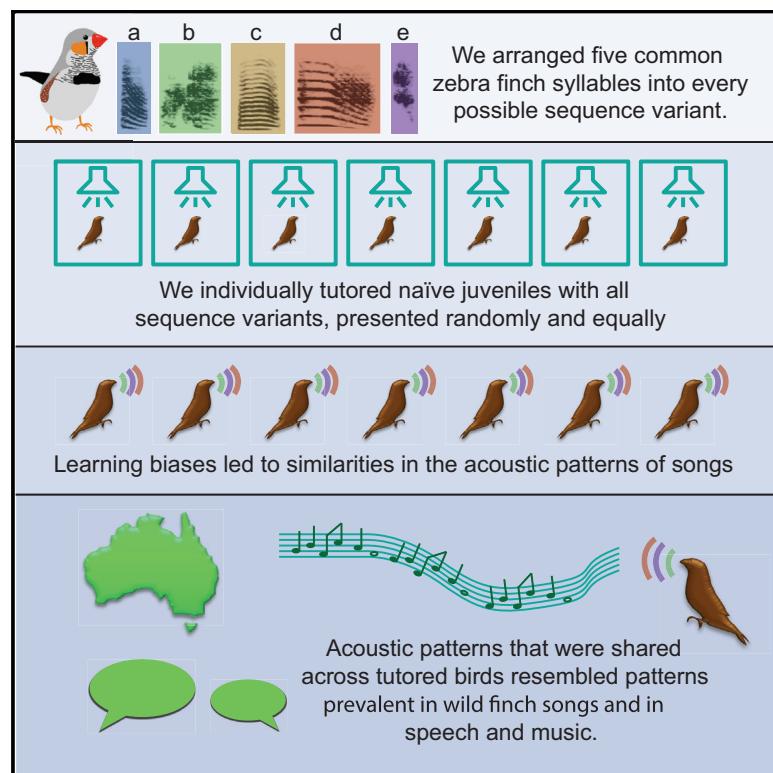


Current Biology

Learning Biases Underlie “Universals” in Avian Vocal Sequencing

Graphical Abstract



Authors

Logan S. James, Jon T. Sakata

Correspondence

logansmithjames@gmail.com (L.S.J.),
jon.sakata@mcgill.ca (J.T.S.)

In Brief

James and Sakata reveal that naive zebra finches individually tutored with randomized acoustic sequences produce convergent acoustic patterns that are non-random and similar to those observed in wild finch song and in speech and music. Their data demonstrate that learning biases contribute to commonalities in acoustic patterning.

Highlights

- Finches were tutored with randomized sequences to reveal biases in sequence learning
- Tutored birds converged onto similar acoustic sequences
- Convergent patterns were similar to those commonly produced by wild birds
- Convergent patterns were also similar to those prevalent in speech and music

Learning Biases Underlie “Universals” in Avian Vocal Sequencing

Logan S. James^{1,*} and Jon T. Sakata^{1,2,3,*}

¹Department of Biology, McGill University, Montreal, QC H3A 1B1, Canada

²Centre for Research on Brain, Language, and Music, McGill University, Montreal, QC H3G 2A8, Canada

³Lead Contact

*Correspondence: logansmithjames@gmail.com (L.S.J.), jon.sakata@mcgill.ca (J.T.S.)

<https://doi.org/10.1016/j.cub.2017.10.019>

SUMMARY

Biological predispositions in vocal learning have been proposed to underlie commonalities in vocal sequences, including for speech and birdsong, but cultural propagation could also account for such commonalities [1–4]. Songbirds such as the zebra finch learn the sequencing of their acoustic elements (“syllables”) during development [5–8]. Zebra finches are not constrained to learn a specific sequence of syllables, but significant consistencies in the positioning and sequencing of syllables have been observed between individuals within populations and between populations [8–10]. To reveal biological predispositions in vocal sequence learning, we individually tutored juvenile zebra finches with randomized and unbiased sequences of syllables and analyzed the extent to which birds produced common sequences. In support of biological predispositions, birds tutored with randomized sequences produced songs with striking similarities. Birds preferentially started and ended their song sequence with particular syllables, consistently positioned shorter and higher frequency syllables in the middle of their song, and sequenced their syllables such that pitch alternated across adjacent syllables. These patterns are reminiscent of those observed in normally tutored birds, suggesting that birds “creolize” aberrant sequence inputs to produce normal sequence outputs. Similar patterns were also observed for syllables that were not used for tutoring (i.e., unlearned syllables), suggesting that motor biases could contribute to sequence learning biases. Furthermore, zebra finches spontaneously produced acoustic patterns that are commonly observed in speech and music, suggesting that sensorimotor processes that are shared across a wide range of vertebrates could underlie these patterns in humans.

RESULTS AND DISCUSSION

Songbirds such as the zebra finch offer a powerful opportunity to conduct controlled experiments of vocal sequence learning and

uncover biological predispositions in vocal learning. This is because songbirds, like humans, learn the sequencing of acoustic elements (“syllables”) during development [5–8, 11]. Furthermore, while species like the zebra finch can learn to produce a wide range of vocal patterns, there exists some evidence that particular acoustic patterns and sequences of syllables are more common than expected [9, 10]. However, experimental tutoring is required to determine the contribution of learning predispositions (independent from cultural forces) to the production of common vocal patterns.

To reveal biases in vocal sequence learning, we tutored naive juvenile zebra finches (i.e., reared without song exposure during their critical period for song learning) with randomized sequences of syllables that lacked sequencing consistency and then assessed the degree to which tutored individuals developed songs with similar syllable sequences (see STAR Methods). We tutored juveniles with synthesized songs consisting of five species-typical syllables (labeled as “a”–“e”) arranged in every possible five-syllable sequence ($n = 120$ possible sequences; Figures 1A–1C). Importantly, each bird heard the 120 possible sequences in equal proportions (Figure 1C), and the order of sequence presentations was randomized across individuals and days of tutoring. Consequently, the tutoring stimuli provided no bias or predictability in syllable sequencing (e.g., syllable position within a sequence or directionality of syllable transitions). If biological predispositions bias the acquisition of vocal sequences, we should observe significant consistencies in syllable positioning or sequencing across tutored birds, despite being tutored by stimuli that provided no such consistencies.

In support of this hypothesis, experimentally tutored birds produced songs with significant similarities in the positioning and sequencing of syllables. Despite being tutored with randomized sequences, birds demonstrated significant learning of the acoustic structure of syllables (see STAR Methods), and like normally tutored birds, most experimentally tutored birds (45 out of 51 birds) produced one dominant sequence of syllables (“motif”; Figures 2A, S1, S2, and S3). We first investigated similarities in how tutored birds positioned syllables within their motif (i.e., beginning, middle, or end). The positioning of three of the five tutored syllables significantly deviated from chance. For example, of the birds that produced the “d” syllable in their motif, 4% produced the “d” at the beginning, 46% produced the “d” in the middle, and 50% produced the “d” at the end. This skewed distribution of the “d” syllable across positions was significantly different than the distribution expected by chance ($p = 0.0328$; as

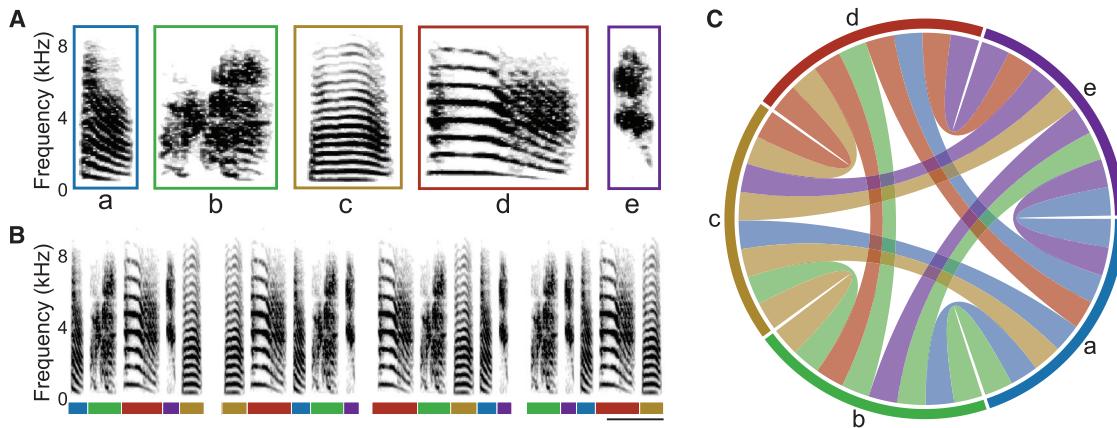


Figure 1. Experimental Tutoring with Randomized and Unbiased Sequences of Syllables

(A) Spectrograms of the five zebra finch “syllables” (i.e., contiguous epochs of sound) in the tutor stimulus set (arbitrarily labeled “a,” “b,” “c,” “d,” and “e”). These five syllables are acoustically distinct from each other and are types commonly observed in zebra finch songs.
 (B) Example of tutoring song bout. Five zebra finch syllable exemplars were arranged into 120 possible five-syllable motif variants in which each syllable appears only once. Four distinct motif variants were presented in each song bout. Scale bar represents 250 ms.
 (C) A circle plot summarizing transition probabilities between syllables within the motifs of the stimulus set. Outer-ring colors represent the five syllables, and line colors represent transitions from that syllable to others (e.g., blue lines represent transitions from “a” to other syllables). The width of the line is proportional to transition probability. All lines for the stimulus set have the same width because all transitions were heard with equal probabilities.

determined using Monte Carlo simulations; see [STAR Methods](#)) and was driven by the facts that the “d” was produced more than expected at the end of the motif and less than expected at the beginning ($p < 0.02$ for each; observed likelihoods in each position were, respectively, 3.43σ and -2.30σ away from the chance value; [Figure 2B](#)). The distributions of the syllables “e” and “b” also significantly differed from chance ([Figure 2B](#)): the syllable “e” was significantly more likely to be produced in the middle of the motif than expected by chance, whereas the syllable “b” was significantly more likely to be produced at the beginning of the motif. Consequently, despite being tutored with stimuli in which each syllable was equally likely to occupy each position, birds preferentially produced syllables in particular positions within the motif.

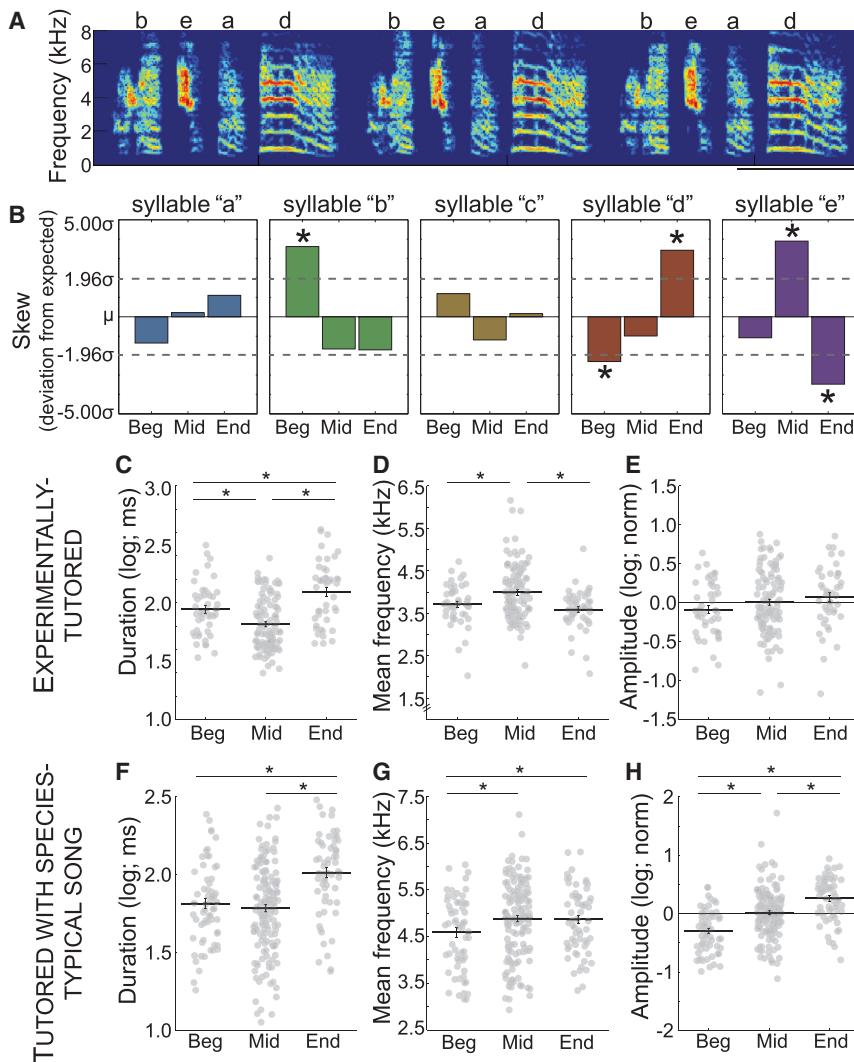
Importantly, these patterns of syllable positioning resemble those observed among populations of normally reared zebra finches [9, 10]. Across wild populations of zebra finches, birds frequently terminate their motif with distance calls (our syllable “d”) and produce a high-pitched syllable (our syllable “e”) in the middle [10]. As such, our data strongly suggest that the biological predispositions in vocal sequence learning could account for natural commonalities in vocal positioning.

Positional variation in the duration, pitch, and amplitude of sounds is prevalent across speech and music [12–15]. For example, utterances in the final position of phrases tend to be longer and lower in pitch than utterances in medial positions, and music phrases tend to have arch-like pitch contours such that notes in the beginning and final positions tend to be lower in pitch. We assessed the degree to which our experimentally tutored birds organized the syllables in their songs with such positional variation. Consistent with these patterns, syllables at the end of the motif were significantly longer than syllables in the middle or beginning of the motif ($p < 0.02$ for each contrast; [Figure 2C](#)). Syllables at the beginning of the motif were also significantly longer than syllables in the middle of the motif ($p = 0.007$). Pitch

(mean frequency) also varied across motif position, with beginning and end syllables being lower in pitch than middle syllables ($p < 0.04$ for each; [Figure 2D](#)). Amplitude did not significantly vary across motif positions ($p = 0.2478$) but generally demonstrated an ascending pattern from beginning to end ([Figure 2E](#)).

To assess whether similar positional variation is observed in the songs of zebra finches that were tutored with species-typical songs, we applied the same analyses to zebra finch songs from an online repository ($n = 61$; see [STAR Methods](#)). We found patterns of positional variation that resembled those found in our experimentally tutored birds. For example, syllable duration significantly varied across position, with end syllables being significantly longer than syllables in the beginning or middle of the motif ($p < 0.0001$ for each contrast; [Figure 2F](#)). Pitch also significantly varied across motif positions ([Figure 2G](#)), and, similar to the songs of experimentally tutored birds, beginning syllables were lower in pitch than middle syllables ($p = 0.0006$). While end syllables were not significantly different in pitch than middle syllables, end syllables were higher in pitch than beginning syllables ($p = 0.0043$). Mean amplitude significantly varied across motif position within the songs of finches tutored with species-typical songs, with end syllables being the loudest, followed by middle syllables and then beginning syllables ($p < 0.0003$ for each contrast; [Figure 2H](#)). A similar, albeit statistically non-significant, pattern was observed in the songs of experimentally tutored birds. Because patterns of positional variation in acoustic features were similar between birds tutored with randomized sequences and birds tutored with normal zebra finch song, this suggests that experimentally tutored birds spontaneously organized their syllables into patterns typically found in zebra finch song.

In addition to the positioning of syllables, we also analyzed the degree to which the sequencing of syllables within the songs of experimentally tutored birds followed predictable patterns. We first analyzed predictability (i.e., asymmetries) in the directionality of syllable transitions within the 10 pairs of tutored syllables



(F-H) The acoustic structure of syllables also significantly varied across motif positions in the songs of birds tutored with species-typical song ($n = 267$ syllables; see STAR Methods). The duration of syllables significantly varied across positions (F), with syllables at the end of the motif being longer than syllables in the middle or at the beginning of the motif ($p < 0.0001$ for each). The pitch of syllables varied by position within the motif (G), with syllables at the beginning of the motif being lower in pitch than syllables in the middle or at the end of the motif ($p < 0.005$ for each). The mean amplitude of syllables significantly varied across positions in the motif (H), with syllables at the end of the motif being louder than syllables in the middle or at the beginning of the motif ($p < 0.003$ for each). Plotted are mean-subtracted amplitude values (see E).

(C-H) Dots represent median values for each syllable, horizontal lines indicate the means of the distributions, and vertical lines indicate standard errors of the distribution. For all panels, the asterisk (*) denotes $p < 0.05$. See also Figures S1–S4 and Table S1.

(e.g., transitioning from "a" to "e" versus from "e" to "a"; Figures 3A–3C). Because birds were tutored with sequences that were symmetric in the directionality of syllable transitions, the null model (i.e., no bias in sequence learning) predicts equal probabilities of transitions in each direction. However, the observed pairwise transitions were significantly asymmetric ($p = 0.0312$; Figure 3D; see STAR Methods), indicating that birds produced similar sequences. One striking result that exemplifies this predictability was that four birds produced the exact same four-syllable sequence, "bead" (Figure 2A). The likelihood that the same four-syllable sequence is produced by four or more birds, by chance, is $<0.2\%$ (Monte Carlo simulation), underscoring that syllable sequencing is non-random and that biological predispositions contribute to vocal sequence learning.

Speech sounds are commonly sequenced in alternating patterns of stressed (generally higher in pitch, louder, and longer) and unstressed syllables [16–18]. However, the extents to which songbirds produce sequences with alternating acoustic patterns and to which biological predispositions in vocal learning contribute to this patterning remain largely unknown. To this end, we first analyzed a large corpus of songs of zebra finches that were tutored with species-typical songs (see STAR Methods). While we did not observe significant patterning of duration ($p = 0.2096$) or amplitude ($p = 0.0838$), birds tutored with normal song alternated the mean frequency of syllables in their motif more than expected by chance (Monte Carlo simulation; $p = 0.0223$; Figures 4A and 4B). To reveal the contribution of learning predispositions to this patterning, we assessed the

Figure 2. Skew in the Positioning of Syllables and Positional Variation in Acoustic Features

(A) Representative spectrogram of an experimentally tutored bird's song, which was composed of the motif "bead." Scale bar represents 250 ms.

(B) Bar graphs depicting the skew in the likelihood that individual syllables were positioned at the beginning ("Beg"), middle ("Mid"), or end ("End") of the motif (see STAR Methods). Skew was computed as the standardized difference between the observed number of times a syllable occupied a particular position and the number expected by chance (across 10,000 iterations of the Monte Carlo simulation). Skew is plotted for all five syllables across these three positions, and the significance of the skew was computed using Monte Carlo simulations ($p < 0.05$). Dashed lines indicate ± 1.96 standard deviations ($p = 0.05$) from the expected (chance) probability (mean of the permutations; solid line). Tutored birds were significantly more likely to position the "b" syllable at the beginning of the motif ($p = 0.0009$), the "d" syllable in the middle ($p < 0.0001$), and the "d" syllable at the end ($p = 0.0014$). Birds were also significantly less likely to produce the "d" syllable at the beginning ($p = 0.0126$) and the "e" syllable at the end of the motif ($p = 0.0001$).

(C-E) The acoustic structure of syllables significantly differed across motif positions for experimentally tutored birds (beginning, middle, or end; $n = 200$ syllables). Syllables at the beginning or end of the motif were significantly longer in duration (C) and lower in pitch (D; mean frequency) than syllables produced in the middle ($p < 0.02$ for each). Syllables at the end of the motif were also significantly longer than syllables at the beginning of the motif ($p < 0.01$). Syllables did not significantly vary in amplitude across positions (E). Amplitude values for each bird are normalized ("norm"; mean subtracted) by the average amplitude of all syllables within his song.

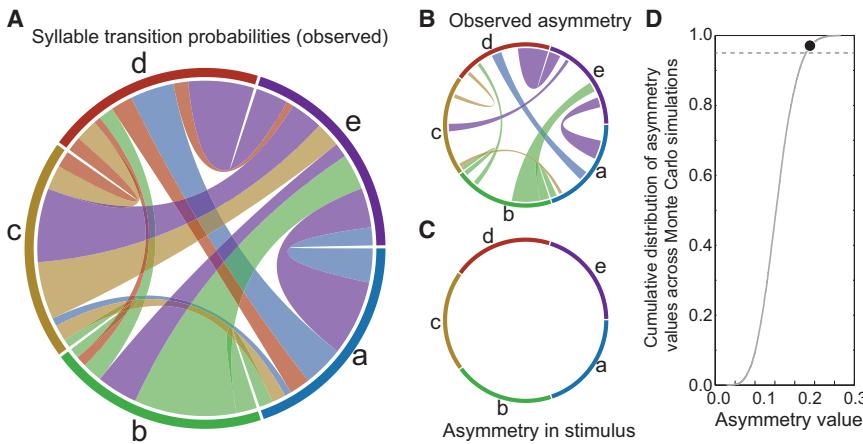


Figure 3. Significant Patterning of Syllable Transitions Following Tutoring with Randomized Sequences

(A) Circle plot depicts the syllable transition probabilities within motifs of tutored birds (see Figure 1C for tutoring stimulus).

(B and C) Asymmetries in the directionality of syllable transitions (e.g., difference in transition probabilities from “a” to “e” and from “e” to “a”) within the motifs of tutored birds (B) and tutoring stimulus (C). Line thickness scales with the magnitude of asymmetry in the directionality of syllable transitions (e.g., thick purple line between “e” and “a” in B represents a bias for transitions from “e” to “a” over “a” to “e”).

(D) Overall, there was a significant asymmetry in the directionality of syllable transitions ($p = 0.0312$; Monte Carlo simulations). Solid line

depicts the cumulative distribution of overall asymmetry values found in the simulations. The dashed line represents the 95th percentile from the simulations, and the dot indicates the overall asymmetry value observed across experimentally tutored birds.

degree to which birds tutored with randomized sequences similarly produced songs with alternating patterns. We found that experimentally tutored birds significantly alternated the pitch ($p = 0.0018$), but not duration or ($p = 0.9311$) amplitude ($p = 0.2570$), of syllables in their motif (Monte Carlo simulation; Figure 4C). These data strongly suggest that learning biases could contribute to the typical alternation of pitch in zebra finch song.

Motor and/or sensory biases could underlie such biological predispositions in learning. For example, the organization of the motor system has been hypothesized to influence and constrain the acquisition and production of acoustic patterns within birdsong, speech, and music [15, 19–22]. To assess the degree to which motor biases contribute to biological predispositions in vocal sequence learning, we analyzed positional variation in the acoustic features of syllables that experimentally tutored birds produced and that were not heard in the tutor stimulus set (“novel syllables”; see STAR Methods; Figures S2 and S3). As observed in the previous analyses of experimentally or normally tutored birds (Figure 2), novel syllables that were produced at the end of the motif were longer than those produced in the middle or beginning of the motif, and novel syllables at the beginning or end of the motif were lower in pitch than those in the middle (Figure S4). Because positional variation in utterances was observed even for vocalizations that were not directly heard during tutoring, these data suggest that motor biases contribute to the acoustic patterns commonly observed in our birds. It is possible, for example, that such acoustic patterns are easier for the motor system to produce [22].

Despite these findings, it remains possible that biases in auditory processing and memory could also contribute to the common vocal patterns. Just as previous studies demonstrate “innate” tuning for species-typical acoustic features [5, 23–25], our data suggest that the auditory system of juvenile zebra finches could preferentially process and learn particular types of acoustic sequences. Recent studies on prosodic processing in zebra finches support this idea. When asked to categorize patterns of human speech sounds or zebra finch syllables, zebra finches preferentially use pitch cues (relative to duration or amplitude) to infer patterns [26, 27]. Interestingly, birds tutored

with randomized sequences spontaneously arranged their syllables such that pitch but not amplitude or duration alternated, highlighting a parallel in the prominence of pitch to production and perception in zebra finches. Similar parallels in production and processing are observed in humans, as speakers of languages in which pitch is used as a primary prosodic cue tend to use pitch cues for speech segmentation [28, 29]. In addition, these data suggest the intriguing possibility that females could prefer males that produce songs with alternation of and positional variation in pitch.

Regardless of the mechanism, we provide compelling experimental evidence that the nervous system is endowed with predispositions that bias vocal sequence learning in a songbird species that lacks specific rules for vocal sequence production [9, 10, 30, 31]. Tutoring birds with random, unbiased sequence inputs led to the production of non-random, biased sequence outputs that were common across individuals, similar to patterns found in birds tutored with species-typical songs, and similar to those commonly observed in speech and music [1, 4, 15, 18, 20, 32]. Consequently, these data indicate that biological predispositions influence not only the learning of acoustic structure [24, 33, 34] but also the acquisition of syllable sequencing. That the patterning of acoustic features is similar between the songs of experimentally and normally tutored birds indicates that juvenile zebra finches “normalize” atypical inputs of syllable sequences into species-typical sequences. This process resembles the “creolization” of human languages and acoustic features of birdsong [30, 33, 35, 36]. Children who are exposed to novel, simplified forms of languages that lack properties commonly observed across established languages reshape and “creolize” this language into one that is more likely to contain common linguistic patterns [35, 36], and juvenile zebra finches that are tutored by conspecifics that produce acoustically atypical songs produce syllables with more normal acoustic structure as adults [33]. It will be important for future studies to assess the extent to which these learning biases can be overcome and to explicitly evaluate the motor and sensory underpinnings of these biases. Given the parallels to patterns in speech and music, our findings suggest that sensorimotor processes

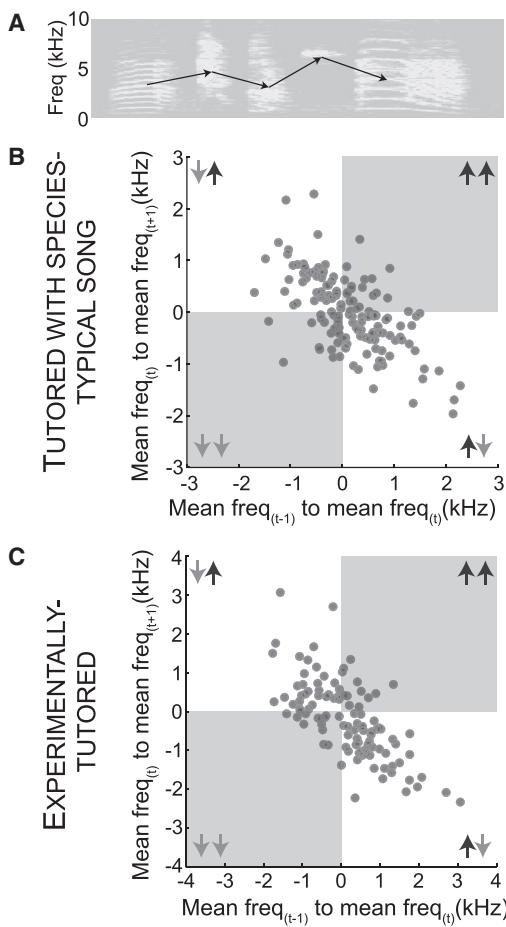


Figure 4. Mean Frequency Significantly Alternates across Syllables within Zebra Finch Motifs

(A) A bird's motif that demonstrates alternation across syllables. Arrows depict the change in mean frequency across adjacent syllables.
 (B) The mean frequency of syllables significantly alternated within the motifs of birds tutored with species-typical song. Plotted is the change in mean frequency (in kHz) from the previous syllable_(t-1) to the current syllable_(t) (x axis) and the change from the current syllable_(t) to the following syllable_(t+1) (y axis). Arrows (in the corners) indicate the direction of change, with the first and second arrows of the pair indicating, respectively, change from syllable_(t-1) to syllable_(t) and change from syllable_(t) to syllable_(t+1). The significant negative relationship indicates that pitch alternated across syllables more than expected by chance.
 (C) Birds tutored with randomized sequences of syllables also significantly alternated the mean frequency of syllables across the motif ($r^2 = 0.476$; $p = 0.0018$; Monte Carlo simulations). Data for experimentally tutored birds are plotted the same way in (B).

that are shared across a wide range of vertebrates underlie “universals” in phonological, prosodic, and musical patterns in humans [19, 20, 37, 38].

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING

● EXPERIMENTAL MODEL AND SUBJECT DETAILS

● METHOD DETAILS

- Song Tutoring
- Song Analysis
- Comparisons with Zebra Finches Tutored with Normal Songs

● QUANTIFICATION AND STATISTICAL ANALYSIS

- Monte Carlo Simulations
- Analysis of Variation in Acoustic Features across Positions
- Statistical Analyses

● DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and one table and can be found with this article online at <https://doi.org/10.1016/j.cub.2017.10.019>.

AUTHOR CONTRIBUTIONS

L.S.J. helped design the experiment, collect and analyze data, and write the manuscript. J.T.S. helped design the experiment, analyze the data, and write the manuscript.

ACKNOWLEDGMENTS

We thank K. Onishi, M. Kao, R. Krahe, and S.C. Woolley for constructive input and feedback throughout the experiment and H. Goad and L. White for formative discussions about speech and language patterns. We also thank L. Matheson, H. Sun, J. Dai, Y. Chen, and G. Lafon for help with data analysis and manuscript preparation; T. Gardner for collecting and making available the large database of zebra finch songs; and three anonymous reviewers for their helpful feedback on the manuscript. Research was supported by funding from the Natural Sciences and Engineering Research Council (J.T.S.; #05016), the Centre for Research on Brain, Language, and Music, and an award from the Heller Family Fellowship (L.S.J.).

Received: July 26, 2017

Revised: September 11, 2017

Accepted: October 6, 2017

Published: November 22, 2017

REFERENCES

1. Chomsky, N. (1980). *Rules and Representations* (Blackwell).
2. Bickel, B. (2013). Distributional biases in language families. In *Language Typology and Historical Contingency*, B. Bickel, L.A. Grenoble, D.A. Peterson, and A. Timberlake, eds. (Benjamins), pp. 415–444.
3. Greenberg, J. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, J. Greenberg, ed. (MIT Press), pp. 73–113.
4. Yang, C., Crain, S., Berwick, R.C., Chomsky, N., and Bolhuis, J.J. (2017). The growth of language: Universal Grammar, experience, and principles of computation. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2016.12.023>.
5. Doupe, A.J., and Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* 22, 567–631.
6. Brainard, M.S., and Doupe, A.J. (2013). Translating birdsong: songbirds as a model for basic and applied medical research. *Annu. Rev. Neurosci.* 36, 489–517.
7. Prather, J., Okanoya, K., and Bolhuis, J.J. (2017). Brains for birds and babies: neural parallels between birdsong and speech acquisition. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2016.12.035>.

8. Lipkind, D., Marcus, G.F., Bemis, D.K., Sasahara, K., Jacoby, N., Takahasi, M., Suzuki, K., Feher, O., Ravbar, P., Okano, K., and Tchernichovski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature* **498**, 104–108.
9. Lachlan, R.F., van Heijningen, C.A.A., Ter Haar, S.M., and Ten Cate, C. (2016). Zebra finch song phonology and syntactical structure across populations and continents—a computational comparison. *Front. Psychol.* **7**, 980.
10. Zann, R.A. (1996). *The Zebra Finch: A Synthesis of Field and Laboratory Studies* (Oxford University Press).
11. Chen, Y., Matheson, L.E., and Sakata, J.T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci. USA* **113**, 6641–6646.
12. Oller, D.K. (1973). The effect of position in utterance on speech segment duration in English. *J. Acoust. Soc. Am.* **54**, 1235–1247.
13. Fougeron, C., and Keating, P.A. (1997). Articulatory strengthening at edges of prosodic domains. *J. Acoust. Soc. Am.* **101**, 3728–3740.
14. Vaissière, J. (1983). Language-independent prosodic features. In *Prosody: Models and Measurements*, A. Cutler, and D.R. Ladd, eds. (Springer), pp. 53–66.
15. Patel, A. (2010). *Music, Language, and the Brain* (Oxford University Press).
16. Liberman, M., and Prince, A. (1977). On stress and linguistic rhythm. *Linguist. Inq.* **8**, 249–336.
17. Langus, A., Mehler, J., and Nespor, M. (2016). Rhythm in language acquisition. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2016.12.012>.
18. Goedemans, R., and van der Hulst, H. (2013). Rhythm types. In *The World Atlas of Language Structures Online*, M.S. Dryer, and M. Haspelmath, eds. (Max Planck Institute for Evolutionary Anthropology).
19. Zollinger, S.A., and Suthers, R.A. (2004). Motor mechanisms of a vocal mimic: implications for birdsong production. *Proc. Biol. Sci.* **271**, 483–491.
20. Tierney, A.T., Russo, F.A., and Patel, A.D. (2011). The motor origins of human and avian song structure. *Proc. Natl. Acad. Sci. USA* **108**, 15510–15515.
21. Podos, J. (1996). Motor constraints on vocal development in a songbird. *Anim. Behav.* **51**, 1061–1070.
22. MacNeilage, P.F., Davis, B.L., Kinney, A., and Matyear, C.L. (2000). The motor core of speech: a comparison of serial organization patterns in infants and languages. *Child Dev.* **71**, 153–163.
23. Araki, M., Bandi, M.M., and Yazaki-Sugiyama, Y. (2016). Mind the gap: neural coding of species identity in birdsong prosody. *Science* **354**, 1282–1287.
24. Marler, P., and Peters, S. (1977). Selective vocal learning in a sparrow. *Science* **198**, 519–521.
25. Gómez, D.M., Berent, I., Benavides-Varela, S., Bion, R.A.H., Cattarossi, L., Nespor, M., and Mehler, J. (2014). Language universals at birth. *Proc. Natl. Acad. Sci. USA* **111**, 5837–5841.
26. Spierings, M.J., and ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proc. Biol. Sci.* **281**, 20140480.
27. Spierings, M., Hubert, J., and Ten Cate, C. (2017). Selective auditory grouping by zebra finches: testing the iambic-trochaic law. *Anim. Cogn.* **20**, 665–675.
28. Bion, R.A.H., Benavides-Varela, S., and Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Lang. Speech* **54**, 123–140.
29. Jusczyk, P.W., Houston, D.M., and Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognit. Psychol.* **39**, 159–207.
30. Gardner, T.J., Naef, F., and Nottebohm, F. (2005). Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* **308**, 1046–1049.
31. Plamondon, S.L., Goller, F., and Rose, G.J. (2008). Tutor model syntax influences the syntactical and phonological structure of crystallized songs of white-crowned sparrows. *Anim. Behav.* **76**, 1815–1827.
32. Savage, P.E., Brown, S., Sakai, E., and Currie, T.E. (2015). Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci. USA* **112**, 8987–8992.
33. Fehér, O., Wang, H., Saar, S., Mitra, P.P., and Tchernichovski, O. (2009). De novo establishment of wild-type song culture in the zebra finch. *Nature* **459**, 564–568.
34. 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.
35. Senghas, A., and Coppola, M. (2001). Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychol. Sci.* **12**, 323–328.
36. Thomason, S., and Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics* (University of California Press).
37. Mol, C., Chen, A., Kager, R.W.J., and Ter Haar, S.M. (2017). Prosody in birdsong: a review and perspective. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2017.02.016>.
38. Liberman, A.M., and Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition* **21**, 1–36.
39. Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B., and Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176.
40. Roper, A., and Zann, R. (2006). The onset of song learning and song tutor selection in fledgling zebra finches. *Ethology* **112**, 458–470.
41. Brainard, M.S., and Doupe, A.J. (2002). What songbirds teach us about learning. *Nature* **417**, 351–358.
42. Dunn, A.M., and Zann, R.A. (1996). Undirected song in wild zebra finch flocks: contexts and effects of mate removal. *Ethology* **102**, 529–539.
43. Volman, S.F., and Khanna, H. (1995). Convergence of untutored song in group-reared zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **109**, 211–221.
44. Tchernichovski, O., and Nottebohm, F. (1998). Social inhibition of song imitation among sibling male zebra finches. *Proc. Natl. Acad. Sci. USA* **95**, 8951–8956.
45. Derégnaucourt, S., and Gahr, M. (2013). Horizontal transmission of the father's song in the zebra finch (*Taeniopygia guttata*). *Biol. Lett.* **9**, 20130247.
46. Tchernichovski, O., Lints, T., Mitra, P.P., and Nottebohm, F. (1999). Vocal imitation in zebra finches is inversely related to model abundance. *Proc. Natl. Acad. Sci. USA* **96**, 12901–12904.
47. Tchernichovski, O., Mitra, P.P., Lints, T., and Nottebohm, F. (2001). Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* **291**, 2564–2569.
48. Clayton, N.S. (1989). The effects of cross-fostering on selective song learning in Estrildid Finches. *Behaviour* **109**, 163–174.
49. Sturdy, C.B., Phillmore, L.S., and Weisman, R.G. (1999). Note types, harmonic structure, and note order in the songs of zebra finches (*Taeniopygia guttata*). *J. Comp. Psychol.* **113**, 194–203.
50. Ten Cate, C. (1991). Behaviour-contingent exposure to taped song and zebra finch song learning. *Anim. Behav.* **42**, 857–859.
51. Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Anim. Behav.* **46**, 149–159.
52. Mooney, R. (2009). Neural mechanisms for learned birdsong. *Learn. Mem.* **16**, 655–669.
53. Scharff, C., and Nottebohm, F. (1991). A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* **11**, 2896–2913.

54. Vu, E.T., Mazurek, M.E., and Kuo, Y.C. (1994). Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14**, 6924–6934.
55. Glaze, C.M., and Troyer, T.W. (2006). Temporal structure in zebra finch song: implications for motor coding. *J. Neurosci.* **26**, 991–1005.
56. Derégnaucourt, S., Poirier, C., Kant, A.V., Linden, A.V., and Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* **107**, 210–218.
57. Stepanek, L., and Doupe, A.J. (2010). Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J. Neurophysiol.* **104**, 2474–2486.
58. Sakata, J.T., and Brainard, M.S. (2006). Real-time contributions of auditory feedback to avian vocal motor control. *J. Neurosci.* **26**, 9619–9628.
59. Manly, B. (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology* (Chapman & Hall).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Excel file with sequence information and acoustic measures for all birds in the analyses	Mendeley Data	https://doi.org/10.17632/sywzmd483f.1
Experimental Models: Organisms/Strains		
Zebra finch <i>Taeniopygia guttata</i>	Exotic Wings and Pet Things: Ontario, Canada as well as lab reared	N/A
Software and Algorithms		
MATLAB	MathWorks	https://www.mathworks.com/
JMP	SAS	https://www.jmp.com/
Sound Analysis Pro	[39]	http://soundanalysispro.com/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Jon T. Sakata (jon.sakata@mcgill.ca).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Adult zebra finches for breeding were obtained from commercial breeders (Exotic Wings and Things, Ontario, Canada). Naive juvenile zebra finches were bred and raised in our lab in custom-built sound-attenuating chambers (TRA Acoustics, Ontario, Canada). Fathers were removed from the nest ≤ 5 days post-hatch (dph), and, thereafter, experimental juveniles were raised solely by their mother or a foster mother (females do not produce song) [10, 33]. The critical period for song learning for zebra finches opens ~ 20 dph; thus, song exposure before 5 dph does not lead to significant song learning [33, 40–42]. After reaching nutritional independence (~ 35 –50 dph), juvenile males were individually housed in sound-attenuating chambers for tutoring. Juveniles were individually housed because social interactions among juveniles can influence song development [43–45]. Animals used in this experiment were not part of any other experiment prior to tutoring. All birds were housed on a 14:10 light:dark cycle and provided with food and water *ad libitum*. All procedures were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

METHOD DETAILS

Song Tutoring

Birds were tutored using operantly triggered playbacks of synthesized song stimuli for >30 days [46, 47]. Song stimuli were synthesized to resemble natural zebra finch song, albeit with no sequence consistency across renditions. To create song stimuli for tutoring, five typical zebra finch song elements (termed “syllables,” contiguous epochs of sound separated by short silent gaps (>5 ms)) were selected from the songs of different males from multiple lab populations. These five syllables corresponded to syllable types most commonly observed in natural populations of zebra finches, four of which were termed the “primary elements” of zebra finch phrases [10]. Using the nomenclature of Zann (1996) [10], these syllables can be described as a short sweep (“a”), a noise-noise syllable (“b”), a short stack (“c”), a distance call (“d”), and a high note (“e”; Figure 1). To create tutoring stimuli, syllables were band-pass-filtered (0.3–10 kHz), normalized by their maximum amplitude, and ramped to avoid sharp onsets and offsets in amplitude.

We tutored juveniles with stimuli that contained randomized and unbiased syllable sequences to assess biological predispositions in sequence learning ($n = 51$). To this end, we constructed all possible five-syllable sequence (“motif”) variants that contained each of the syllables only once ($n = 120$ motif variants; Figure 1). We created motif variants in this manner because individual syllables are rarely produced more than once in a zebra finch’s motif [33, 48, 49]. Because zebra finches typically produce bouts of song that consists of multiple renditions of a motif, we synthesized song stimuli such that each song bout contained four motifs. However, in contrast to normal zebra finch song bouts in which the same motif is repeated, synthesized song bouts consisted of four distinct motif variants. The sequence of motif variants per song bout was randomized across days for each bird and across birds. Moreover, to ensure that each motif variant was heard equally and randomly, we presented the stimuli in blocks of 120 motifs so that every motif variant was heard once before any variant was repeated.

The timing of syllables and motif variants in song stimuli followed a species-typical pattern: syllables within the motif were each separated by a silent gap of 30 ms, whereas the gap between syllables across adjacent motifs (i.e., from offset of last syllable in the motif to onset of first syllable of subsequent motif) was 100 ms.

Birds were operantly tutored using perch hops [50, 51], and each perch hop triggered the playback of one song bout. To improve learning outcomes, song playbacks were spaced out such that juveniles were allowed to hear only 10 operantly triggered song playbacks (i.e., 40 motif variants) within each of three time periods in the day (morning, noon, and afternoon [46]), for a maximum of 120 motifs (one of each variant) per day. Birds were housed individually throughout tutoring and until 4 months of age.

Sound Analysis Pro 2011 (SAP; <http://soundanalysispro.com>) was used for song tutoring. Stimuli were played out of an Avantone Pro Mixcube speaker (Avantone, NY) connected to a Crown XLS 1000 amplifier (Crown Audio, IN). Song playback was operantly triggered by perch hops using custom-built perches connected to a National Instruments PCI-6503 I/O card (National Instruments, TX).

Song Analysis

We recorded the spontaneously produced songs of tutored birds using an omnidirectional microphone (Countryman Associates, Menlo Park, CA) and an amplitude-based recording system (SAP; 44.1 kHz; bandpass-filtered from 0.3–10 kHz). Songs were visualized and analyzed offline using custom-written software in MATLAB (Mathworks, Natick, MA). We analyzed the adult songs ($n > 30$ songs/bird) of tutored birds when they were 4 months old, an age at which song is typically stable and stereotyped [41, 52]. We identified and labeled syllables and motifs following automated, amplitude-based segmentation of audio files. As indicated above, zebra finch songs consist of syllables that are arranged into stereotyped sequences called “motifs,” and each adult zebra finch repeats a single motif within his song. An individual’s motif is readily identifiable because it is repeated multiple times within a song bout [10, 41, 52, 53] and because the duration of gaps between motifs is usually longer and more variable than gap durations within the motif [54, 55]. To identify motifs, authors independently examined multiple renditions of each bird’s song. Both researchers identified and agreed upon a stereotyped motif for 45 of the 51 tutored birds, and the songs of these birds resembled the songs of normally reared zebra finches. The remaining 6 tutored birds either produced songs with sufficient sequence variability to make motif identification difficult or produced songs with multiple motif variants (i.e., songs in which authors could not agree upon a motif; see **Figure S1** for examples). Such variability is not unexpected since operant tutoring can lead to aberrant learning outcomes in a subset of birds (e.g., [8, 56]). The songs of these six birds with more variable sequencing were not analyzed because of ambiguity in demarcating the beginning and end of the motif.

The 45 birds used in the analysis produced motifs with, on average, 4.4 ± 0.2 syllables (mean \pm SEM; range: 2–8), with average gap durations (i.e., intervals between syllables within the motif) of 27.1 ± 1.2 ms. Thirty-seven of the 45 birds (82%) produced a direct transition from the end of one motif to the beginning of the subsequent motif, and the average inter-motif interval (duration from the offset of the last syllable in the motif to the onset of the first syllable in the subsequent motif) for these birds was 77.5 ± 7.5 ms.

After identifying motifs, we labeled each syllable within the motif using quantitative and qualitative approaches. We first identified the median rendition of each distinct syllable of a bird’s motif. To identify the median syllable rendition, we used MATLAB scripts to quantify acoustic features of every rendition of each syllable in the bird’s motif across >30 randomly selected songs (duration, mean frequency, spectral density, spectral temporal entropy, and loudness entropy; [57, 58]), converted these feature values into z-scores for each syllable in each bird’s motif, and then identified the individual syllable rendition that was closest to the median centroid of the distributions. Once the median rendition for each syllable was identified, five individuals with expertise studying birdsong were presented with individual spectrograms of each median syllable rendition and asked to determine whether the syllable matched those in the tutoring stimulus set (syllables “a” through “e”) or should be considered distinct from the five tutor syllables (i.e., “novel”). Experts were blind to the identity of the bird and the position of the syllable within the motif (i.e., raters categorized individual syllables that were spliced out from the motif). The label chosen by at least three of the five raters was selected as the final label for the syllable. There was a single syllable in which there was no consensus across raters, and this syllable was categorized as a “novel” syllable. On average, each rater’s labels were consistent with the final labels (i.e., labels used in analyses) for 90% of the syllables.

The degree of syllable learning in tutored birds with stereotyped motifs ($n = 45$) was comparable to that observed in previous studies of experimental song tutoring [10, 11]; on average, 72% of all syllables produced by these birds were categorized as syllables that were learned from the tutor stimulus set. To further assess the extent of learning, we employed quantitative analyses of acoustic similarity using SAP [39]. We measured the spectral similarity of each pupil’s motif to a motif from the tutor stimulus. The SAP “similarity score” does not take into account similarities in the sequencing of syllables between compared motifs; therefore, we randomly picked one of the 120 motifs in the tutor stimulus to use for all comparisons. On average, the motifs of our tutored birds had SAP similarity scores of $55.9 \pm 2.6\%$ (mean \pm SEM), and these scores were significantly and positively correlated with our estimate of the number of syllables learned from the tutor stimulus set ($r = 0.64$, $p < 0.0001$). Furthermore, these similarity scores were comparable to those reported for birds that were operantly tutored in other studies; for example, Derégnaucourt et al. (2013) found that pupils tutored by operant conditioning of natural song stimuli had an average SAP similarity score of $61.1 \pm 3.4\%$ (values extracted from their figure; [56]).

In addition, the songs of our experimentally tutored birds were more similar to the tutor stimulus set than were the songs of zebra finches not tutored with these stimuli. For this analysis, we used SAP to measure the similarity of motifs of zebra finches from five different laboratories to our tutor stimulus set (http://people.bu.edu/timothyg/song_website/index.html; see “Comparisons with zebra finches tutored with normal songs” below for further details). These similarity scores provide an estimate of the extent to which birds could, by chance, resemble the tutor stimulus set. Overall, the motifs of birds not tutored by our tutor stimulus set had similarity scores of $32.4 \pm 1.9\%$, which was significantly lower than the scores of our tutored birds ($t_{104} = 7.39$, $p < 0.0001$). Altogether, these analyses indicate that the extent of learning of our experimentally tutored birds was significantly greater than chance and comparable to the degree of learning observed in related studies (e.g., [56]).

To assess the robustness of the findings reported in the main text, we also conducted the same set of analyses on motifs that were independently identified by two songbird researchers who were unaware of the experiment. These researchers were asked to identify motifs by examining spectrograms of the adult songs of tutored birds. Overall, there was 86% similarity in syllable composition of motifs across the two researchers and 88% similarity to the dataset analyzed in the main text. Moreover, results for these analyses were highly consistent with those reported in the main text, with 89% of analysis yielding $p < 0.05$ or $p < 0.10$ for all significant effects reported in the main text.

Comparisons with Zebra Finches Tutored with Normal Songs

To corroborate patterns that were commonly observed in our experimentally tutored birds, we analyzed the extent of positional variation in and alternation of song features within the songs of zebra finches tutored with normal songs. We examined zebra finch songs from a repository that consisted of songs from five different laboratories (http://people.bu.edu/timothyg/song_website/index.html); songs from our laboratory were not included in the repository. While specific information about how these birds were tutored was not provided, these recordings were obtained from labs that generally raise and tutor birds with typical zebra finch song (i.e., stereotyped sequences of syllables). We independently examined the repository and identified a subset of birds in which we could readily identify a motif. Songs were included in the analysis if (1) the song file contained at least two full motifs (since only one file existed per bird) and (2) we agreed on the exact sequence of syllables that comprised the motif ($n = 61$ birds).

Because these songs were obtained from birds raised in different laboratories, we examined the extent to which positional variation in or alternation of song features varied across labs. There was no significant interaction between lab of origin and motif position or alternation in the dataset ($p > 0.2$ for each), suggesting that the nature and extent of positional variation and alternation were comparable across lab populations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Monte Carlo Simulations

We conducted Monte Carlo simulations to determine the extent to which adult birds tutored with randomized sequences produced similar acoustic patterns within their motifs [59]. These simulations were used to compute the likelihood of observing a particular pattern by chance and were important because individual birds produced different repertoires of syllables. These simulations were also important to account for the number of times each syllable was produced within the motif (i.e., syllables were occasionally produced more than once in a motif). The simulation consisted of 10,000 iterations, and for each iteration we randomized the sequence of syllables within each bird's motif. We then computed the likelihood of observing a particular phenomenon across motifs (e.g., percent of motifs that ended with a particular syllable) during each iteration, and then generated a distribution of these likelihoods across the 10,000 iterations. These Monte Carlo simulations were the basis for generating null distributions against which we assessed the significance of our observations.

Skew in Syllable Position

Monte Carlo simulations were first used to quantify the degree to which the observed skew in syllable positioning within the motif (beginning, middle, and end) deviated from chance. The first syllable of the motif was designated as positioned at the beginning, the final syllable of the motif was designated as positioned at the end, and the remaining syllables were designated as positioned in the middle. We first analyzed the degree to which the distribution of each syllable across motif positions deviated from random using Likelihood ratio tests. Using the data from Monte Carlo simulations, we computed the expected distribution (i.e., null distribution) of syllables across motif positions if syllable sequencing was random. For example, to generate the null distribution for the "d" syllable, we computed the proportion of motifs (across birds) in which the "d" syllable was situated in the beginning, middle, or end of the motif for each iteration of the Monte Carlo simulation. We then calculated the average proportion across all 10,000 iterations for each of the three positions to generate the expected probability of observing a "d" syllable in each of these positions. Because the sequence of syllables within each bird's motif was randomized for each iteration of the simulation, this allowed us to compute the expected probability of a "d" occupying a particular position if sequence learning was random (i.e., if there were no biological predispositions for sequence learning). We then multiplied these average proportions by the number of birds that produced the "d" syllable to generate the expected number of birds that would have produced the "d" syllable at the beginning, middle, or end of their motif if sequencing was random. Finally, we compared these expected numbers to the observed number of birds that produced the "d" syllable at the beginning, middle, or end of the motif using a Likelihood ratio test. This procedure was repeated for each of the five tutor syllables.

To supplement the Likelihood ratio tests, we used Monte Carlo simulations to assess whether the observed likelihoods that individual syllables were produced at the beginning, middle, or end of the motif were each significantly different than chance. Using the syllable "d" as an example again, we observed that, of birds that produced the "d" syllable in their motif, 4% produced the "d" at the beginning of the motif, 46% of birds produced the "d" in the middle of the motif, and 50% of birds produced the "d" at the end of the motif. To assess whether each of these observed percentages were significantly different than chance, we created three null distributions from the Monte Carlo simulation, with distributions representing the likelihood of observing a "d" at the beginning, in the middle, or at the end by chance. We then compared the observed likelihoods of observing a "d" at the beginning, middle, and end against their respective null distribution. We deemed the observed percentage as significant if it lay outside the 95th percentile of values from the simulation (two-tailed). For example, 50% (13 of the 26) of the birds that produced a "d" in their motif positioned the "d" at the end of the motif, but in only 0.14% (14 of the 10,000) simulated iterations did we observe that $\geq 50\%$ of motifs ended with "d"; in other words, by chance alone, it was highly unlikely that 50% of birds ended their motif with a "d." Consequently, our results indicate a significant bias for zebra finches to produce the "d" syllable at the end of the motif. Conversely, only 3.8% (1 of the 26) of

these birds produced the “d” at the beginning of their motif, and in only 1.2% of the simulated iterations did we observe that $\leq 3.8\%$ of motifs began with “d”; therefore, our results indicate a significant bias for birds to *not* produce a “d” at the beginning of the motif.

Transitions between Syllables

We also analyzed the extent of patterning in syllable transitions within the motif. For example, we analyzed the degree to which birds exhibited consistent directionality (i.e., predictability) in sequencing for each of the 10 pairs of tutored syllables. To this end, we compared the observed asymmetry in syllable sequencing to the distribution of sequence asymmetries across the iterations of the Monte Carlo simulation. To compute sequence asymmetry, we first calculated, for each pair of syllables, the proportion of transitions in one direction (e.g., “e” to “a”) out of the total transitions in either direction (e.g., sum of “e” to “a” and “a” to “e”); this value should be 0.5 if transitions were equally likely in both directions. We computed the asymmetry in pairwise transitions as the absolute deviation of the observed proportion from 0.5 (“asymmetry value”). For example, across the 12 birds that produced motifs with a direct transition between “e” and “a” (in either direction), 8 of those were from “e” to “a”; therefore, the asymmetry value for the “e”–“a” pair was 0.167 ($((8/12)-0.5)$). To compute the overall degree of asymmetry in syllable sequencing, we averaged the asymmetry values across the 10 syllable pairs (“overall asymmetry value”). To compute the significance of the overall asymmetry value that we observed, we compared the observed overall asymmetry value to the distribution of overall asymmetry values derived from the Monte Carlo simulation. To generate this distribution, we calculated the overall asymmetry value on each of the 10,000 iterations of the simulation, and then computed the likelihood of observing an overall asymmetry value that was equal to or greater than the overall asymmetry value observed in our data (Figure 3D).

Alternation in Acoustic Structure

We ran Monte Carlo simulations to assess the significance of acoustic alternations within each bird’s motif (Figures 3E and 3F). We measured the mean frequency, duration, and mean amplitude, three features that are commonly found to alternate in speech, for each rendition of each syllable in every bird’s motif [57, 58]. Then we computed the median value for each feature across all renditions of the syllable and used these values in the analysis. We plotted the change in mean frequency, for example, from the previous syllable_(t-1) to the current syllable_(t) against the change from the current syllable_(t) to the next syllable_(t+1) across all birds to measure the extent of pitch alternation between adjacent syllables in the motif. If there is acoustic alternation from syllable-to-syllable, we would expect a negative relationship between the change from syllable_(t-1) to syllable_(t) and from syllable_(t) to syllable_(t+1). For each acoustic feature, we then computed the r^2 value of the relationship between the two changes (“observed r^2 ”) and compared this observed r^2 value to the distribution of r^2 values obtained for each of the 10,000 simulated iterations. For example, the observed r^2 value for the alternation in mean frequency was 0.476, and in only 18 of the 10,000 iterations did we observe an r^2 value equal to or greater than the observed r^2 ($p = 0.0018$). Therefore, the observed extent of pitch alternation in the songs of experimentally tutored birds was significantly greater than that expected by chance alone.

Analysis of Variation in Acoustic Features across Positions

We analyzed how acoustic features of syllables varied depending on the position in the motif (beginning, middle, or end) in our birds tutored with randomized sequences of syllables as well as birds tutored with species-typical song. We measured the duration, mean frequency, and mean amplitude of each rendition of each syllable in every bird’s motif, and then analyzed how the medians of these features varied depending on syllable position. We used mixed effects models with position (Beginning, Middle or End) as the independent variable and individual bird as a random variable (to account for the fact that multiple syllables were analyzed per bird) to analyze systematic variation in acoustic structure across positions. Such mixed effects models (with birdID as a random variable) are repeated-measures analyses that provide the flexibility to accommodate the facts that most birds have multiple syllables defined as “middle syllables” and that different birds have different numbers of middle syllables. Tukey’s HSD tests were used for post hoc contrasts. Significant differences between motif positions are identical if we compare values for beginning and end syllables to the average value for all middle syllables for each bird (i.e., each bird has a single data point per position).

For birds tutored with the same set of syllables (i.e., our experimentally tutored birds), we examined the degree to which positional variation in acoustic structure was due to variation in the types of syllables placed at different positions in the motif as well as variation in the acoustic structure of individual syllables according to position. Indeed, acoustic features of individual speech sounds systematically vary according to position [12, 13]. To differentiate between these possibilities, we conducted an additional analysis in which “syllable identity” (“a,” “b,” “c,” “d” or “e”) was included as an additional factor in the model. Acoustic features did not significantly vary across position when “syllable identity” was included in model, suggesting that positional variation in acoustic structure is primarily driven by variation in the placement of particular syllables at different parts of the motif.

Statistical Analyses

All analyses were done using MATLAB (Mathworks, Natick, MA) or JMP v. 11 (Cary, NC). We set $\alpha = 0.05$ for all tests.

DATA AND SOFTWARE AVAILABILITY

The datasets generated from this experiment have been deposited into Mendeley Data, available at <https://doi.org/10.17632/sywzmd483f.1>.

Current Biology, Volume 27

Supplemental Information

**Learning Biases Underlie “Universals”
in Avian Vocal Sequencing**

Logan S. James and Jon T. Sakata

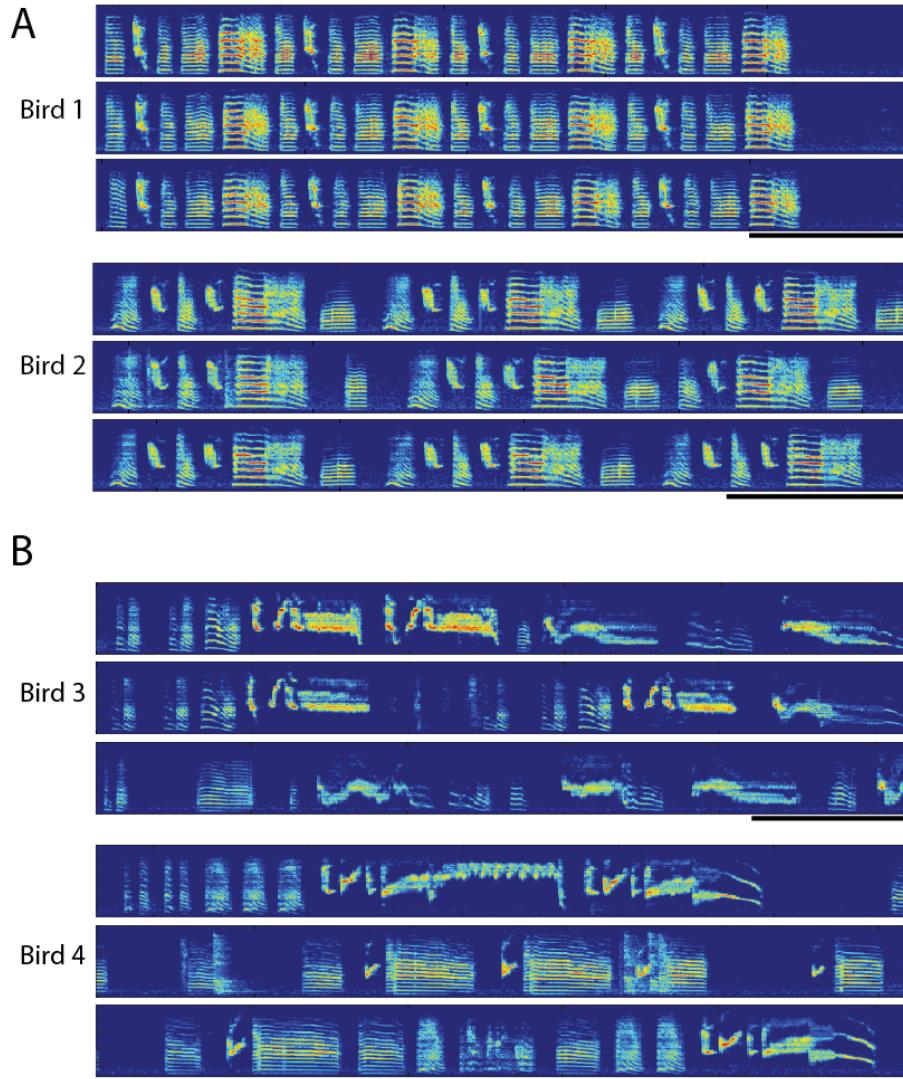


Figure S1: Examples of birds who were included and excluded from analysis (Related to Figure 2)

Spectrograms (time on the x-axis, frequency on the y-axis (0-10 kHz), color representing amplitude) of song bouts of example birds that were included (A) or excluded (B) from the analysis based on the ability to reliably detect stereotyped motifs. A) Representative spectrograms of three song renditions for two tutored birds (“Bird 1” and “Bird 2”) that produced consistent motifs at 4 months of age. B) Examples of three song renditions from two tutored birds that produced variable sequences in their adult songs (“Bird 3” and “Bird 4”), making it difficult to reliably detect stereotyped motifs that are typical of zebra finch song. Scale bars= 500 ms.

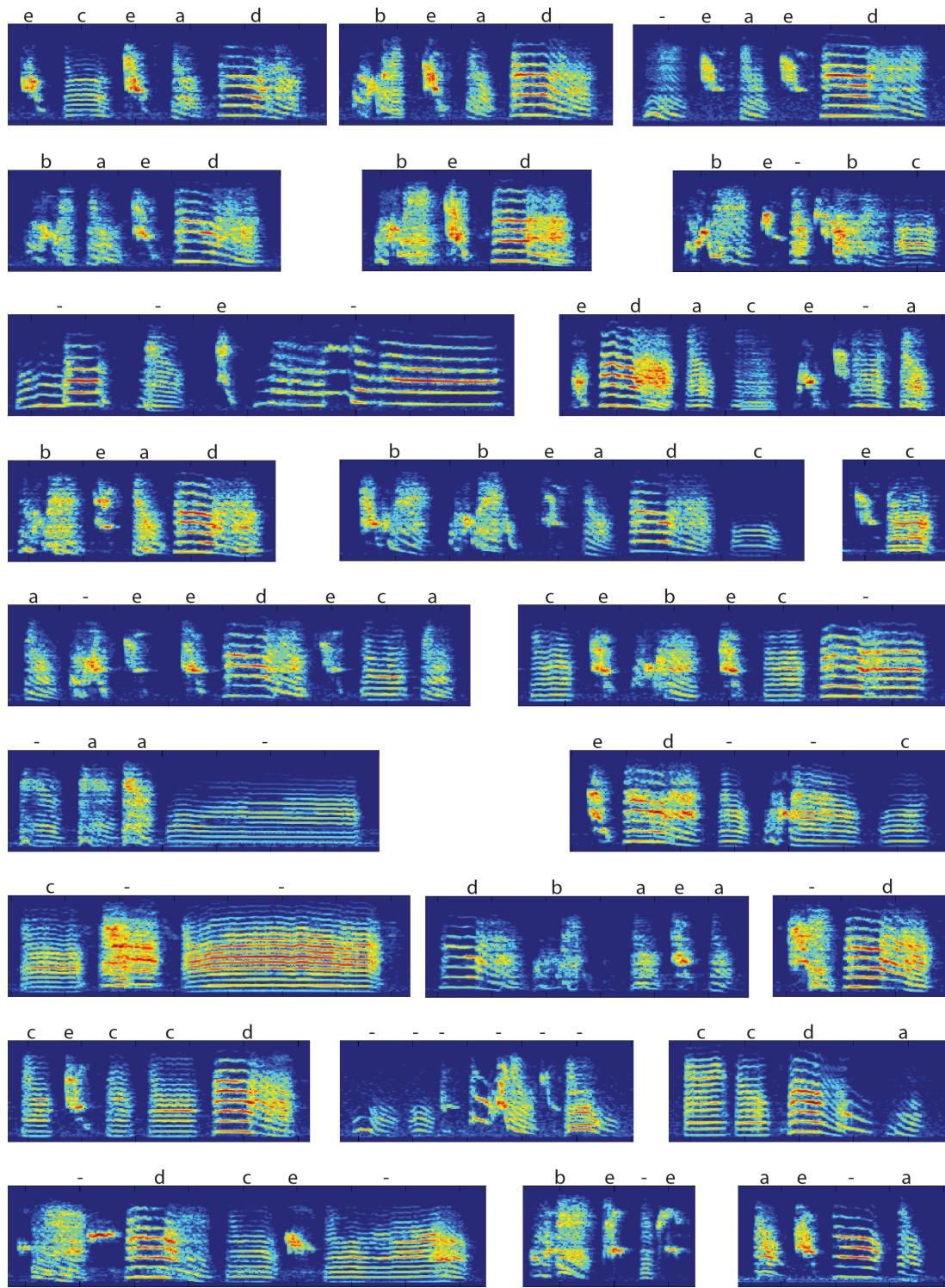


Figure S2: Example spectrograms of motifs of all birds included in the analysis part 1 (Related to Figures 2 and 3)

Letters and symbols above spectrograms denote the syllable labels used in our analysis ("-" denotes syllables that were not considered part of the tutoring set ("novel syllables")). N=24 of 45 birds.

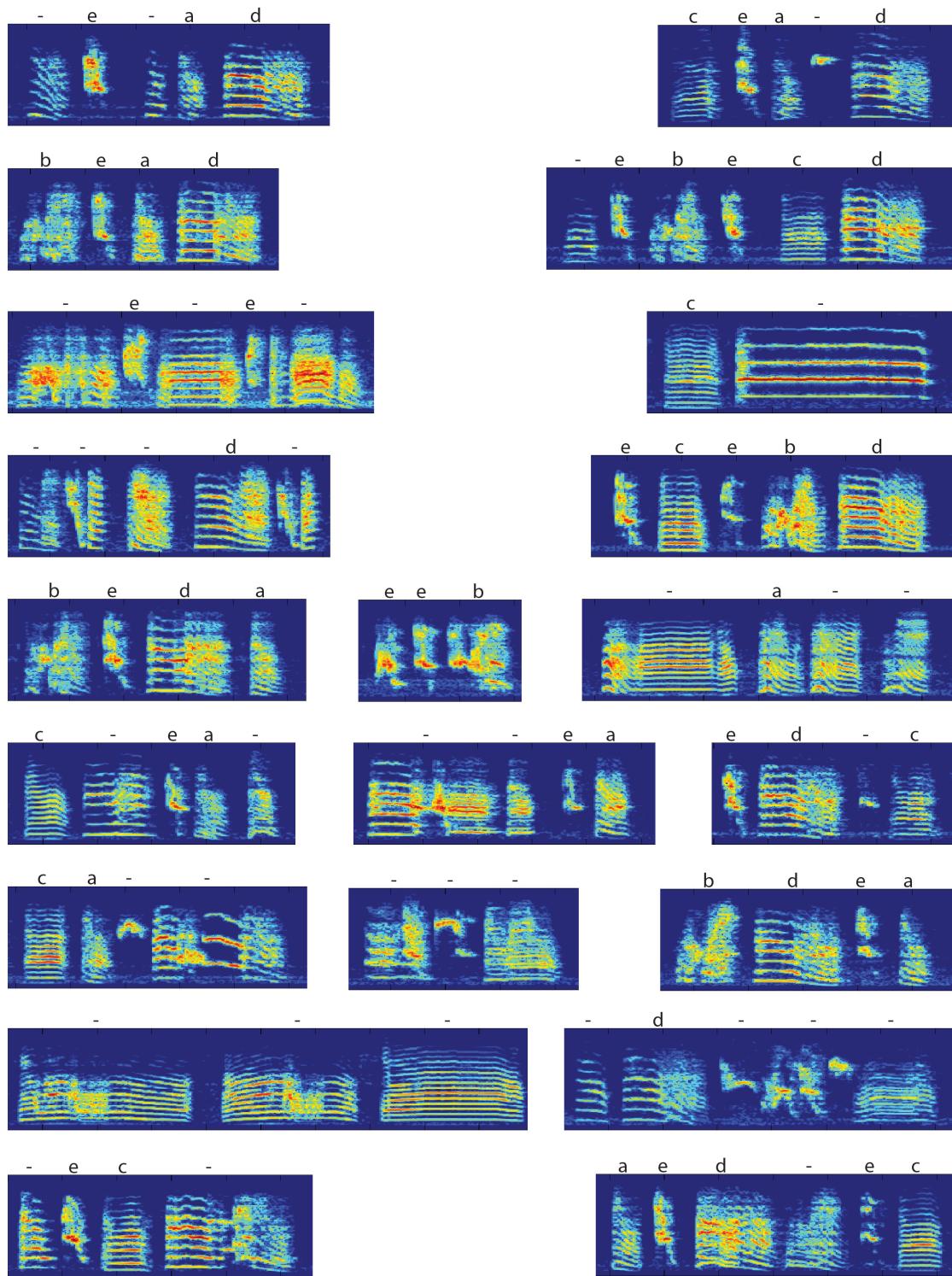


Figure S3: Example spectrograms of motifs of all birds included in the analysis part 2 (Related to Figures 2 and 3)

Letters and symbols above spectrograms denote the syllable labels used in our analysis (“-“ denotes syllables that were not considered part of the tutoring set (“novel syllables”)). N=21 of 45 birds.

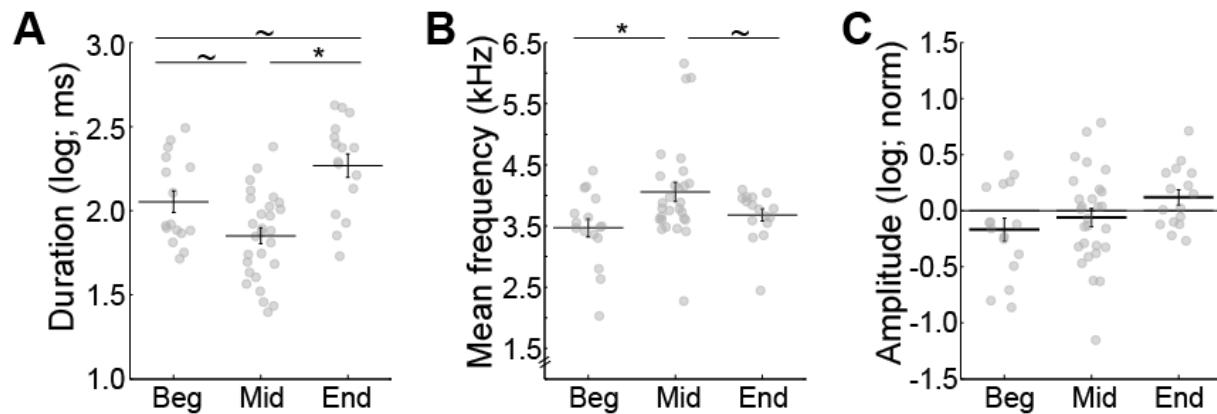


Figure S4: Positional variation in acoustic features of syllables that were produced by birds but not present in tutor stimulus (“novel syllables”; n=60; Related to Figure 2).

A) Novel syllables positioned at the end of the motif were significantly longer than those in the middle ($p<0.0001$; Tukey's HSD) and marginally longer than those at the beginning ($p=0.0669$). Novel syllables at the beginning were marginally longer than those in the middle ($p=0.0574$). B) Novel syllables positioned in the middle of the motif were higher in mean frequency than those at the beginning of the motif ($p=0.0084$) or those at the end of the motif ($p=0.0542$). C) Amplitude of novel syllables did not vary across motif positions. Amplitude values for each bird are normalized (“norm”; mean-subtracted) by the average amplitude of all syllables within his song. For all panels, dots represent median values for each syllable, horizontal lines indicate the means of the distributions, vertical lines indicate standard errors of the distribution, “*” denotes $p<0.05$, and “~” denotes $p<0.07$.

Source	Comparison	Duration	Mean Frequency	Amplitude
Experimental	Beg vs Mid syllables	<i>Beg>Mid: 65.5%</i>	<i>Mid>Beg: 65.5%</i>	
	Beg vs End syllables	<i>End>Beg: 66.7%</i>		
	Mid vs End syllables	<i>End>Mid: 72.7%</i>	<i>Mid>End: 72.7%</i>	
ZF Library	Beg vs Mid syllables		<i>Mid>Beg: 64.1%</i>	<i>Mid>Beg: 66.9%</i>
	Beg vs End syllables	<i>End>Beg: 78.7%</i>	<i>End>Beg: 70.5%</i>	<i>End>Beg: 78.7%</i>
	Mid vs End syllables	<i>End>Mid: 67.6%</i>		<i>End>Mid: 68.3%</i>

Table S1: Summary of trends regarding significant differences in acoustic features between motif positions (Related to Figure 2). Values in the table represent the percent of syllables that followed patterns with respect to other syllables in their corresponding motifs. Percentages are reported on a per syllable basis (instead of, for example, on a per bird basis) because birds produce multiple middle syllables in their song; therefore, for this calculation, each beginning or end syllable is compared to each of the multiple middle syllables to compute the percent of comparisons that follow a trend. Percentages are only provided for contrasts that are significantly different (mixed effects model followed by Tukey's HSD, $p < 0.05$). “Experimental” birds refer to zebra finches that we tutored with randomized sequences (see Figure 1), whereas “ZF Library” birds refer to those tutored with species-typical songs that we obtained from an online repository (see STAR Methods).