identity:

(b) $begs_{playback} \sim intercept + (1|brood\_id)$

In all subsequent models, we evaluated interactions between age and relevant song measures to estimate the ontogeny of discrimination (see Table S1). To test song discrimination in allopatry, we included the playback species as a fixed effect in the above model. To compare song discrimination in sympatry and allopatry, we included an interaction term between the playback species and either sympatry, as a categorical variable, or distance to sympatry, as a continuous variable (see Table S2). To test the influence of playback similarity to the local dialect, we included the similarity category ("local-similar", "local-dissimilar", "collared-similar") as a fixed effect (see Table S3). Finally, to test the effect of song exposure on song responses at Tovetorp, we included interactions between song exposure, estimated as the total number of song phrases divided by the total recording time, and the Euclidian distance between the playback and the social father's song as fixed effects (see Table S4). Due to convergence errors, we did not include random effects for models including only Tovetorp.

For all models, we tested the significance of fixed effects and interactions using log-likelihood ratio tests and, to evaluate interactions, we performed post-hoc tests using the functions *emmeans* and *emtrends* in the emmeans package v1.4.3.01.[53] For interactions between two categorical variables, we evaluated the difference between levels of one variable across levels of the other variable using the *contrast* function in emmeans. For interactions between two continuous variables, we evaluated the effect of one variable at set levels of the other variable (mean, mean ± SD).

The majority of experiments (84% of N = 347) were conducted at nests with nestlings age 11 or 12, but around twice as many experiments were conducted on age 12 (n = 195) compared to age 11 nestlings (n = 98). In addition, nests with age 11 nestlings were moderately less likely to produce begging calls (48% of nests) compared to nests with age 12 nestlings (59% of nests). As a result, the observed increase in song discrimination with age (Figures 3B and 3C; Table S3) could be partially explained by sample size

and begging rate differences across age categories. To evaluate this possibility, we bootstrapped our findings by generating 100 random sub-samples of experiments at nests with age 12 nestlings. Each sub-sample consisted of 80 nests ([begging rate at age 11 nests]/[begging rate at age 12 nests] x [sample size of age 11 nests]), randomly selected with replacement. We applied identical zero-inflated quasi-poisson regression models to all 100 datasets and evaluated the difference in song discrimination between age 11 and age 12 as for the entire dataset. The effect of age on song discrimination was evaluated as the proportion of model runs with a significant difference in discrimination between ages 11 and 12.