Temporal stability in songs across the breeding range of the Mourning Warbler may be due to learning fidelity and transmission biases

Jay Pitocchelli[1*], Adam Albina[2], R. Alexander Bentley[3], David Guerra[4], Mason Youngblood[5]

[1] Biology Department, Professor Emeritus, Saint Anselm College, Manchester, NH
[2] Computer Science Department, Saint Anselm College, Manchester, NH
[3] Department of Anthropology, University of Tennessee, Knoxville, Tennessee
[4] Physics Department, Saint Anselm College, Manchester, NH
[5] Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY
[*] Corresponding Author: jpitocch@anselm.edu

Accepted in *Ornithology*: https://doi.org/10.1093/ornithology/ukae046

ABSTRACT

We found a stable pattern of geographic variation in songs across the breeding range of the Mourning Warbler over a 36 yr period. The Western, Eastern, Nova Scotia, and Newfoundland regiolects found in 2005-2009 also existed from 1983-1988 and 2017-2019. Each regiolect contained a pool of syllables that were unique and different from the other regiolects. The primary syllable types that defined each regiolect were present throughout the study, but there were changes in the frequencies of variants of these syllable types in each regiolect. We developed an agent-based model of birdsong learning within each regiolect to explore whether these frequency changes were consistent with unbiased copying or two forms of transmission bias: frequency bias and content bias. Strong content bias, possibly for more complex syllables, best models the temporal dynamics across regiolects. In combination with a high estimated learning fidelity, this may explain why regiolects and syllable types were stable for 36 years. We also examined whether variation in physical parameters of song over time could be attributed to acoustic adaptation to breeding habitat, using Landsat variables as a proxy for vegetation characteristics of each male's breeding territory. The physical parameters of the songs, which changed little over time, revealed no coherent relationships with the Landsat variables and therefore little evidence for acoustic adaptation.

Keywords: birdsong, cultural evolution, *Geothlypis philadelphia*, Mourning Warbler, regiolects

LAY SUMMARY

- We studied changes in Mourning Warbler songs, syllables, and physical parameters of songs from three time periods throughout the breeding range: 1983–1988, 2005–2009 and 2017–2019.
- We found that the same pattern of geographic variation in songs existed in each time period.
- The stability may be due to strong content bias where birds in different regions of the breeding range preferred to learn certain syllable types unique to each region.
- Physical parameters of song also remained largely unchanged over the study period.
- Song parameters were not selected for optimal transmission through the breeding habitat.

INTRODUCTION

Geographic and temporal changes in birdsong are thought to be the results of social learning and cultural evolution (Lynch 1996, Aplin 2019). Change in songs may occur in the qualitative characteristics like phrases, syllables, and notes and or it might also occur in the frequency and duration parameters. The mechanisms responsible for changes in syllables and songs are considered equivalent to drivers of change in genetic evolution. Drift, mutation, selection, and learning biases have been identified as agents of cultural evolution (Lynch 1996, Williams 2021). Frequency and duration parameters of songs may change in response to habitat-dependent selection for optimal sound transmission through the breeding habitat (Morton 1975). Habitat-dependent selection on sound transmission is known to drive change and signal divergence in the physical parameters of songs among populations (Slabbekorn and Smith 2002).

We studied cultural evolution in both the qualitative and physical characteristics of songs of the Mourning Warbler (*Geothlypis philadelphia*). We used empirical data from a long-term database of songs recorded over 36 years from throughout the breeding range. Our goals were to describe change versus stability in patterns of macrogeographic variation over time, use agent-based modeling to assess potential mechanisms of change, and test for possible acoustic adaptation of song parameters to breeding habitat using Landsat variables as a proxy for vegetation characteristics of breeding territories.

Pitocchelli (2011) described a pattern of macrogeographic variation in Mourning Warbler songs with four different regiolects based on recordings from 2005-2009: Western, Eastern, Nova Scotia and Newfoundland. One of our goals was to evaluate the stability of the regiolects, their geographic boundaries, and an admixture zone in central Ontario and the Great Lakes states, by comparing the pattern from 2005-2009 with a historical pattern from 1983-1988 and a later pattern from 2017-2019. To explore the drivers of cultural evolution in each regiolect, we used a framework called simulation-based inference (a.k.a. generative inference) that is used to infer underlying processes from patterns in observed data (Kandler and Powell 2018, Cranmer et al. 2020). More specifically, we fit parameters from an agent-based model of birdsong learning to our real data using BayesFlow (v1.1.4 in Radev et al. 2023), a method that uses deep neural networks to approximate posterior distributions (Radev et al. 2020) and has been successfully used to infer cognitive processes (von Krause et al. 2022, Schumacher et al. 2023). One of the biggest challenges in inferring underlying processing from population level data is accounting for equifinality, or the fact that different transmission mechanisms can yield similar outcomes (Barrett 2019).

We modified an existing agent-based model of birdsong (Lachlan et al. 2018, Youngblood and Lahti 2022) to account for two forms of cultural transmission bias which have clear empirical evidence in songbirds: frequency bias (Lachlan et al. 2018) and content bias (Soha and Marler 2000, Youngblood and Lahti 2022). Frequency bias (e.g. conformity bias vs. anti-conformity bias) occurs when the commonness or rarity of cultural variants affects their adoption, whereas content bias occurs when some cultural variants are more likely to be learned because of their content (e.g. complexity) (Rendell et al. 2011). Transmission biases are generally thought to be genetically-evolved predispositions for learning information or behavior that enhances fitness (Kendal et al. 2009). Simulations show that frequency bias, content bias, and unbiased copying lead to differences in song diversity that are visually and statistically discriminable (Lachlan et al. 2018). Our goal was to determine the combination of transmission biases and other model parameters that best recreate the cultural dynamics of Mourning Warbler song.

According to the Acoustic Adaptation Hypothesis, physical parameters of vocalizations will be shaped and selected by the environment to optimize sound transmission through the breeding habitat. Birdsongs with lower frequencies, narrower bandwidths, fewer song elements, longer song elements, and longer intervals between elements suffer less reverberation and attenuation and occur more often in closed habitats with dense vegetation (Boncoraglio and Saino 2007, Derryberry 2011). We studied the relationship between physical parameters of songs and vegetation characteristics of breeding habitat during the 2005-2009 and 2017-2019 time periods using high resolution Landsat satellite data as a measure of vegetation density on breeding territories. If transmission of the physical parameters of Mourning Warbler songs were affected by habitat characteristics, then variation in the Landsat variables should explain variation in song parameters in different habitats (Smith et al. 2013). Finding a significant relationship between increasing vegetation density, lower frequencies, longer song elements, and longer intervals between elements in Mourning Warbler songs would support the Acoustic Adaptation Hypothesis.

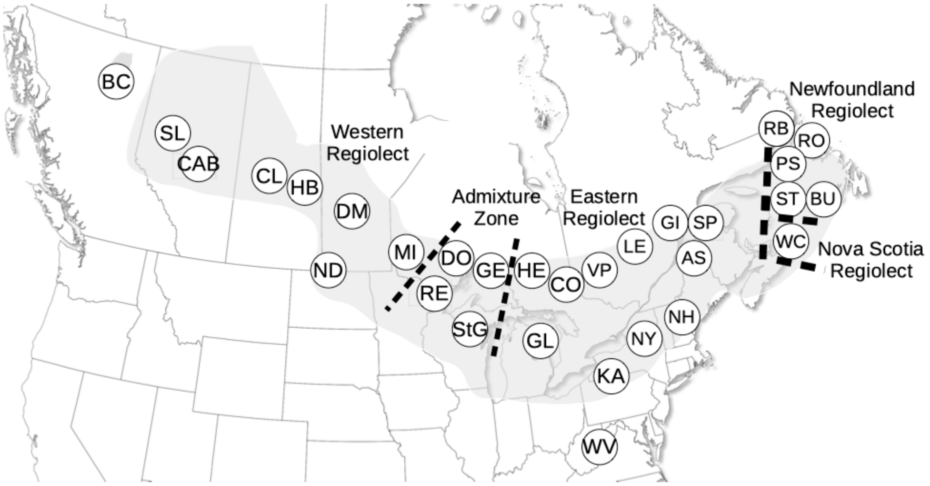

Figure 1. Study sites for all three time periods. BC –Fort Nelson, BC. SL – Slave Lake, AB. CAB – three localities near Lodgepole, AB. CL – Candle Lake, SK. HB – Hudson Bay, SK. ND – Turtle Mountains, ND. DM – Duck Mountain Provincial Park, MB. MI – Minitaki and Dryden, ON. RE – Remer and Waskish, MN. DO – Dorion, ON. StG – St. Germain, WI. GE – Geraldton, ON. HE – Hearst and Calstock, ON. GL – Glennie, MI. CO – Cochrane, ON. WV – Durbin and Cranberry Mtn., WV. KA – Kane and Bradford, PA. VP – Val Paradis, QC. LE – Lebel, QC. NY – Chazy, NY. NH – Franconia and Jefferson, NH. GI – Girardville and Normandin, QC. AS – Fort Kent and Ashland, ME. SP – St. Paule and Rimouski, QC. WC – Wreck Cove, NS. ST – Stephenville Crossing, NL. BU – Burgeo, NL. PS – Port Saunders, NL. RB – Red Bay, NL. RO - Roddickton, NL.

METHODS

Recordings, Study Areas, and Time Periods

We analyzed songs of 1391 males from 29 localities throughout the breeding range of the Mourning Warbler (Figure 1, Supplementary Material Table S1). Recordings were made during three different time periods: 1983-1988, 2005-2009 and 2017-2019. J. P. used a Dan Gibson E. P. M. 300 microphone and a Uher-4000 report reel to reel tape recorder from 1983–1986. Later recordings from the 1980s were made with the same microphone and a Sony Walkman Professional WM-D6 cassette recorder. J. P. used a Marantz PDM670 digital recorder, Telinga Pro Universal parabola, and Sennheiser ME62 microphone for all 2005-2019 recordings. Mourning Warblers have a single song repertoire (Pitocchelli 2020) so the best sonogram from each male was selected for this study. All recordings were digitized in 24 bit WAV format, 44 kHz sampling rate, framelength of 512 points, Hann window function and filter bandwidth of 124 Hz using Raven Pro 1.5 (Bioacoustics Research Program 2014).

Songs and Syllables

We analyzed change in the songs, syllable types and variants of syllable types in this study. Most birds sang two-part songs with introductory and concluding phrases (Figure 2). The introductory phrase was composed of a single syllable, repeated several times. Syllables of the concluding phrase were different from the introductory syllable, and were repeated twice or more times, or were two different syllables. We used a hierarchical classification of syllable types and variants of each syllable type similar to Lynch et al. (1989). The first occurrence of a syllable was labeled a syllable type. Differences among syllable types included overall shape of notes, positions of notes within the syllables, and number of notes per syllable. Minor variations in the number or shape of notes in each syllable type were catalogued as variants of that type (examples in

Supplementary Material Figure S1). Syllables from the 1983-1988 and 2017-2019 time periods were entered into the syllable catalogue created by Pitocchelli (2011). The primary method for separating syllable types and variants was subjective, visual pattern recognition of differences between syllables by J. P. Students in J.P.'s biostatistics class were independent observers who corroborated the initial syllable classification. For 70% of the syllable variants the students were in 100% agreement with the professor (Pitocchelli 2011). The correlator in Raven 1.5 (Bioacoustics Research Program 2014) was also used to confirm similarities among syllable variants with higher peak correlations among variants of the same syllable type. New syllables from the 1983-1988 and 2017-2019 recordings were added to the original 2005-2009 catalogue by visual inspection by J.P. and corroborated by M.Y.

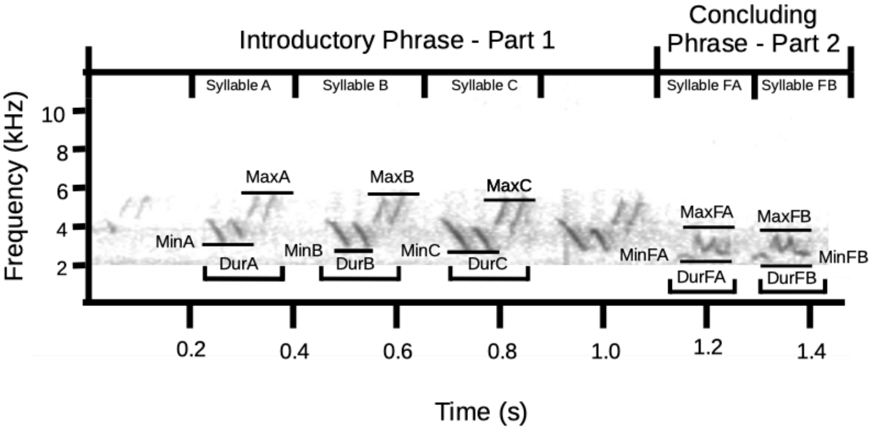

Figure 2. Phrases, syllables and measured physical parameters of songs. Max and Min – maximum and minimum syllable frequencies respectively, Dur – syllable duration.

Analysis of Macrogeographic Variation

We analyzed and compared the pattern of macrogeographic variation in song over the three time periods. All songs were assigned a song type named after the syllable type found in the introductory phrase. This assignment was chosen because the introductory phrase often contained the most syllables in the song and was the longest phrase of the song. We calculated the percent of males singing each song type at each locality and plotted pie charts of these percentages on separate maps for each time period.

We conducted a clinal analysis on the song types and syllables data to test the isolation by distance model. A Mantel test (Mantel 1967) was used to examine the relationship among geographic distances and Euclidean song distances among localities. Geographic distances were extracted from Garmin GPS data. Euclidean song distances among localities were calculated using the percentages of syllable types and variants from each locality using a UPGMA (unweighted pair-group method with arithmetic averaging) analysis in SPSS 27.0 (IBM Corp. 2020). Three Mantel tests, one each for the 1983-1988, 2005-2009 and 2017-2019 time periods, were used to examine the relationship between a matrix of geographic distances and a corresponding matrix of Euclidean song distances. Clinal variation and isolation by distance would have occurred if there was a significant positive relationship among song distances and geographic distances. Each Mantel test was one-tailed with 10,000 permutations using zt 1.1 software from Bonnet and Van de Peer (2002).

Males with alien songs immigrating to a different population introduce new songs that could change the composition of that population's song types (Fayet et al. 2014). Hybrid songs composed of syllables from different song types could also create new sources of song variation in a population. We conducted elementary statistical analyses of these effects on the composition of songs in each regiolect over time. These results provide only preliminary insights into the effects of immigration and hybrid songs so we reported them in the Supplementary Materials.

Agent-Based Modeling

We used a modified form of an agent-based model recently used to detect content bias in House Finches (*Haemorhous mexicanus,* Youngblood and Lahti 2022). Descriptions and prior distributions of parameters for the model are listed in Table 1. The model of cultural transmission is initialized with N_B birds across the four regiolects, each of which is represented by a separate grid that approximates geography. The Western and Eastern regiolects are 50 cells wide and 20 cells high (1000 cells in total), while the Nova Scotia and Newfoundland regiolects are 10 by 10 (100 cells in total) (see Supplementary Material Figure S3 for a diagram). The N_B birds are split into four batches so that the average density of each grid is the same (e.g. if $N_B = 88{,}000$, then the Western and Eastern regiolects have 40,000 birds and Nova Scotia and Newfoundland have 4,000 birds). Each bird is randomly assigned a geographic index (g) that represents their cell within their regiolect. Note that the spatial component of our model is only meant to be a coarse approximation of the species range—included to introduce geographic clustering and allow for transmission between regiolects (details below).

Table 1. Descriptions and prior distributions for each of the parameters explored in the agent-based model. The prior distributions for frequency bias and content bias were truncated to be positive.

Parameter	Prior	Truncated	Description
Population size (N_B)	$U(5{,}000, 50{,}000)$	NA	Uniform between 5,000 and 50,000
Demonstrators (D)	$U(2, 8)$	NA	Uniform between 2 and 8
Innovation rate (μ)	$B(1, 40)$	[0, 1]	Beta with $\alpha = 1$ and $\beta = 40$
Frequency bias (α)	$N(1, 0.2)$	[0, Inf)	Normal with $\mu = 1$ and $\sigma = 0.2$, truncated below 0
Content bias (β)	$N(0, 2)$	[0, Inf)	Normal with $\mu = 0$ and $\sigma = 2$, truncated below 0

The model is also initialized with N_S possible syllable types. N_S was set to 312—the total number of syllable types present across the four regiolects in the real data, as estimated by an unseen species model (Supplementary Material Figure S2) (Hsieh et al. 2016). Repertoire sizes (r_S) are set to 3, the typical value in Mourning Warblers across most of their range. In the first year of the simulation, repertoires are filled by pseudorandomly sampling from the real distribution of syllable types in 1985 weighted by their counts. Each possible syllable type is assigned an attractiveness index (M) drawn from a truncated normal distribution with $\mu = 1$, $\sigma = 0.5$, and a lower bound of 0.

At the start of each year a new generation learns syllables from a random set of D demonstrators, sampled from the 10% of birds who are closest to the focal bird in that regiolect. If the focal bird is in a cell that borders a neighboring regiolect, then transmission between regiolects is simulated in two steps. First, the potential migrant demonstrators (i.e. the 10% of birds from the neighboring regiolect who are closest to the border) are compiled. Second, each original demonstrator is replaced by a random migrant demonstrator, with probability equal to the proportion of males in each regiolect who sing alien songs from neighboring populations (see Supplementary Material Table S3 and Figure S3 for more information). The number of demonstrators, D, is drawn from a uniform distribution between two and eight.

The probability of a bird learning a syllable x, $P(x)$, depends upon the frequency of that syllable among the demonstrators (F_X) and the attractiveness of the content of the syllable (M_X), normalized against the probability of adopting other syllables in the population:

$$P(x) = \frac{F_x^{\alpha} \times M_x^{\beta}}{\sum_{y=1}^{n} F_y^{\alpha} \times M_y^{\beta}}$$

To simulate frequency bias, F_X is raised to the exponent α, where $\alpha > 1$ corresponds to conformity bias and $\alpha < 1$ corresponds to anti-conformity bias. To simulate content bias, M_X is raised to the exponent β, where $\beta = 0$ is neutrality and $\beta > 0$ corresponds to content bias. The probabilities, $P(x)$, are used as sampling weights to fill each bird's repertoire, which represents the number of unique syllable types a bird sings. This means that if a subset of syllables become dominant and have a very high probability of adoption, some birds may end up with only one or two types in their repertoire. During the learning process, new syllable variants are innovated with probability μ, where innovated syllables are drawn from the initial pool of possible syllables (N_S). Mortality is simulated at the end of each year by randomly selecting 50% of the agents to remove. There are no studies of mortality rates for the Mourning Warbler. We arbitrarily used 50%, near the middle of the 63.1% reported for the Golden-cheeked Warbler (*Setophaga chrysoparia*, Reidy 2007) and 35% for Prairie Warblers (*Setophaga discolor*, Nolan et al. 2020).

At the beginning of the year, a new generation is added to the population to bring the size up to the population size of the previous year multiplied by 0.989 (to reflect a 1.1% annual decline; Sauer et al. 2017). The frequency distribution of syllable types is collected from the model in 2005-2009 and 2017-2019, from the same number of birds recorded in the real dataset in the same time period. Simulation-based inference was conducted using BayesFlow (v1.1.4) in Python (v3.9.10) (Radev et al. 2023), a method that uses amortized neural networks to approximate the posterior distributions of parameters in a generative model (Radev et al. 2020). We used BayesFlow (v1.1.4) because it has been shown to outcompete various forms of approximate Bayesian computation (e.g. MCMC, random forests) in both accuracy and computational cost. Simulation-based inference was conducted with the following steps:

1. 100,000 iterations of the model were run from 1983-2019 to generate simulated frequency distributions from 2005-2009 and 2017-2019 in each regiolect for different values of the parameters: N_B, D, μ, α, and β. The results were split into a training set (n = 95,000) and a testing set (n = 5,000).
2. BayesFlow (v1.1.4) was run on the training set for 500 epochs with a batch size of 32 and a learning rate of 0.0001. Sixteen informative summary statistics were learned from the raw model output using a hierarchical network comprised of a deep set transformer to compress the individual frequency distributions (Bloem-Reddy and Teh 2020), a time series transformer to compress the two timepoints within each regiolect (Wen et al. 2023), and a simple two-layer network to combine the data across the four regiolects. These summary statistics were then mapped to the parameters of the agent-based model using an invertible neural network. This architecture is based on default modules that come with BayesFlow (v1.1.4), and further details can be found in the analysis code.
3. The trained neural networks were used to produce 10,000 posterior predictions for each parameter in the agent-based model, after being provided with the observed frequency distributions from 2005-2009 and 2017-2019 in each regiolect.

This procedure is able to recover known values of the innovation rate (μ), number of demonstrators (D), and the two biases (α and β) from the testing set with high accuracy (r^2 = 0.889, 0.775, 0.428, and 0.793, respectively). Initial population size (N_B) left a weaker statistical signal in the data (r^2 = 0.261), so we will not make strong conclusions about this parameter. The observed and simulated frequency distributions from each regiolect in 2017-2019 can be seen in Supplementary Material Figure S4, and the loss curve from the training of the deep neural networks can be seen in Supplementary Material Figure S5.

To assess whether syllables with particular characteristics had higher persistence, we conducted Bayesian logistic regression with all of the syllable types detected in 1983-1988 using the *rstanarm* package (v2.26.1) in R (v4.3.1) (Goodrich et al. 2020). We use the following predictor variables for each syllable type: duration (s), minimum frequency (Hz), maximum frequency (Hz), and concavity (changes in the sign of the slope of the mean frequency trace/s). Concavity, a basic indicator of syllable complexity (Ju et al. 2019, Youngblood and Lahti 2022), was manually calculated from representative spectrograms. Predictor variables were averaged across all of the observations of each syllable type and scaled and centered prior to modeling. Whether or not the syllable type persisted to 2017-2019 (logical: true/false) was used as the outcome variable. We used Student's t distributions with a scale of 2.5 as priors to allow for a relatively wide range of parameter estimates. The Bayesian logistic regression was run with 16,000 iterations across four chains.

Physical Parameters of Songs

We studied temporal variation in the physical parameters of songs. Raven 1.5 (Bioacoustics Research Program 2014) was used to measure syllable maximum (Max) and minimum (Min) frequencies in Hz of the A, B and C syllables in the introductory phrase: MaxA, MinA, MaxB, MinB, MaxC, MinC (Figure 2). Syllable duration (Dur) of these three syllables in the introductory phrase (DurA, DurB, DurC) and entire song duration (DUR) were measured in s (Figure 2). The total number of syllables per song was divided by the entire song duration (DUR) to obtain a syllable rate for each song. Syllable rate was analyzed separately from the other parameters. Principal Components Analysis (PCA) using a variance-covariance matrix was performed on the song parameters to condense the original variables into a new set of Principal Component scores for each song. All variables were log10-transformed for the PCA because of the magnitude of the differences between frequencies in Hz and duration variables in s: DurAlg (log10 - DurA), MinALg (log10 - MinA), MaxALg (log10 - MaxA), DurBLg (log10 - DurB), MinBLg (log10 - MinB), MaxBLg (log10 - MaxB), DurCLg (log10 - DurC), MinCLg (log10 - MinC), MaxCLg (log10 - MaxC), DURLg (log10 - DUR). Variables for the concluding phrase were omitted because the introductory phrase was the longest part of most songs and many males did not have a concluding phrase. Including the shorter, concluding phrase would have eliminated too many songs from the statistical analyses.

The PCA produced a new set of PC1, PC2 and PC3 scores for each male's song that were used in subsequent analyses. PC1 explained 55.0% of the variance in the data followed by 30.1% for PC2 and 5.9% for PC3 (Supplementary Material Table S2). Maximum and minimum frequencies of the song syllables from the introductory phrase had high positive loadings while duration variables had low negative loadings on PC1. Males with high PC1 scores had songs with higher frequencies but shorter syllables and songs, similar to results for Paradise Flycatchers (*Terpsiphone mutata*, van Dongen and Mulder 2006). All loadings on PC2 were positive so all song parameters increased with increasing PC2 scores. PC3 loadings were mixed but dominated by a positive loading for song duration. Songs with higher PC3 scores had longer songs.

We conducted ANOVAs of the average PC1 score across the three time periods within each regiolect to determine whether the physical parameters differed among the three time periods. PC1 was chosen because it accounted for most of the variation in the song parameters (Supplementary Material Table S2). Student-Newman-Keuls (SNK) post hoc tests were used to detect significant differences in song parameters across time periods in each regiolect.

There were three different ANOVAs for MO1, MO1P and MO18 songs from the Eastern, Nova Scotia and Newfoundland regiolects respectively. The Western regiolect differed from the other regiolects in that it was made up of three different song types, mixed together within some samples, and in close geographic proximity to each other. Although they belonged to the same regiolect, the MO6, MO16, and MO24 song types were defined by unique syllables that were as qualitatively distinct from each other as they were from song types from each of the other regiolects. It is possible that the physical parameters of these song types may be evolving differently from each other within the Western regiolect. Therefore, we performed separate ANOVAs on the PC1 scores of the MO6, MO16 and MO24 song types to assess change in each of them over time. A sequential Holm-Bonferroni correction was used to control for the effects of using multiple test procedures (Holm 1979).

Acoustic Adaptation

If acoustic adaptation has played a role in song divergence in Mourning Warblers, then we predict that territories with closed, dense vegetation should have lower frequencies, slower delivery of repeated elements, longer elements or syllables, and longer songs. We used PC1 scores of songs to determine if this relationship existed between these song parameters and vegetation density on breeding territories. Since songs with higher PC1 scores had higher frequencies, shorter syllables, and shorter songs, we expected a negative relationship between PC1 scores and increasing vegetation density. We separately calculated syllable rate (number of syllables per s) as a measure of syllable delivery and used it to determine whether syllable delivery decreased with increasing vegetation density.

Landsat data were used as a proxy for ecological characteristics of each breeding territory. Latitude and longitude of each males' territory were collected with a Garmin Dakota 10 and a Garmin Oregon 650t for 2005-2009 and 2017-2019 respectively. Landsat data were downloaded using these coordinates with USGS EarthExplorer (USGS 2022). The 30 m resolution, Tier 1, Landsat images from the Thematic Mapper represent a proxy for vegetation measurements in each territory. The Landsat images were chosen with restrictions of daytime observation with less than 10% cloud cover and from the summer months of June - August to best represent the vegetation during the breeding season. Individual bands from Landsat 4-8 were downloaded and composite images were produced in ArcGIS Pro so that the bands from the different generations of the satellites were aligned with the spectral definitions set by Landsat 4 and 5. Bands were set to: B1 - blue to distinguish soil from vegetation and deciduous from coniferous vegetation, B2 - green to emphasizes peak vegetation, B3 - red to discriminate vegetation slopes, B4 - near IR to emphasize biomass, B5 - short wave IR 1 to indicate moisture content of soil and vegetation, B6 - thermal IR to give an estimate of soil moisture, and B7 - short wave IR 2 to indicate moisture content of soil and vegetation (USGS 2022). Mosaics of the composite bands of each year the bird songs were collected were plotted on the same map. The extraction of the Landsat Bands from the maps used a circular buffer with a radius of 50 m to match the approximate size of the bird's territory of 0.8 hectares (Pitocchelli 2020). The Landsat tiles inside the buffer for each bird song recording were extracted and focal statistics were applied to find the mean of each Landsat band in each territory. Following the method outlined by Smith, et. al. (2013), random forest-based analysis was conducted to investigate the relationship between PC1 scores from the songs and the seven averaged and normalized Landsat bands at each territory. Finally, we used the bands to compute new indices of vegetation at each territory based on work by Thi Loi et. al. (2017) and Nguyen et. al. (2021). The Normalized Difference Vegetation Index (NDVI) is an indicator of the health of the vegetation:

$$NDVI = \frac{(B4 - B3)}{(B4 + B3)}$$

The Normalized Difference Moisture Index (NDMI) is an indicator of moisture content in the vegetation:

$$NDMI = \frac{(B4 - B5)}{(B4 + B5)}$$

The Green Normalized Difference Vegetation Index is a good indicator of chlorophyll differences in a crop:

$$GNDVI = \frac{(B4 - B2)}{(B4 + B2)}$$

The Bare Soil Index (BSI) is an of bare soil but is also able to measure the amount of ground cover in a region, with a negative BSI indicating dense vegetation:

$$BSI = \frac{\big((B5 + B3) - (B4 + B1)\big)}{\big((B5 + B3) + (B4 + B1)\big)}$$

Higher scores for NDVI, NDMI, and GNDVI indicate closed habitats with higher density of vegetation while BSI decreases with vegetation density.

We did not have GPS coordinates for field work in 1983-1988 so statistical analyses were performed only on data from the 2005-2009 and 2017-2019 time periods. Data from these two time periods were analyzed separately because Mourning Warblers from 2017-2019 could not breed on the same territories as earlier birds due to their preference for ephemeral, disturbed second growth which changes over time due to ecological succession (Pitocchelli 2020). We used multiple regression analysis to determine whether there was a significant relationship between Landsat indices of vegetation density and the PC1 scores. PC1 score was the dependent variable and the Landsat measurements were the independent variables. One regression analysis was performed on PC1 scores from the 2005-2009 time period and a second analysis was performed on PC1 scores from the 2017-2019 time period. We repeated this approach with the syllable rate data from the two different time periods. The log-transformed syllable rate (SyllRateLg) was the dependent variable and the Landsat indices were the independent variables. We log transformed SyllRate because of the order of magnitude of difference with the Landsat measurements. The sequential Holm-Bonferroni correction was used to control for the effects of using multiple test procedures (Holm 1979). The r^2 or coefficient of determination was a measurement of the per cent of variation in PC1 scores or syllable rate explained by the regression and combined Landsat indices. A significant multiple regression and high r^2 values would show the Landsat variables explained significant variation in the songs and support the Acoustic Adaptation Hypothesis.

RESULTS

First, we found that the persistence of syllable types and syllable variants was remarkably stable over the 36 year duration of the study. Average survival of syllable types across regiolects was the highest at 95%. The average survival of syllable variants was 54.9%. The highest survival of variants occurred in the Western regiolect vs. the lowest in Nova Scotia (Table 2). The low average for Nova Scotia was probably due to lower sample size compared to the other regiolects. Using criteria from Byers et al. (2010), a song from 2017-2019 was considered a match to a song from 1983-1988 if it possessed the same syllable types, syllable variants and in the same sequence as a song from 1983-1988. Song survival across the regiolects was lowest, averaging 23.7% (Table 2).

Table 2. Survival of songs, syllable types and variants that survived from the 1983-1988 to 2017-2019 time periods.

Regiolect	Songs (n^a)	Syllable Types (n^b)	Syllable Variants (n^c)
Western	29.3% (n = 41)	100% (n = 10)	69.4% (n = 36)
Eastern	18.6% (n = 43)	100% (n = 7)	66.7% (n = 30)
Nova Scotia	20.0% (n = 10)	100% (n = 2)	33.3% (n = 9)
Newfoundland	28.0% (n = 25)	80.0% (n = 5)	50.0% (n = 26)
Average[d]	23.7%	95.0%	54.9%

[a] – number of different songs from 1983-1988 in each regiolect.
[b] - number of different Syllable Types from 1983-1988 in each regiolect.
[c] - number of different Syllable Variants from 1983-1988 in each regiolect.
[d] - n = 4, number of regiolects.

The pattern of geographic variation in the song types was also stable over time. The four regiolects first described by Pitocchelli (2011) contained song types, defined by syllable types in their introductory phrases, which were unique to each region of the breeding range. MO6, MO16 and MO24 song types from the Western regiolect contained MO6ST, MO16ST and MO24ST syllable types in each of the different time periods (Figure 3). The MO1 song type with the MO1ST of the Eastern regiolect was also present in 1983-1988, 2005-2009, and 2017-2019 (Figure 3). Nova Scotia regiolects, defined by the MO1P song type and MO1PST syllable type, persisted from the 1983-1988 time period to the 2017-2019 time period. MO18 song types with MO18DST and MO18UST syllable types were found in Newfoundland in each time period (Figure 4). The distribution and boundaries of the relative frequencies of the song types, shown with pie charts at each locality, remained stable from 1983-1988 (Figure 5), 2005-2009 (Figure 6), and in 2017-2019 (Figure 7). Pitocchelli (2011) documented an admixture zone between the Western and Eastern regiolects in the Great Lakes states and central Ontario in 2005-2009. This zone contained a mix of males that sang the Western MO6 song type, the Eastern MO1 song type or hybrid songs with a mix of MO1 and MO6 syllables. The zone occurred in the same area in 1983-1988, 2005-2009 and 2017-2019 (Figures 5 - 7).

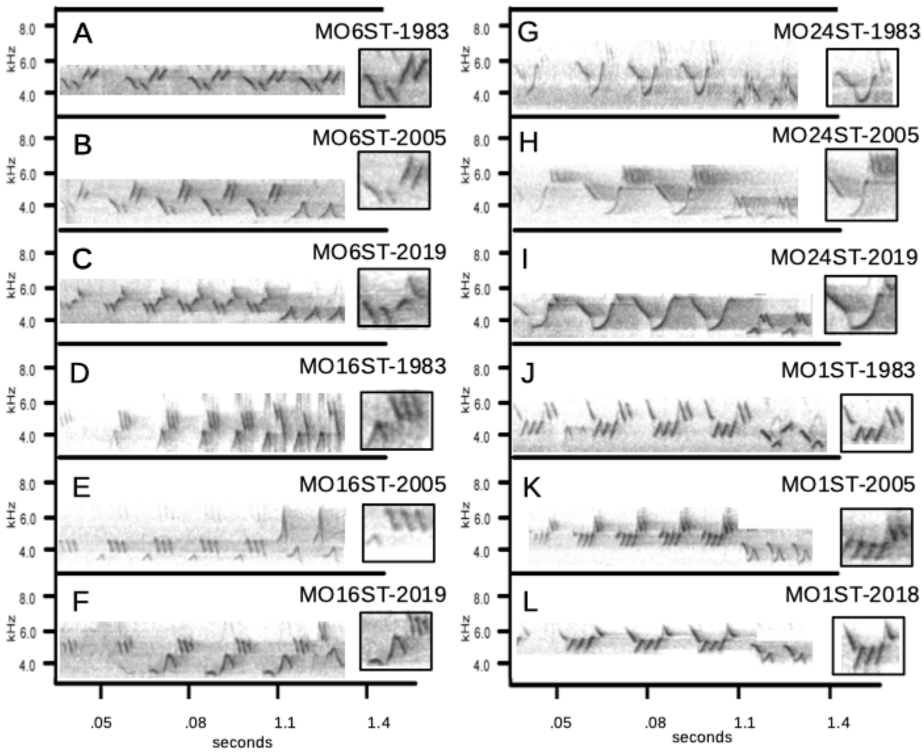

Figure 3. Examples of songs from the Western (MOST6, MOST16, MOST24) and Eastern regiolects (MOST1) for each time period. Enlarged images are variants of the syllable type that defined each song. A. MO6GG variant of MO6ST from male SL1, Slave Lake, AB. B. MO6B1 variant of MO6ST from SL20, Slave Lake, AB. C. MO6E variant of MO6ST from SL19, Slave Lake, AB. D. MO16BB variant of MO16ST from ED1, Edson, AB. E. MO16BB variant of MO16ST from DM4, Duck Mtn., MB. F. MO16B variant of MO16ST from DM19, Duck Mtn., MB. G. MO24 variant of MO24ST from ON57, Dorion, ON. H. MO24 variant of MO24ST from DO29, Dorion, ON. I. MO24A variant of MO24ST from DO8, Dorion, ON. J. MO1 variant of MO1ST from QC16, Lebel, QC. K. MO1D variant of MO1ST from Lebel, QC. L. MO1GTent variant of MO1ST from LE7, Lebel, QC.

Table 3. The percent of syllable variants in each regiolect that fluctuated, declined, increased, were stable, new in 2005 – 2009, new in 2017 – 2019. *n* – number of different syllable variants in each regiolect.

Regiolect	Fluctuates	Declines	Increases	Stable	New in 2005-2009[a]	New in 2017-2019
Western	11.5%	47.0%	15.0%	0.0%	62.0%	24.5%
	(*n* = 200)	(*n* = 200)	(*n* = 200)	(*n* = 200)	(*n* = 71)	(*n* = 200)
Eastern	13.0%	42.2%	8.5%	0.0%	48.3%	31.1%
	(*n* = 270)	(*n* = 270)	(*n* = 270)	(*n* = 270)	(*n* = 89)	(*n* = 270)
Nova Scotia	5.3%	57.9%	7.9%	0.0%	84.2%	31.6%
	(*n* = 38)	(*n* = 38)	(*n* = 38)	(*n* = 38)	(*n* = 19)	(*n* = 38)
Newfoundland	12.4%	48.7%	5.3%	0.0%	69.6%	32.7%
	(*n* = 113)	(*n* = 113)	(*n* = 113)	(*n* = 113)	(*n* = 56)	(*n* = 113)
Average[b]	10.5%	48.9%	9.2%	0.0%	66.0%	30.0%

[a] - Sample sizes are the number of syllable variants only from localities in 2005-2009 that were also sampled in 1983-1988.
[b] - *n* = 4, number of regiolects.

Clinal variation and isolation by distance for syllable types and variants, during each time period were confirmed by Mantel tests. Mantel tests produced positive correlations between Euclidean distances based on syllables and geographic distances for the 1983-1988 (One-tailed Mantel test, $r = 0.31$, $P = 0.020$), 2005-2009 (One-tailed Mantel test, $r = .56$, $P = 0.001$) and 2017-2019 (One-tailed Mantel test, $r = 0.56$, $P = 0.001$) time periods. The correlations were slightly higher in the 2005-2009 and 2017-2019 time periods.

Although the acoustic structure of syllable types, geographic boundaries of regiolects based on syllable types, and the admixture zone were stable over time, there was change in the frequencies of syllable variants in each regiolect. Three different types of change occurred among the syllable variants: fluctuations, declines, and increases (Table 3). Most syllable variant frequencies declined in every regiolect at an average of 48.9%. Syllable variants with fluctuating frequencies or increasing frequencies were comparatively lower. Although average declines in syllable variants were high, they were constantly being replaced by new syllable variants throughout the study period. The average appearance of new variants among the regiolects was 66% in 2005 – 2009. The average appearance of new variants at the end of the study was lower at 30.0% but the time period for replacements from 2005-2019 was shorter than from 1983-2009. No variants in any of the regiolects were stable over time.

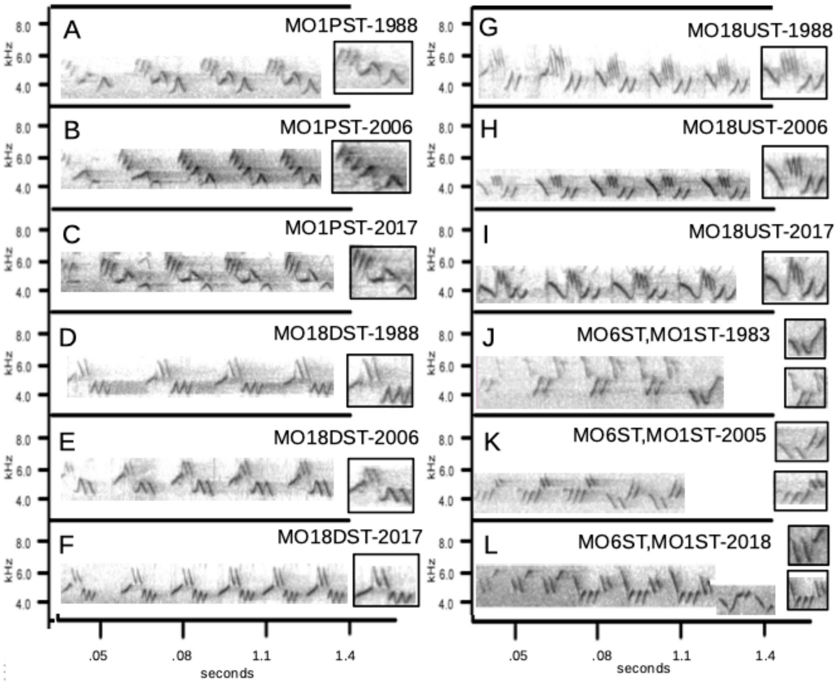

Figure 4. Songs from each time period from the Nova Scotia (MO1PST) and Newfoundland (MO18DST, MO18UST) regiolects and the West-East admixture zone by time period. Enlarged images are variants of the syllable type that defined each song. A. MO1PP variant of MO1PST from male WC6, Wreck Cove, NS. B. MO1PP variant of MO1PST from WC4, Wreck Cove, NS. C. MO1PP2 variant of MO1PST from WC6, Wreck Cove, NS. D. MO18A variant of MO18DST from PS1, Port Saunders, NL. E. MO18D variant of MO18DST from PS28, Port Saunders, NL. F. MO18A variant of MO18DST from BB2, Bellburns, NL. G. MO18L variant of MO18UST from PS4, Port Saunders, NL. H. MO18HH variant of MO18UST from PS27, Port Saunders, NL. I. MO18B2 variant of MO18UST from PS7, Port Saunders, NL. J. MO1ST x MO6ST hybrid song from ON16, Calstock, ON. K. MO1ST x MO6ST hybrid song from DO27, Dorion, Ontario. L. MO1ST x MO6ST hybrid song from DO27, Dorion, Ontario.

The agent-based model indicates content bias is consistent with these patterns. Figure 8 shows the posterior distributions for population size (A), innovation rate (B), frequency bias (C), and content bias (D) from simulation-based inference with the agent-based model. Contrary to our prediction, the temporal dynamics across the four regiolects are consistent with a strong content bias rather than unbiased copying. The median estimate for content bias is 1.96, with a posterior that clearly departs from neutrality ($\beta = 0$). The results for frequency bias are less clear. The median estimate for frequency bias is 0.88, which is consistent with anti-conformity, but the posterior is wide and has a great deal of probability mass (19.7%) above neutrality ($\alpha = 1$). We think that the posterior distribution is too wide to make conclusions about the presence or absence of anti-conformity in this species ($\alpha < 1$), but we can at least tentatively rule out conformity ($\alpha > 1$). The median estimate for initial population size is $N_B = 12{,}692$, although this parameter had the most prediction error during simulation-based inference. Innovation rate has a strong signal with a median estimate of 0.044 (95% CI: 0.021-0.073), which is around the range estimated for House Finches (μ = 0.0067; 95% CI: 0.0011-0.17) and Swamp Sparrws (*Melospiza georgiana*) (μ = 0.0185; 95% CI: 0.0062-0.047) with similar methods (Youngblood and Lahti 2022, Lachlan et al. 2018). This suggests that Mourning Warblers learn song with high fidelity. Finally, the posterior probability for the number of demonstrators (not plotted because it is not continuous) is 0.62 at D = 4 and 0.14 at D = 3 and D = 5, indicating that Mourning Warblers likely learn song from four demonstrators. Simulations show that the fitted parameter values lead to frequency distributions of syllable types close to the observed data (Supplementary Material Figure S4).

A low innovation rate will increase cultural stability, as fewer new syllable types are introduced into the population over time. To assess whether content bias has a similar effect, we ran 100 posterior simulations under two conditions: (1) using the median estimates from each parameter, including the strong content bias, and (2) using the median estimates from each parameter, excluding content bias by setting β = 0. The strong content bias appears to increase cultural stability by reducing turnover in syllable types (Figure S6).

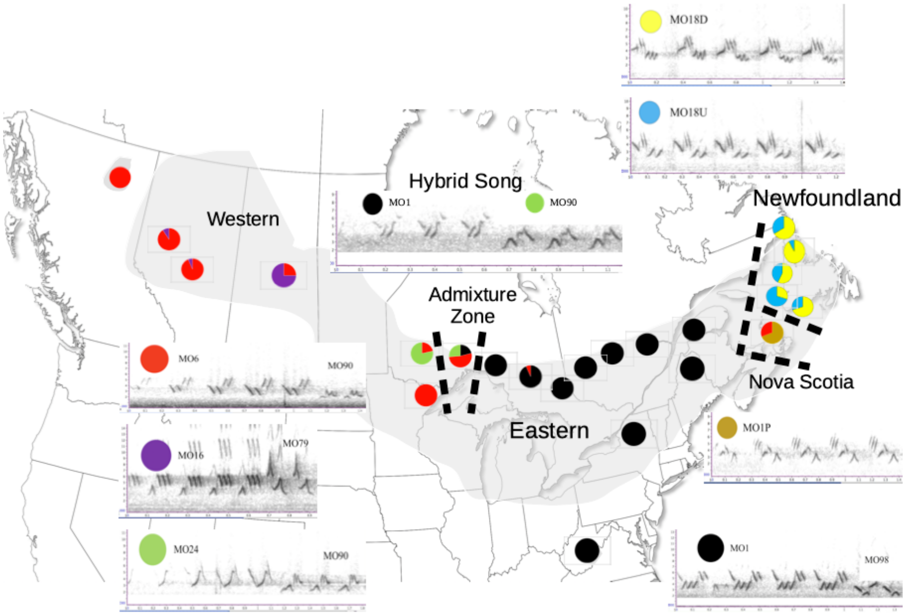

Figure 5. Pie charts of the relative frequency of song types at each study locality from 1983-1988. Each song type is named after the dominant syllable type that occurred most often in each song. Red - Western MO6 song type, Purple - Western MO16 song type, Green – Western MO24 song type, Black – Eastern MO1 song type, Gold – Nova Scotia MO1P song type, Blue – Newfoundland MO18U song type, Yellow – Newfoundland MO18D song type.

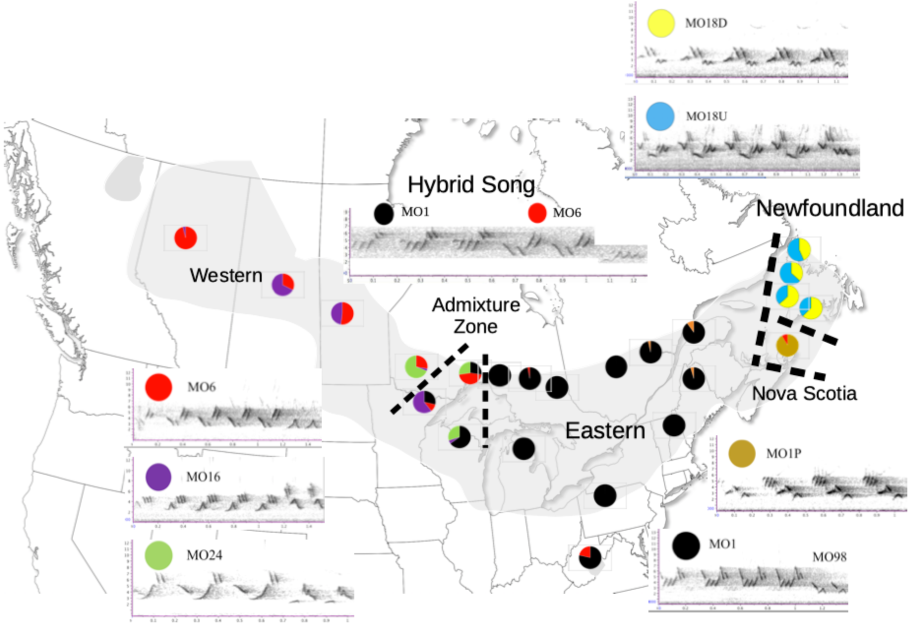

Figure 6. Pie charts of the relative frequency of song types at each study locality from 2005-2009. Each song type is named after the dominant syllable type that occurred most often in each song. Red - Western MO6 song type, Purple - Western MO16 song type, Green – Western MO24 song type, Black – Eastern MO1 song type, Gold – Nova Scotia MO1P song type, Blue – Newfoundland MO18U song type, Yellow – Newfoundland MO18D song type.

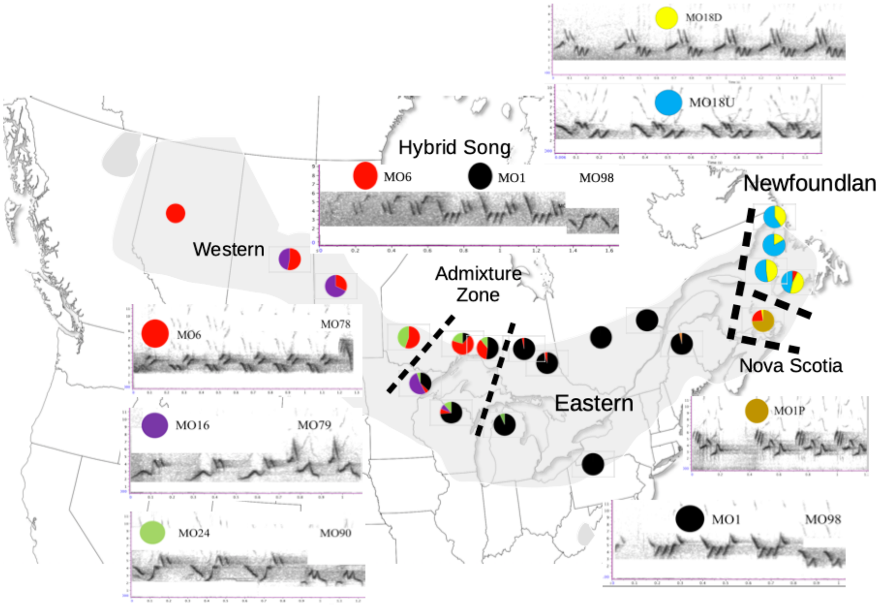

Figure 7. Pie charts of the relative frequency of song types at each study locality from 2017-2019. Each song type is named after the dominant syllable type that occurred most often in each song. Red - Western MO6 song type, Purple - Western MO16 song type, Green – Western MO24 song type, Black – Eastern MO1 song type, Gold – Nova Scotia MO1P song type, Blue – Newfoundland MO18U song type, Yellow – Newfoundland MO18D song type.

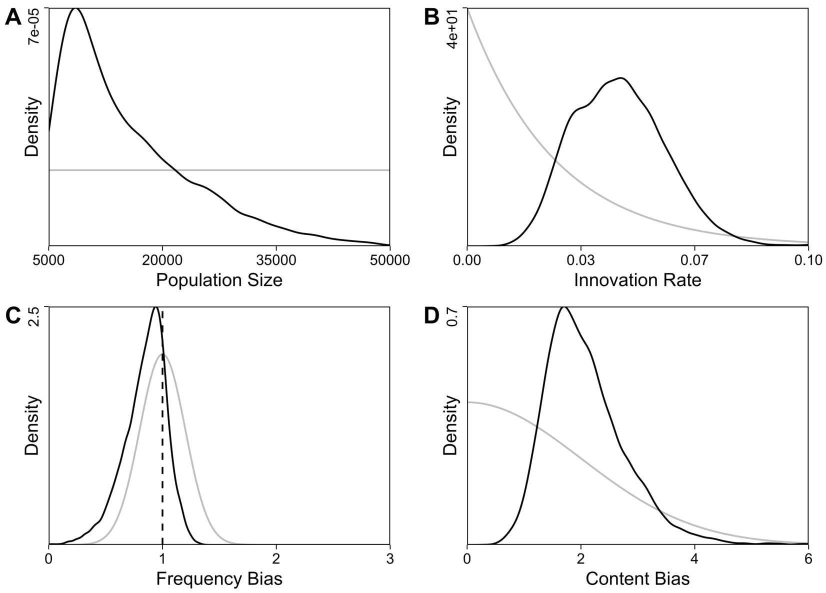

Figure 8. The posterior distributions for initial population size (N_B, panel A), innovation rate (μ, panel B), frequency bias (α, panel C) and content bias (β, panel D) across the four regiolects, plotted against the priors in grey. The expected parameter value at neutrality for frequency bias (α = 1) is marked with the dashed vertical lines. β = 0 is neutrality for content bias.

Table 4. The median, 95% credible interval and standard deviation for each parameter included in the Bayesian logistic regression. Concavity is the number of changes in the sign of the slope-1s. 95% credible intervals that do not overlap with 0 indicate a statistically significant effect and are marked with an asterisk.

Feature	Mean	SD	95% CI	
Duration (s)	-0.21	0.41	[-0.88, 0.46]	
Max. frequency (Hz)	-0.44	0.49	[-1.25, 0.35]	
Min. frequency (Hz)	0.09	0.44	[-0.63, 0.83]	
Concavity	0.87	0.32	[0.34, 1.41]	*

The results of the Bayesian logistic regression are shown in Table 4. The persistence of syllable types from 1985 to 2019 was positively predicted by concavity, with no significant effects for duration, minimum frequency, and maximum frequency. This suggests that the strong content bias detected by the simulation-based inference may be targeted towards syllable complexity. All effective sample sizes are greater than 1,000 and all R-hat values are equal to 1, indicating that the chains have converged. Several t-tests were run to follow up on the effect of syllable concavity. Novel syllable types from 2005 and 2019 have significantly higher concavity than those that existed in the 1980s ($t = 3.12$, $p = 0.002$). Additionally, the overall concavity of syllable types in 2005 is significantly higher than in 1985 ($t = 2.37$, $p = 0.019$), but syllable types in 2019 do not have higher concavity than those in 2005 ($t = -1.22$, $p = 0.222$). More complex syllables appear to have an advantage, but this only caused the overall syllable complexity to increase between 1985 and 2005.

The physical parameters of the songs within each regiolect remained largely unchanged over time. The average PC1 scores from each time period were not significantly different in each regiolect after a sequential Holm-Bonferroni correction except for the MO16 song type. The ANOVAs of PC1 scores for the Western MO24 songs ($F = 2.56$, df = 2 and 92, $P = 0.83$), Nova Scotia MO1P songs ($F = 2.52$, df = 2 and 59, $P = 0.089$) and Newfoundland MO18 songs ($F = 0.27$, df = 2 and 223, $P = 0.76$) were not significant. The ANOVA of PC1 scores for MO6 songs from the western regiolect was significant ($F = 5.19$, df = 2 and 230, $P = 0.006$) but the post-hoc SNK test did not find differences among the means for the 1983-1988, 2005-2009 and 2017-2019 time periods ($P > 0.05$). ANOVA for the MO1 Eastern song was significant ($F = 4.26$, df = 2 and 466, $P = 0.015$) but not after a sequential Holm-Bonferroni correction (greater than the corrected alpha of $P < 0.013$).

Only PC1 scores from 2005-2009 differed from the 2017-2019 time period for western MO16 songs (ANOVA, $F = 12.42$, df = 1 and 82, $P = 0.001$). Data from 1983-1988 for western MO16 songs were lacking due to small sample sizes.

The results from Mantel tests of the relationship between song distances based on physical parameters of song from the PC scores and geographic distances showed clinal variation and isolation by distance. The results were positive and significant for 1983-1988 (One-tailed Mantel test, $r = 0.64$, $P = 0.001$), 2005-2009 (One-tailed Mantel test, $r = 0.56$, $P = 0.001$) and 2017-2019 (One-tailed Mantel test, $r = 0.61$, $P = 0.001$) time periods. Physical parameters of song diverged gradually with increasing geographic distance between localities, similar to the results of the Mantel tests for syllables.

Statistical tests of the Acoustic Adaptation Hypothesis produced mixed and different results for the 2005-2009 and the 2017-2019 time periods. The multiple regressions for PC1 scores and for syllable rate for 2005-2009 were significant but the results were ambiguous (Table 5). PC1 scores had positive relationships with NDVI, NDMI and increased with increasing vegetation density. There was a negative relationship between PC1 and GNDVI and positive relationship with BSI which coincided with increases in vegetation density. The r^2 value was low indicating that the model and independent variables accounted for only 4.9% of the variation in song frequencies, syllable and song durations represented by PC1 scores. The results for syllable rate were similar with significant but mixed results. Syllable rate increased with increasing vegetation for NDVI and NDMI but decreased with increasing vegetation as measured by GNDVI and BSI. Only 5.9% of the variation could be explained by the Landsat variables (Table 5). There was no significant relationship between Landsat variables and PC1 scores for the 2017-2019 data. Syllable rate in 2017-2019 was inversely related to all Landsat variables except GNDVI. The relationships were significant but the r^2 was only 4.1% (Table 5).

Table 5. Results of multiple regression of PC1 scores, log-transformed syllable rate, and Landsat variables for the 2005-2009 and 2017-2019 time periods: unstandardized *R* coefficients for the long-transformed Landsat variables (with coefficient's SE in parentheses), significance levels after Holm-Bonferroni correction (Holm 1979) and r^2 values in percentages.

Song variables	**Landsat variables**					
2005-2009	NDVI	NDMI	GNDVI	BSI	Significance	r^2
PC1[a]	2.10	2.94	-2.41	3.14	$P < 0.000$	4.90%
	(0.65)	(1.64)	(0.58)	(2.01)		
SyllRatelg[b]	0.18	0.38	-0.22	0.45	$P < 0.000$	5.90%
	(0.05)	(0.13)	(0.05)	(0.16)		
2017-2019	NDVI	NDMI	GNDVI	BSI	Significance	r^2
PC1[c]	-0.56	-7.89	0.28	-8.73	$P = 0.730$	1.60%
	(5.00)	(4.60)	(5.12)	(7.00)		
SyllRatelg[d]	-0.23	-0.69	0.12	-0.60	$P < 0.000$	4.10%
	(0.42)	(0.38)	(0.43)	(0.58)		

[a] – $F = 6.53$, df = 4 and 508.
[b] – $F = 7.91$, df = 4 and 508.
[c] – $F = 2.15$, df = 4 and 522.
[d] – $F = 5.63$, df = 4 and 523.

DISCUSSION

Some major findings of our study were that song types, syllable types, macrogeographic pattern of regiolects, song parameters, and geographic boundaries of the regiolects have been stable throughout the breeding range for 36 yrs. The Western, Eastern, Nova Scotia and Newfoundland regiolects described in 2011 (Pitocchelli 2011) also existed in 1983-1988 and during sampling from 2017-2019. The high survivorship of the syllable types and syllable variants showed that their acoustic structures in each regiolect were maintained over time. The lower survivorship of entire songs is probably because songs are longer and more complex with a unique syntax compared to syllables and are less likely to be copied in the exact same sequence over time. The much higher turnover rate in songs compared to syllables has been reported in several other species including the Corn Bunting (*Emberiza calandra*, Holland et al. 1996), unaccented songs of the Chestnut-sided Warbler (*Setophaga pensylvanica*, Byers et al. 2010), and the House Finch (Ju et al. 2019). The stability of Mourning Warbler regiolects are in contrast to species with rapid turnover in one to five yrs in dialects like Yellow-rumped Caciques (*Cacicus cela*, Trainer 1989) or Village Indigobirds (*Vidua chalybeate*, Payne 1985). Mourning Warbler regiolects also survived longer than species with moderate turnover rates in dialects like the Hermit Warbler Type I songs from northern Oregon over 11 yrs (*Setophaga occidentalis*, Janes and Ryker 2013), Corn Buntings over 14 yrs (Holland et al. 1996), unaccented songs of Chestnut-sided Warblers over 19 yrs (Byers et al. 2011). Mourning Warbler regiolects are more like the longer-term survivorship of dialects of the Brown-headed Cowbird (*Molothrus ater*, O'Loghlen et al. 2011), Rufous-collared Sparrow (*Zonotrichia capensis,* Handford 1988), and White-crowned Sparrow (*Zonotrichia leucophrys*, Harbison et al. 2004).

The persistence of the geographic boundaries of regiolects of the Mourning Warbler across the breeding range is notable considering that expansions and contractions in dialect boundaries over time are common (McGregor and Thompson 1988, Harbison et al. 1999). Rapid turnover in song dialect structure of the Village Indigobirds was accompanied by boundary changes (Payne 1985). Gradual changes in dialect boundaries have occurred in some populations of White-crowned Sparrows from Washington (Chilton and Lein 1996) and Corn Buntings in North Cornwall England (Holland et al. 1996). Stable regiolect boundaries of the Mourning Warbler are comparable to some dialects of California White-crowned Sparrows (Trainer 1983) and the Rufous-collared Sparrow (Handford 1988). One reason why boundaries change is due to changes in land use practices that disturb breeding habitat and disrupt dialects while boundaries in undisturbed areas remain unaffected (Holland et al. 1996). Mourning Warblers breed in ephemeral, disturbed second growth of boreal forests. The boreal forest and availability of disturbed second growth is constantly changing due to wildfires and human disturbances like logging, mineral and gas exploration (Pitocchelli 2020). But the boundaries of the regiolects and the admixture zone in central Canada have remained intact. The effects of habitat change vs. stability on dialect boundaries does not explain the lack of change of regiolect boundaries of the Mourning Warbler.

Cultural evolutionary dynamics within a population play an important role in the emergence of geographic variation in song (Derryberry 2011). Our simulations indicate that content bias, possibly targeted towards syllable complexity, is the most likely driver of the dynamics within regiolects across 36 years. There is also a weak signal of anti-conformity in the data, although the posterior distribution is too wide for us to make definitive conclusions. Additionally, our simulation shows that this stability can persist despite low levels of dispersal between regiolects (e.g., average 4.71% dispersal of Western songs into the Eastern regiolect, Supplementary Material Table S4, Figure S3) as specified in our simulation. Although new syllable variants continued to appear over time, almost every new syllable was a variant of a model syllable type unique to each regiolect (e.g., variants of the MO1ST in the Eastern regiolect, Supplementary Material Figure S1). Birds may have created new syllable variants by improvisation or copying errors but they rarely produced profoundly new syllable types and different song types that could lead to new regiolects. Only 2% of syllables in 2017 – 2019 ($n = 983$) were new syllable types not observed in the 1983 – 1988 or 2005 – 2009 time periods. The estimated innovation rate from the agent-based modeling was also low (~4%) and around the range of other species assessed with similar methods (Youngblood and Lahti 2022, Lachlan et al. 2018), suggesting that Mourning Warbler song is learned with high fidelity. A low innovation rate combined with a strong content bias acting on different syllable types within each regiolect may explain their stability over time. Youngblood and Lahti (2022) found content bias for syllable complexity in the House Finch but their study was from one region in the breeding range. To our knowledge, these results for the Mourning Warbler are the first to show evidence for content bias throughout the entire breeding range.

Evidence that lower frequencies, longer song elements, and slower delivery of repeated elements occurred in closed habitats with dense vegetation in 2005-2009 and 2017-2019 was inconsistent and equivocal. In 2005-2009, there was a significant relationship between PC1 scores and the Landsat variables but these results did not support predictions from the Acoustic Adaptation Hypothesis. If songs with lower frequencies, longer songs and longer syllables were found on breeding territories with dense, closed vegetation then we should have expected a negative relationship between PC1 scores and NDVI, NDMI, GNDVI and a positive relationship with BSI. Our results were contradictory. Higher frequency songs and shorter songs were associated with increasing vegetation density and positive coefficients for NDVI and NDMI in contradiction to the Acoustic Adaptation Hypothesis. But, in the same regression model, lower frequency songs and longer songs were associated with dense vegetation, a negative coefficient for GNDVI, and positive coefficient for BSI in support of acoustic adaptation. The multiple regression model for syllable delivery measured as syllable rate produced similar contradictory results. Syllable rate increased with increasing vegetation for NDVI and NDMI but decreased with increasing vegetation for GNDVI and BSI. The low r^2 values also suggested these models and the Landsat variables explained little of the variation in song parameters in 2005-2009 (Table 5).

There was no evidence that songs with lower frequencies, longer syllables or longer songs occurred on territories with increasing vegetation density in 2017-2019. The relationship between PC1 scores and Landsat indices was not significant (Table 5). The relationship between syllable rate and Landsat variables from 2017-2019 was significant but equivocal. Syllable rate was negatively related to NDVI, NDMI and therefore decreased with increasing vegetation density as predicted by the Acoustic Adaptation Hypothesis. However, the positive relationship with GNDVI and negative relationship with BSI indicated syllable rate was increasing with increasing vegetation density. These conflicting results and low r^2 of 4% do not support acoustic adaptation in 2017-2019. For multiple regression analyses in both time periods, 50% of the relationships supported the Acoustic Adaptation Hypothesis while 50% of the relationships did not or the multiple regression model was not significant. These results were also accompanied by low r^2 values. We concluded that evidence was insufficient to support acoustic adaptation of Mourning Warbler songs to the environment.

Species with little or no differences in breeding habitat throughout the breeding range tend to lack or have minor geographic differences in physical parameters of song (Boncoraglio and Saino 2007, Brumm and Naguib 2009). In contrast, song parameters vary in species that breed in different habitats in different parts of the breeding range (Ruegg et al. 2006, Smith et al. 2013). Lack of support for the Acoustic Adaptation Hypothesis for the Mourning Warbler is likely due to habitat selection and similarity in breeding habitat across the breeding range. We did not investigate other factors such as morphological variation and ambient noise that may also affect song transmission (Ryan and Brenowitz 1985, Podos et al. 2004, Derryberry 2009). Pitocchelli (1992) did not find significant geographic variation in size such as weight, skull, or other skeletal measurements from the breeding range so the effects of morphological variation on sound transmission is unlikely. The effect of ambient noise has yet to be investigated.

Derryberry (2011) noted the important role that temporal changes in songs within populations plays in the evolution of geographic variation among populations. Temporal change in the physical parameters of Mourning Warbler songs was negligible across regiolects. Immigration of alien males and their songs into different populations and hybrid songs could also produce new variation within regiolects, but their rare occurrence failed to produce change anywhere in the breeding range (Supplementary Materials). In contrast, syllable variants were parts of the songs that changed the most and continued to change over time in all regiolects. The emergence of a new syllable type and song from novel syllable variants via innovation or improvisation has the greatest potential to generate variation in songs and regiolects of the Mourning Warbler. Sweeping changes in songs due to the rapid spread of new variants has been reported in the doublet-ending song of the White-throated Sparrow (*Zonotrichia, albicolis*, Otter et al. 2020). The spread of the new variant is thought to first have occurred through tutoring on the wintering grounds, later spreading rapidly across the breeding range. Change caused by new variants could occur within regiolects of the Mourning Warbler but it would be highly unlikely to occur at such a rapid rate. Mourning Warblers do not sing on the wintering grounds (Pitocchelli 2020). The rarity of successful, long-term establishments of new variants combined with a high rate of decline in variants (Table 3) may further constrain the rate of change in Mourning Warblers. Change in Mourning Warbler regiolects involving novel variants would likely occur over a much longer time frame.

Acknowledgments

We thank Shawn Bugbee, Ken Burrel, Hannah Stevens, Darroch Whitaker for providing valuable information on breeding sites across the United States and Canada. The Interlibrary Loan staff at the Geisel Library assisted with the literature searches. We would also like to thank Stony Brook Research Computing and Cyberinfrastructure, and the Institute for Advanced Computational Science at Stony Brook University for access to the SeaWulf computing system.

Funding Statement

This study was funded in part by the Saint Anselm College Summer Research Grant program and a sabbatical leave for J. P. Access to the SeaWulf computing system was made possible by grants from the National Science Foundation (#1531492 and Major Research Instrumentation award #2215987), with matching funds from Empire State Development's Division of Science, Technology and Innovation (NYSTAR) program (contract C210148).

Author Contributions

J.P. conceived of the idea for the study and conducted the field work. M.Y. developed the agent-based model and conducted simulation-based inference and related supplemental analyses. All authors participated in data analyses and writing or editing different aspects of the paper.

Data Availability

Songs and data are archived in Open Science Framework at https://doi.org/10.17605/OSF.IO/7G4C9. All code and data used in the agent-based modeling analysis can be found at Zenodo (https://doi.org/10.5281/zenodo.10698484) and GitHub repository (https://github.com/masonyoungblood/MourningWarblers).

LITERATURE CITED

Aplin, L. M. 2019. Culture and cultural evolution in birds: a review of the evidence, Animal Behaviour 147:179-187. https://doi.org/10.1016/j.anbehav.2018.05.001.

Barrett, B. J. (2019). Equifinality in empirical studies of cultural transmission. Behavioural Processes 161:129–138. https://doi.org/10.1016/j.beproc.2018.01.011.

Bioacoustics Research Program, Cornell Lab of Ornithology (2014). Raven Pro: Interactive sound analysis software 1.5. http://www.birds.cornell.edu/raven.

Bloem-Reddy, B., & Teh, Y. W. (2020). Probabilistic symmetries and invariant neural networks. Journal of Machine Learning Research 21:1–6.

Boncoraglio, G., and N. Saino. (2007). Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. Functional Ecology 21:134-142.

Bonnet, E., Van de Peer Y. (2002). zt: a software tool for simple and partial Mantel tests. Journal of Statistical Software. 7:1–12.

Brumm, H., and M. Naguib. (2009). Environmental acoustics and the evolution of bird song. Advances in the Study of Behavior 40:1-33.

Byers, B. E., K. L. Belinsky, and R. A. Bentley. (2010). Independent cultural evolution of two song traditions in the Chestnut-Sided Warbler. The American Naturalist 176:476-489.

Chilton, G., and M. R. Lein. (1996). Long-term changes in songs and song dialect boundaries of Puget Sound White-Crowned Sparrows. The Condor 98:567–580. https://doi.org/10.2307/1369569.

Cranmer, K., Brehmer, J., Louppe, G. (2020). The frontier of simulation-based inference. Proceedings of the National Academy of Sciences 117:30055-30062. doi:10.1073/pnas.1912789117.

Derryberry, E. P. (2009). Ecology shapes birdsong evolution: variation in morphology and habitat explains variation in White-Crowned Sparrow Song. The American Naturalist 174:24-33.

Derryberry, E. P. (2011). Male response to historical and geographical variation in bird song. Biology Letters 7:57-59.

Fayet, A. L., J. A. Tobias, R. E. Hintzen, and N. Seddon. 2014. Immigration and dispersal are key determinants of cultural diversity in a songbird population. Behavioral Ecology 25:744 – 753. doi:10.1093/beheco/aru047.

Goodrich, B., J. Gabry, I. Ali, and S. Brilleman, (2020). rstanarm: Bayesian applied regression modeling via Stan. (R package 2.21.1). https://mc-stan.org/rstanarm.

Harbison, H., D. A. Nelson, and T. Hahn. (1999). Long-term persistence of song dialects in the Mountain White-Crowned Sparrow. The Condor 101:133-148.

Handford, P. (1988). Trill rate dialects in the Rufous-crowned Sparrow, *Zonotrichia capensis*, in northwestern Argentina. Canadian Journal of Zoology 66:2658-2670.

Holland, J., McGregor, P. K., and Rowe, C. L. (1996). Changes in microgeographic song variation of the Corn Bunting *Miliaria calandra*. Journal of Avian Biology 27:47-55.

Holm, S. (1979). A simple sequential rejective multiple test procedure. Scandinavian Journal of Statistics 6:65-70.

Hsieh, T. C., Ma, K. H., and Chao, A. (2016). iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution, *7*:1451–1456. https://doi.org/10.1111/2041-210X.12613

IBM Corp. (2020). IBM SPSS Statistics for Macintosh 27.0. Armonk, NY.

Janes, S. W., and L. Ryker. (2013). Rapid change in a Type I song dialect of Hermit Warblers *(Setophaga occidentalis)*. The Auk 130:30–35. https://doi.org/10.1525/auk.2012.11273

Ju, C., F. C. Geller, P. C. Mundinger, and D. C. Lahti (2019). Four decades of cultural evolution in House Finch songs. The Auk: Ornithological Advances 136:1–18. https://doi.org/10.1093/auk/uky012

Kandler, A., and A. Powell. (2018). Generative inference for cultural evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 373(1743). https://doi.org/10.1098/rstb.2017.0056

Lachlan, R. F., O. Ratmann, and S. Nowicki. (2018). Cultural conformity generates extremely stable traditions in bird song. Nature Communications 9. https://doi.org/10.1038/s41467-018-04728-1

Lynch, A. 1996. The population memetics of birdsong. In Ecology and evolution of acoustic communication in birds. (eds D. E. Kroodsma & E. H. Miller), pp. 181–197. Ithaca, NY: Cornell University Press.

Lynch, A. G. M. Plunkett, A. J. Baker, and P. F. Jenkins. (1989). A Model of Cultural Evolution of Chaffinch Song Derived with the Meme Concept. The American Naturalist 133:634-653.

Mantel N. (1967). The detection of disease clustering and a generalized regression approach. Cancer Research. 27:209–220.

McGregor, P. K., and D. B. A. Thompson (1988). Constancy and change in local dialects of the Corn Bunting. Ornis Scandinavica 19:153-159.

Morton, G. A. (1975). Ecological sources of selection on avian sounds. American Naturalist 109:605–624.

Nolan Jr, V., E. D. Ketterson, and C. A. Buerkle (2020). Prairie Warbler (*Setophaga discolor*), version 1.0. In Birds of the World (A. F. Poole, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://doi.org/10.2173/bow.prawar.01.

Nguyen, C., A. Chidthaisong, P. K. Diem, and L-Z. Huo. (2021). Modified bare soil index to identify bare land features during agricultural fallow-period in southeast Asia using Landsat 8, Land 10 231:1-17.

O'Loghlen, A. L., V. Ellis, D. R. Zaratziam, L. Merrill, and S. I. Rothstein. (2011). Cultural evolution and long-term song stability in a dialect population of Brown-headed Cowbirds. The Condor 113:449 461.

Otter, K. A., A. Mckenna, S. E. LaZerte, and S. M. Ramsay. 2020. Continent-wide shifts in song dialects of White-Throated Sparrows. Current Biology 30:3231 – 3235. https://doi.org/10.1016/j.cub.2020.05.084.

Payne, R. B. (1985). Behavioral continuity and change in local song populations of Village Indigobirds, *Vidua chalybeata*. Zeitschrift für Tierpsychologie. 70:1-44.

Pitocchelli, J. (1992). Plumage and size variation in the Mourning Warbler (*Oporornis philadelphia*). The Condor 94:198-209.

Pitocchelli, J. (2011). Macrogeographic variation in the song of the Mourning Warbler (*Oporornis philadelphia*). Canadian Journal of Zoology 89:1027-1040.

Pitocchelli, J. (2014). Song discrimination by male Mourning Warblers (*Geothlypis philadelphia*) and Implications for population divergence across the breeding range. Canadian Field-Naturalist 128:408-415.

Pitocchelli, J. (2020). Mourning Warbler (*Geothlypis philadelphia*), version 1.0. In Birds of the World (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://doi.org/10.2173/bow.mouwar.01

Podos, J., S. K. Huber, and B. Taft. (2004). Birdsong: the interface of evolution and mechanism. Annual Review of Ecology, Evolution, and Systematics 35:55–87.

Podos, J., and Warren, P. S. (2007). The evolution of geographic variation in birdsong. Advances in the Study of Behavior. 37:403–458.

Radev, S. T., Mertens, U. K., Voss, A., Ardizzone, L., & Köthe, U. (2020). BayesFlow: Learning complex stochastic models with invertible neural networks. IEEE Transactions onNeural Networks and Learning Systems 33. https://doi.org/10.1109/TNNLS.2020.3042395

Radev, S. T., Schmitt, M., Schumacher, L., Elsemüller, L., Pratz, V., Schälte, Y., Köthe, U., & Bürkner, P.-C. (2023). BayesFlow: Amortized Bayesian workflows with neural networks. The Journal of Open Source Software 8:5702. https://doi.org/10.21105/joss.05702

Reidy, J. L. 2007. Golden-cheeked Warbler nest success and nest predators in urban and rural landscapes. Ph.D. Thesis. University of Missouri-Columbia.

Rendell, L., L. Fogarty, W. J. E. Hoppitt, T. J. H. Morgan, M. M. Webster, and K. N. Laland, (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. Trends in Cognitive Sciences 15:68–76. https://doi.org/10.1016/j.tics.2010.12.002.

Ruegg, K., H. Slabbekoorn, S. Clegg, and T. B. Smith. (2006). Divergence in mating signals correlates with ecological variation in the migratory songbird, Swainson's thrush (*Catharus ustulatus*). Molecular Ecology 15:3147–3156.

Ryan, M. J., and E. A. Brenowitz..(1985). The role of body size, phylogeny, and ambient noise in the evolution of bird song. American Naturalist 126:87-100.

Schumacher, L., Bürkner, P.-C., Voss, A., Köthe, U., & Radev, S. T. (2023). Neural superstatistics for Bayesian estimation of dynamic cognitive models. Scientific Reports 13:13778. https://doi.org/10.1038/s41598-023-40278-3

Slabbekoorn, H., and T. B. Smith. (2002). Birdsong, ecology and speciation. Philosophical Transactions of the Royal Society B: Biological Sciences 357:493–503.

Smith, T. B., R. J. Harrigan, A. N. G Kirschel, W. Buermann, S. Saatchi , D. T. Blumstein, S. R. de Kort, and H. Slabbekoorn. (2013). Predicting bird song from space. Evolutionary Applications 6:865–874.

Soha, J. A., and P. Marler. 2000. A species-specific acoustic cue for selective song learning in the White-crowned sparrow. Animal Behaviour 60:297 –3 06.

Thi Loi, D., T-Y Chou, and F. Yao-Min. (2017). Integration of GIS and Remote Sensing for Evaluating Forest Canopy Density Index in Thai Nguyen Provence, Vietnam., International Journal of Environmental Science and Development, Vol. 8, No. 8.

Trainer, J. M. (1983). Changes in song dialect distributions and microgeographic variation in song of White-Crowned Sparrows *(Zonotrichia leucophrys nuttalli)*. The Auk 100:568–582. https://doi.org/10.1093/auk/100.3.568

Trainer, J. M. (1989). Cultural evolution in song dialects of Yellow-rumped Caciques in Panama. Ethology 80:190-204.

USGS (2022). What are the best Landsat spectral bands for use in my research? https://earthexplorer.usgs.gov/.

van Dongen, W. F. D., and R. A. Mulder. Habitat density, song structure and dialects in the Madagascar Flycatcher *Terpsiphone mutata*. Journal of Avian Biology 37:349-356.

von Krause, M., Radev, S. T., & Voss, A. (2022). Mental speed is high until age 60 as revealed by analysis of over a million participants. Nature Human Behaviour 6:700–708.

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2023). Transformers in time series: A survey. Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. 6778–6786. https://doi.org/10.24963/ijcai.2023/759

Williams, H. (2021). Mechanisms of cultural evolution in the songs of wild bird populations. Frontiers in Psychology 12:643343. doi: 10.3389/fpsyg.2021.643343

Youngblood, M., and D. Lahti. (2022). Content bias in the cultural evolution of house finch song. Animal Behaviour, 185:37–48. https://doi.org/10.1016/j.anbehav.2021.12.012

Supplementary Material

Table S1. Study localities, sample sizes for the 1983-1988, 2005-2009, 2017-2019 time periods, latitude and longitude.

Localities	1983-1988	2005-2009	2017-2019	Latitude	Longitude
BC - Fort Nelson, BC	13	n/a	n/a	58.8466°N	122.8876°W
SL - Slave Lake, AB	18	26	33	55.3918°N	114.9199°W
CAB - Central Alberta	19	n/a	n/a	53.1266°N	115.4456°W
CL - Candle Lake, SK	8	7	n/a	53.7035°N	105.3859°W
HB – Hudson Bay, SK	4	36	32	52.9743°N	102.3466°W
DM - Duck Mtn., MB	n/a	20	27	51.6848°N	100.9298°W
ND - Turtle Mountains, ND	n/a	6	n/a	48.9890°N	100.4566°W
RE - Remer, MN	1	24	18	47.0738°N	94.1000°W
MI - Dryden, ON	15	30	27	50.1060°N	92.7236°W
DO - Dorion, ON	18	30	45	49.0268°N	88.6168°W
GE - Geraldton, ON	6	7	27	49.4729°N	87.1912°W
HE - Hearst, ON	16	34	33	49.9308°N	83.9615°W
StG - St. Germain, WI	n/a	19	24	45.9663°N	89.5939°W
GL - Glennie, MI	n/a	23	30	44.6570°N	83.5514°W
CO - Cochrane, ON	15	28	23	49.0767°N	80.3386°W
LE - Lebel, QC	18	49	35	49.2346°N	77.2952°W
GI - Girardville, QC	8	30	19	49.0373°N	72.6277°W
SP - Rimouski, QC	6	19	n/a	48.5939°N	67.5611°W

WV - Cranberry Mtn. WV	5	11	n/a	36.2068°N	80.2279°W
KA - Kane, PA	n/a	22	17	41.7915°N	78.7355°W
NY - Chazy, NY	1	n/a	n/a	44.8793°N	73.4782°W
NH - Jefferson, NH	n/a	16	n/a	44.3573°N	71.4719°W
AS - Ashland, ME	19	22	27	46.4360°N	68.3850°W
WC – Wreck Cove, NS	13	24	33	46.5572°N	60.6038°W
ST - Stephenville, NL	10	26	21	48.4076°N	58.3624°W
BU - Burgeo, NL	21	17	28	48.4695°N	57.9567°W
RO - Roddickton, NL	11	23	30	51.0909°N	56.6245°W
PS – Port Saunders, NL	7	30	24	50.4228°N	57.3807°W
RD - Red Bay, NL	5	2	n/a	51.7386°N	56.4085°W

Table S2. Log 10-transformed physical parameters of song and PC score loadings.

Song Variables	PC1	PC2	PC3
DurALg	-0.775	0.579	0.006
MinALg	0.858	0.405	-0.025
MaxALg	0.435	0.764	0.066
DurBLg	-0.795	0.573	-0.037
MinBLg	0.860	0.432	-0.028
MaxBLg	0.437	0.779	0.053
DurCLg	-0.786	0.563	-0.026
MinCLG	0.865	0.416	-0.029
MaxCLg	0.445	0.769	0.050
DURLg	-0.033	0.011	0.992

% of the variation	55.0%	30.0%	5.9%

Dur_Lg – log10-transformed duration of the A, B, C syllables. Min_Lg – log10-transformed minimum frequency of the A, B, C syllables. Max_Lg – maximum frequency of the A, B, C syllables. DURLg – log10-transformed entire song duration. All original variables were log 10 transformed because of the order of magnitude of measurement differences between Hz and s.

Immigration and Hybrid Songs

We conducted elementary statistical analyses of the effects of hybrid songs and immigration on variation within regiolects over time. Hybrid songs are a mix of syllables from different song types (Figures 4 - 7). They are unique, aberrant songs that may introduce new variation into a regiolect. We calculated the percentage of males with hybrid songs in each regiolect and assessed their trends over time. Immigration is another potential source of new variation in a regiolect. Males with alien songs immigrating to a different population introduce new songs that could change the composition of that population's song types (Fayet et al. 2014). There are no studies of site fidelity or dispersal of males returning to the breeding range (Pitocchelli 2020) so we used the percentage of males with alien songs in each regiolect (e.g., per cent of birds with Western songs recorded in the Eastern regiolect) as a simple measure of immigration for each time period. We used those values to calculate an average immigration percentage in each regiolect for all three time periods.

All cases of hybrid songs occurred in the Western and Eastern regiolects and they were observed in each time period. These songs contained a mix of Western and Eastern syllables (Figures 4 - 7). The percentages of hybrid songs within each time period were low, varied little over time, and did not change the dominant song types in any regiolects (Table S3). The immigration of alien songs across regiolect boundaries was also very low. The highest average levels of exchange occurred between the Western and Eastern regiolects. The ranges of the percentages showed fluctuations over time (Table S4). Instances of immigration in the Maritimes occurred between the Eastern and Newfoundland regiolects and Nova Scotia and Newfoundland regiolects. Average exchanges were low and immigration did not occur in each time period. Immigration has been an important agent of change within song populations of Great Tits (*Parus major*, Fayet et al. 2014) and Hermit Warblers (*Setophaga occidentalis*, Furnas et al. 2020). However, song exchanges via immigration between Mourning Warbler regiolects were very low compared to these previous studies. It is unlikely that immigration has, or will lead to change in the dominant song types in any of the regiolects.

Furnas, B. J., R. H. Landers, and R. C. K. Bowie. 2020. Wildfires and mass effects of dispersal disrupt the local uniformity of type I songs of Hermit Warblers in California. The Condor 137:1 - 14.

Table S3. Occurrence of hybrid songs in each regiolect over time

Time Period	Western Regiolect	Eastern Regiolect
1983-1988	3.13% (*n* = 96)	1.06 % (*n* = 94)
2005-2009	3.35% (*n* = 179)	0.36% (*n* = 280)
2017-2019	3.30% (*n* = 182)	0.85% (*n* = 235)
Average	3.26%	0.76%

Table S4. Average percentage of alien songs and their destination regiolects over the three time periods[a]. Ranges for the time periods are in parentheses.

Destination

Source	Western	Eastern	Nova Scotia	Newfoundland
Western	n/a	4.71% (1.06% - 10.21%)	0.00%	0.00%
Eastern	6.75% (4.17% - 9.50%)	n/a	0.00%	0.34% (0.00% - 1.02%)
Nova Scotia	0.00%	0.00%	n/a	0.65% (0.00% -1.94%)
Newfoundland	0.00%	0.00%	1.01% (0.00% - 3.03%)	n/a

[a] - The average percentage of alien songs occurring in a new, destination regiolect was the average over the three time periods in each regiolect. Sample sizes used to obtain a percentage of alien songs in each regiolect for each time period are the same as in Table S3.

Syllable Types and Variants

We followed Pitocchelli's (2011) methods for classifying syllable types and variants. Syllable types varied in basic shape and configuration of their notes. Syllable variants shared the basic shape and order of notes but differed in the number of notes or an addition of a new note. For example, there are three variants of the MO1ST below. The MO1M1 variant is missing the first downslur (decreasing frequency) note but retains the remaining notes. The MO1G2Tent retains the first four notes but has a new, tent-shaped note at the end. The MO1B5Low retains the basic MO1ST shape but has five upslur (increasing frequency) notes in the middle of the syllable. We also added survivorship, timing of loss, first time occurrence for these variants to indicate that many are in a state of constant flux.

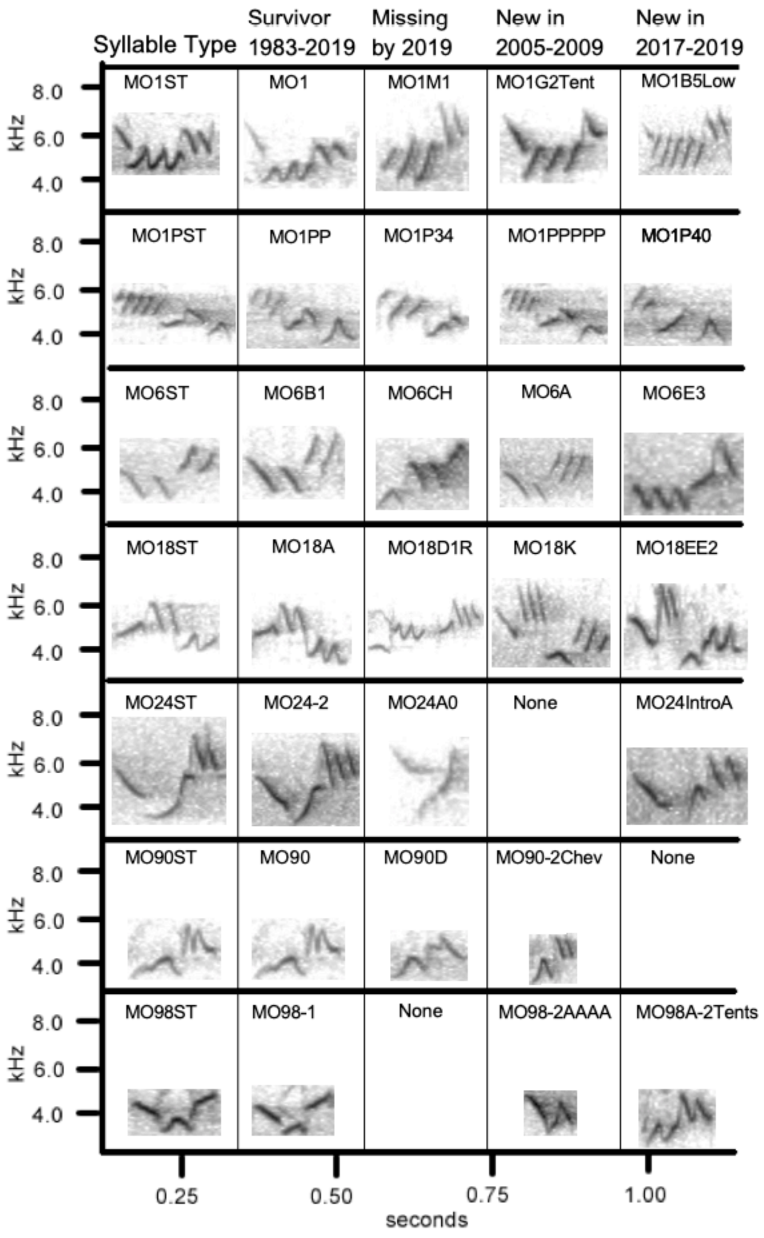

Figure S1. Examples of syllable types, variants of syllable types from Mourning Warbler songs and their occurrence over 36 yrs.

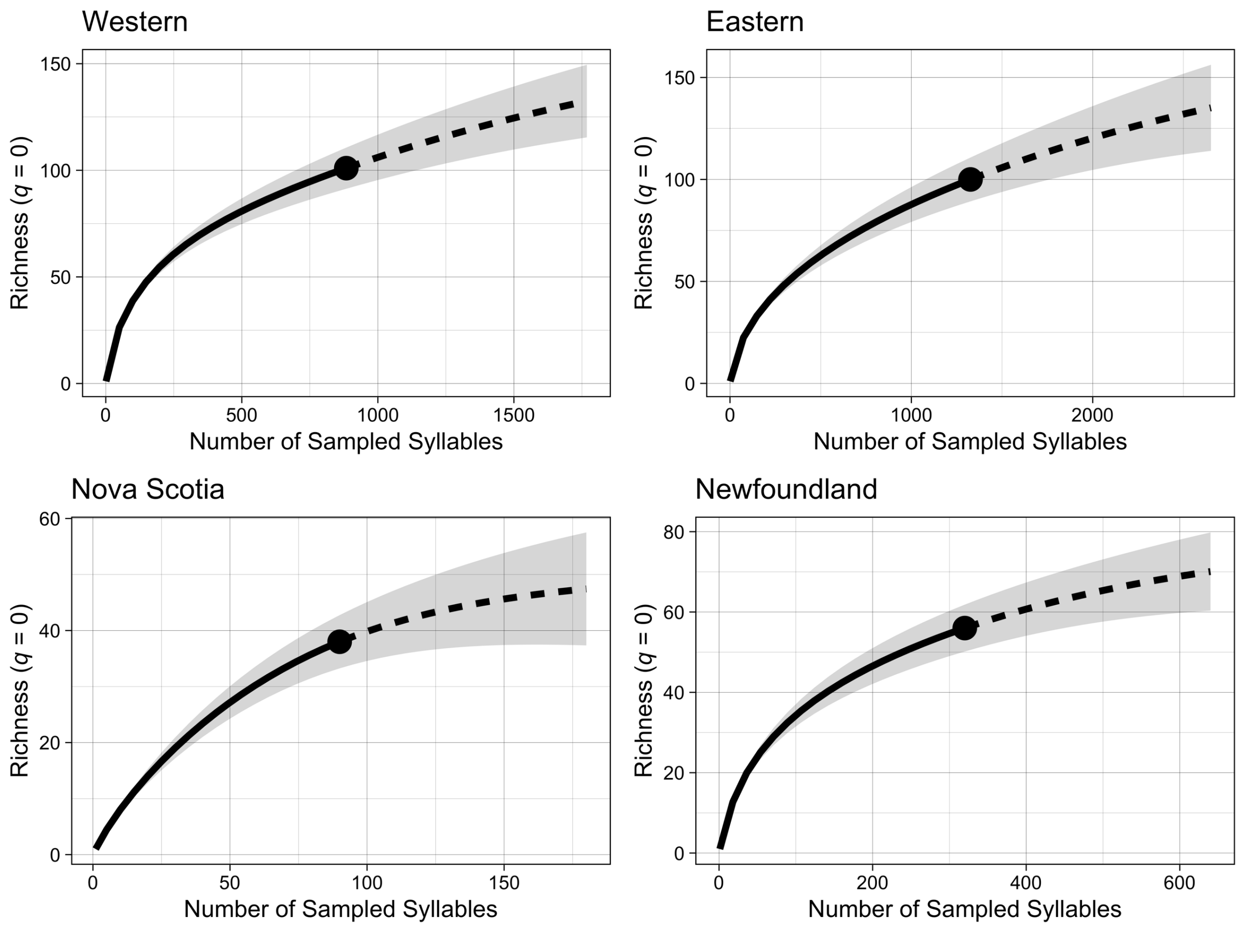

Figure S2. The results of unseen species models applied to the full frequency distribution of syllables in each regiolect across all yrs, run using the *iNEXT* package in R (Hsieh et al. 2016). The point in each plot marks the observed richness (*y*-axis) for the total number syllables sampled (*x*-axis). The solid lines mark interpolated richness, the dashed lines mark extrapolated richness, and the shaded areas mark 95% confidence intervals. The interpolated richness values are the number of types you observe if you subsample that number of tokens from the real data. The extrapolated values are the projected number of types you would observe if you were able to continue sampling tokens from the same population. More details can be found in Hsieh et al. (2016: https://doi.org/10.1111/2041-210X.12613).

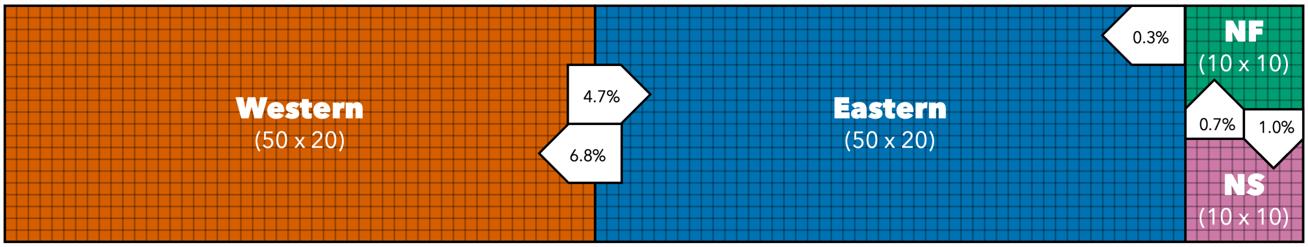

Figure S3. A diagram of the four grids used to simulate geography in the agent-based model, labelled by regiolect. The white arrows denote probability of cultural transmission between neighboring regiolects, based on the proportion of males in each regiolect who sing alien songs averaged across the three time periods (see Supplementary Material Table S3 above for more details).

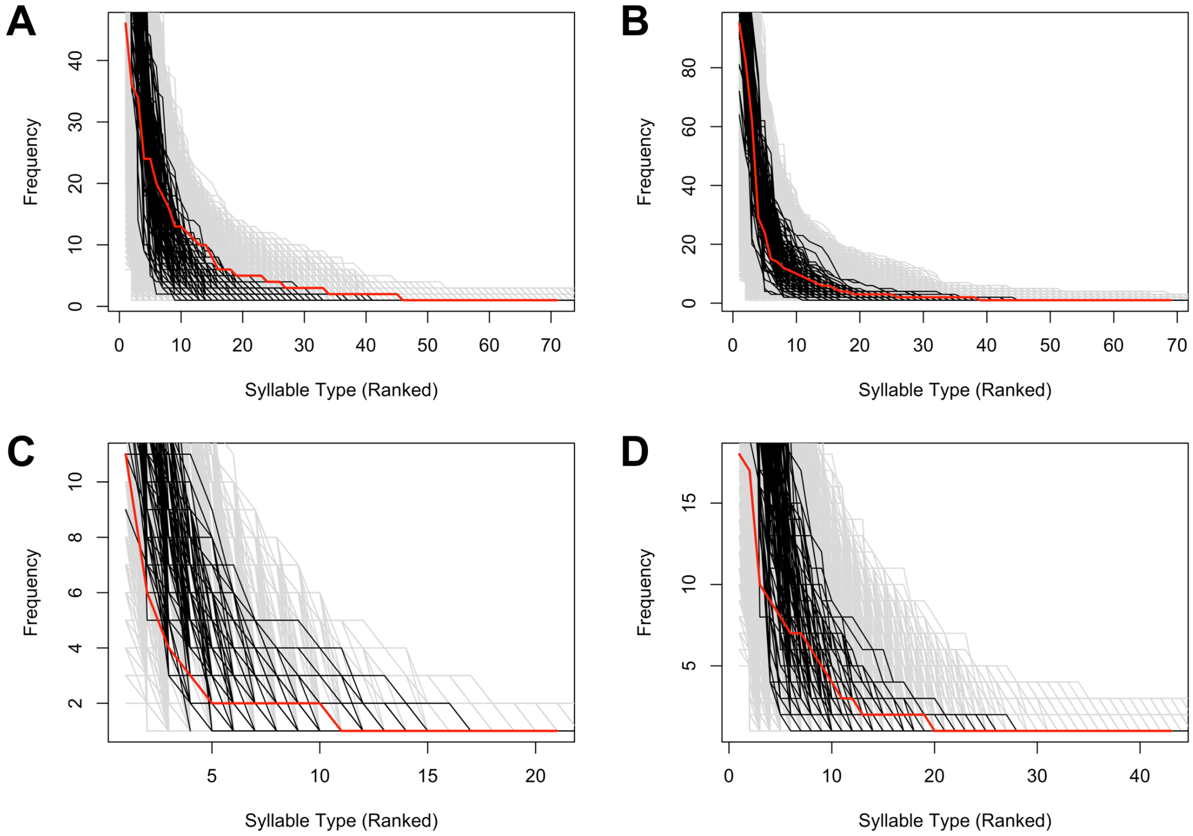

Figure S4. The observed (red) and simulated (grey and black) frequency distributions from 2019 for each of the four regiolects (A: Western; B: Eastern; C: Nova Scotia; D: Newfoundland). The black lines denote the 100 simulated distributions with parameter values closest to the estimated parameter values in terms of Euclidean distance. In all cases, the posterior simulations are reasonably close to the observed distribution.

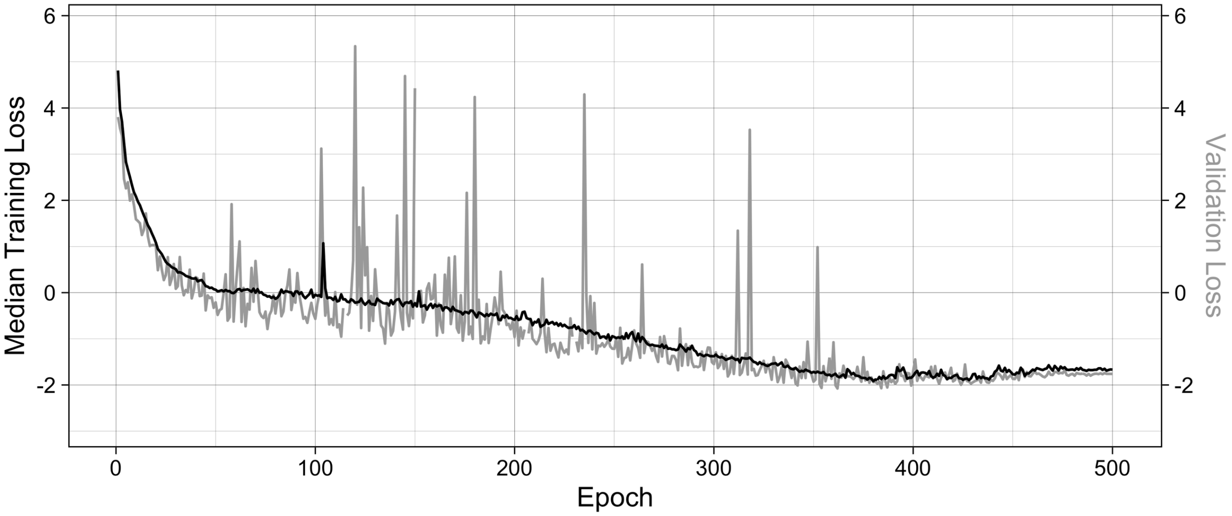

Figure S5. The median training loss of the BayesFlow model (left y-axis, in black) and the validation loss on the testing set (right y-axis, in grey) over the 500 epochs of training. The curves have roughly the same trajectory and plateau towards the end of training, showing that the model has converged and has good generalization to the testing set. Note that the validation loss curve is noisier and lower (on average) than the median training loss curve because it shows only the final value in each epoch, rather than the median value across the entire epoch.

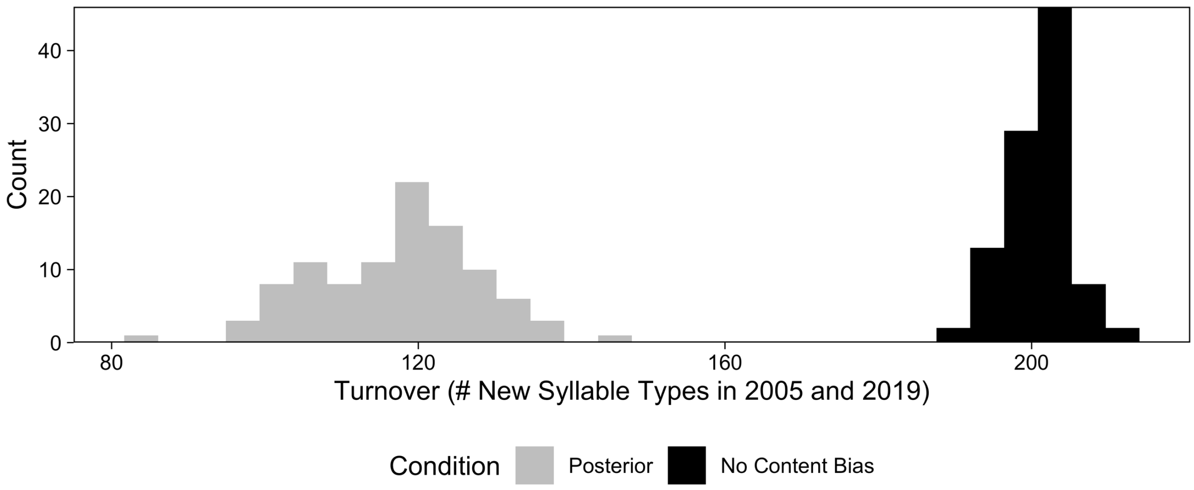

Figure S6. The observed level of turnover from the fitted agent-based model. Turnover here is simply the number of new syllable types that appeared in the population in 2005 and 2019, according to Acerbi and Bentley (2014: https://doi.org/10.1016/j.evolhumbehav.2014.02.003). The data in grey come from 100 simulations from the agent-based model, using the median point estimates for each parameter of the fitted model. The data in black come from 100 simulations from the agent-based model, using the median point estimates for each parameter of the fitted model except for content bias (β), which was set to 0.