Available online at www.sciencedirect.com

SCIENCE DIRECT®

Perception of geographical variation in song by male Puget Sound white-crowned sparrows, *Zonotrichia leucophrys pugetensis*

DOUGLAS A. NELSON & JILL A. SOHA

Borror Laboratory of Bioacoustics, Department of Evolution, Ecology and Organismal Biology, The Ohio State University

The songs of male Puget Sound white-crowned sparrows form about 13 dialects along the Pacific Northwest coast of North America. In his original description of dialects in this subspecies, Baptista (1977, *Condor*, **79**, 356–370) defined dialects by the terminal trill portion of the song because most males at a given location sing the same trill. Complex syllables in the introductory part of the song vary between males at a sample location and occur over a wider geographical range than trills. Baptista also recognized two superdialects or 'themes' based on differences in the phrase order within songs. We performed three playback experiments to territorial male white-crowned sparrows using both natural and computer-modified stimuli to test which acoustic features of song these birds attend to. Two of the experiments were conducted in nine dialect areas spanning nearly the entire geographical range of the subspecies. The results of the first such experiment indicate that across dialect areas, males give stronger responses (increased song rate and a shorter approach distance) to natural songs from the local dialect than they do to songs from a neighbouring dialect, a different theme, or another subspecies of white-crowned sparrow. Two experiments using modified stimuli conducted either across nine dialects or in one dialect led to the same conclusion: substitution of the trill from a neighbouring dialect into a song elicited weaker responses relative to the local control than did substitution of a neighbouring dialect's introduction. We conclude that the response of male Puget Sound white-crowned sparrows to playback is influenced most by variation in the song's trill. We found no evidence that the birds recognize 'superdialects' as do humans.

© 2004 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

The acoustic structure of bird songs frequently varies geographically (Krebs & Kroodsma 1980; Mundinger 1982). Because male songbirds learn their songs, imitation 'errors' or improvisational inventions during vocal development may introduce novel song variation into geographically isolated populations (Lemon 1975; Mundinger 1982). If males learn their songs where they breed, geographical song variants will tend to persist (Krebs & Kroodsma 1980).

Human observers generally have little difficulty in identifying geographical variation in song, either by sorting sonagrams 'by eye' into categories (perhaps still

the most sensitive technique) or by using quantitative acoustic and statistical methods (Bradbury et al. 2001; Warren 2002). Evaluating the birds' perception of these hypothetical categories is a crucial first step for testing hypotheses about the functional significance of song variation. Many playback studies have demonstrated that male and female songbirds are attentive to geographical variation in song. In most cases, male and female birds give stronger responses to songs with which they are familiar (local songs) than they do to songs from a distant location (Baker 2001). Less attention has been paid to determining the acoustic cues that birds use to distinguish songs from different areas (Pellerin 1982; Baker et al. 1987; Balaban 1988; Thompson & Baker 1993).

The set of experiments we report here examined the perception of geographical variation in song by male Puget Sound white-crowned sparrows. Baptista (1977)

Correspondence: D. A. Nelson, Borror Laboratory of Bioacoustics, Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 318 W. 12th Avenue, Columbus, OH 43210, U.S.A. (email: nelson.228@osu.edu).

described a series of six song dialects in this sparrow along the Pacific Northwest coast of North America (Fig. 1). Baptista described two main geographical patterns in song from his survey, both of which have been corroborated, at least in part, by our recent work. First, the terminal trill of the song varies little among males at a given sampling location, whereas the complex syllables in the song's introduction at the same location are usually more diverse

(Fig. 2). Complex syllables also tend to be more widely distributed geographically than are terminal trills (Baptista 1977). Baptista defined the six dialects based on differences in the terminal trill, and this number has been increased to about 13 dialects based on work in areas bypassed in Baptista's survey (Heinemann 1981; Baker 1987; DeWolfe & Baptista 1995; Chilton & Lein 1996). The second main pattern in song variation he discerned
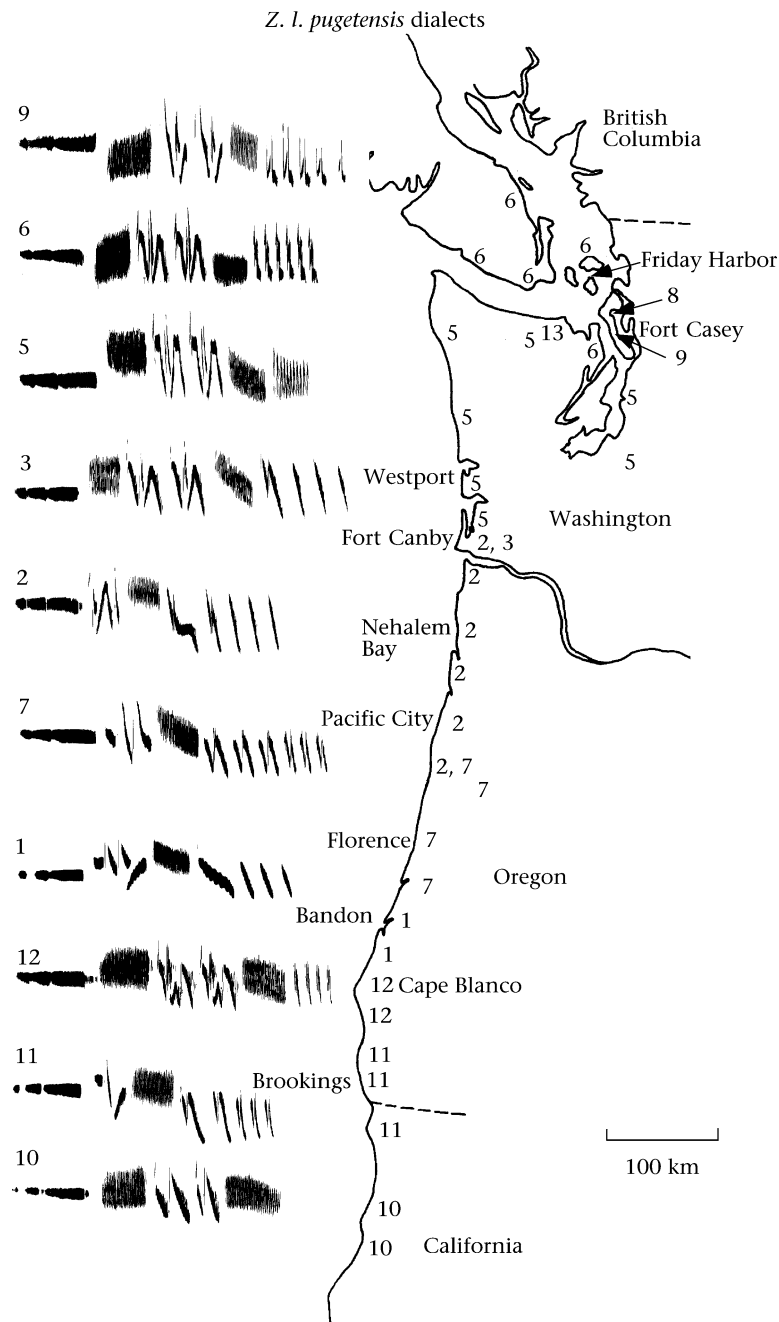

Figure 1. Examples of 10 dialects and their locations along the west coast of North America based on our recordings. Dialect 6 also occurs on mainland British Columbia (Baptista 1977). Dialects 8 (Whidbey and Camano Islands, Washington) and 13 are not illustrated. Names of the 10 test locations used in experiments 1 and 3 are shown. Northern themes occur north of the Columbia River and consist of five phrases in the following order: whistle, buzz, pair of complex syllables, buzz, and terminal trill. Dialects 10 and 12 are disjunct northern themes near the southern end of the subspecies' distribution. Southern themes occur south of the Columbia River and begin with a whistle, followed by a single complex syllable, a buzz, and the terminal trill.
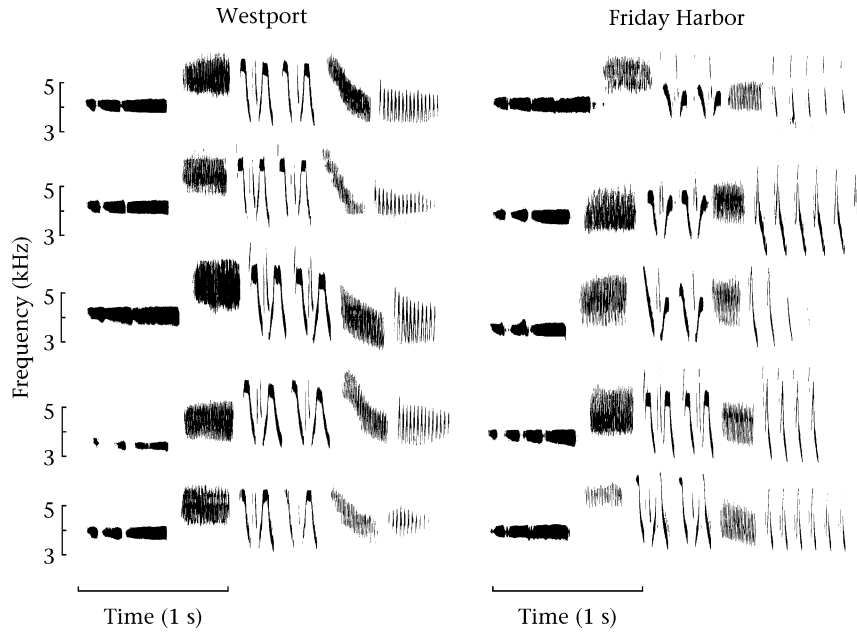
Figure 2. Examples of songs of five males at Westport (dialect 5) and Friday Harbor (dialect 6). The songs from Westport were very uniform while those from Friday Harbor displayed three different types of complex syllables in the introduction. The bottom two Friday Harbor songs contained the same complex syllable type as occurred at Westport. The top two songs at Friday Harbor were representative of 75% of males ($N = 31$) in our sample of recordings. All sound spectrograms in this paper were prepared in Signal using a 128 point FFT (time resolution = 5.1 ms, frequency = 195 Hz).

was the grouping of dialects into superdialects, or what he termed northern and southern 'themes', based on differences in the sequence of phrases (buzzes and complex syllables or note complexes) within the song (Fig. 1).

The white-crowned sparrow has been the focus of attention directed to the hypothesis of whether learned vocal dialects contribute to the genetic subdivision of populations (reviewed in: Kroodsma et al. 1984; Baker & Cunningham 1985). Central to testing this hypothesis is establishing whether the birds perceive vocal variation in a manner that is concordant with the classification used by researchers. In this study, we report the results of field playback experiments designed to test whether male Puget Sound white-crowned sparrows are attentive to variation in the overall structure (theme differences) of their songs, and whether they are particularly attentive to variation in the trill (dialect differences). We tested these hypotheses using experimental designs that incorporated two levels of replication.

Recent years have seen a debate concerning experimental design in playback studies and the level of inference supported by a given experiment (Kroodsma 1989, 1990; McGregor 1992a). One concern is that most, though not all (see Searcy et al. 2002), existing studies of the perception of geographical variation in song have used one or two 'typical' stimulus exemplars to represent each treatment group (e.g. local versus foreign dialect). Inference to the population of songs within a dialect requires the assumption that these few stimuli adequately represent acoustic variation within the dialect. Secondly, most, though not all studies, (Baker et al. 1984; Leader et al. 2002) have tested subjects at a single location. Given that song structure varies among locations in a 'dialect

species', the acoustic cues attended to by the birds at one location may differ from those used at other sites. In all of our experiments we replicated stimuli, and in two experiments we replicated test locations. Where possible, we compare results when locations were or were not replicated.

GENERAL METHODS

Stimulus Preparation

All experiments were between-subjects designs, where each subject was presented with one stimulus treatment. We used a different stimulus for each subject. The details of stimulus preparation specific to each experiment will be described below. One experiment used natural recordings and two experiments used stimuli created by splicing together phrases from different recordings. In all cases, we began with natural songs recorded using a Sony TCD-D10 Pro II digital tape recorder and Sennheiser MKH70 'shotgun' microphone. Most songs used as stimuli had been recorded in previous years, thereby minimizing the probability of subjects being familiar with particular songs. Songs were digitized at 25 kHz with 16-bit amplitude resolution using the Signal software package (Engineering Design 1999). We selected clean recordings, and all songs used in the same experiment were first normalized to the same root-mean-square (average) amplitude level before any synthesis occurred and playback tapes were made. If one song in a set needed high-pass filtering (>1500 Hz) to reduce noise, all were filtered prior to amplitude normalization. All manipulations of stimuli were carried out in Signal using basic commands.

Playback Protocol

Subjects were territorial male Puget Sound white-crowned sparrows tested on their breeding territories from late April until the first week of June with the exception of seven males tested in the last week of June. Most of the birds were paired during this time of year, with mates either building nests or incubating. With the exception of two of the seven males tested after the first week of June, none carried food for offspring during trials. Territories were mapped by observing song posts on days prior to testing. We placed a loudspeaker (Audix PH-3) atop a 1.6-m-high tripod near the centre of the subject's territory. Care was taken to choose a location with trees or shrubs distributed evenly near the speaker for the male to perch in. Songs were broadcast at a constant level (approximately 76 dB sound pressure level, SPL, at 8 m). At least 48 h elapsed between tests of males on neighbouring territories. Presentation order was counterbalanced with stimulus identity in a Latin Square design in each experiment.

Stimuli were recorded on a Sony TC-D5ProII analogue cassette recorder at six songs per minute using a Signal program. Stimulus tapes were coded so the observer did not know the identity of the stimulus, although they could hear the stimulus during playback. One observer sat 10 m from the speaker and observed the male's behaviour during three periods: a 2-min preplayback period, a 2-min playback period, and a 10-min postplayback period. Observations were dictated into a second tape recorder, and notes were transcribed later the same day. Based on previous experience, we knew that most males make a silent, close approach during playback of conspecific song from the centre of their territory, and that responses vary most between stimuli during the postplayback period. We therefore present results for the postplayback period only.

Data Analysis

We measured three response variables during the postplayback period. The song rate (songs/min) was calculated by subtracting the song rate in the pretrial period from that of the postplayback period. Mean distance from the speaker during the postplayback period was calculated by noting the male's position continuously relative to five categories: 0–2 m (median: 1 m), 2–4 m (3 m), 4–8 m (6 m), 8–16 m (12 m) and 16+ m (24 m). Distances were measured with a marked rope. The number of seconds spent in each category was multiplied by the respective median distance, the products summed, and the sum divided by 600 s to yield the mean distance. Distance of closest approach was also recorded as one of the above five categories. We combined the three behavioural response measures into one or two orthogonal principal components (PC), and used the PC scores as the random variables in statistical testing. Separate tests on the original three variables would not be statistically independent (Rice 1989) and would not reflect the multivariate nature of the response (McGregor 1992b). We made a priori pairwise comparisons between treatments using Mann–Whitney U tests after verifying that treatment group variances were equal using Levene's test. Tabled values of rank sums, T, were consulted to obtain exact probabilities and the sequential Bonferroni procedure was used to control the experiment-wise error rate at $P < 0.05$ (Hochberg 1988; Mundry & Fischer 1998). N in all experiments equals the number of subjects, which is equivalent to the number of stimuli. To estimate the power associated with our tests, we followed Cohen (1988) and collected estimates of the effect size, d, from previous studies that made comparisons similar to ours and presented means and standard deviations. We then calculated power for a given N and α level of 0.05 using a power calculator for parametric tests (http://calculators.stat.ucla.edu/powercalc/). We multiplied the calculated value by 0.95 to yield an adjusted power estimate because the Mann–Whitney U test we used has approximately 95% of the power of a parametric test (Dixon & Massey 1969).

EXPERIMENT 1: RECOGNITION OF DIALECTS AND THEMES, WITH DIALECTS REPLICATED

Methods

To establish the responsiveness of male Puget Sound white-crowned sparrows, in general, to geographical variation in song, we tested males in nine dialect areas (as defined by the terminal trill). The hypotheses of interest were: do males throughout the geographical range of this subspecies distinguish between the local dialect and a neighbouring dialect, between the local dialect and a different theme, and between songs of their own and a different subspecies?

Subjects

Trials were conducted between late April and early June during 1999–2001 with the exception of those conducted at Brookings on 29 June 2000. Four males at each location were tested within a few days of one another. D.A.N. performed all the trials except for the four birds tested by J.A.S. at Fort Canby. Tests were run in Oregon at Bandon (Bullard's Beach State Park: dialect 1), Brookings (dialect 11), Cape Blanco State Park (dialect 12), Florence (dialect 7), and Pacific City (Robert Straub State Park: dialect 2); and in Washington at Fort Canby State Park (dialect 3), Westport (Westhaven State Park: dialect 5), Friday Harbor (dialect 6), and Fort Casey State Park (dialect 9).

Stimuli

We used nine sets of stimuli, each containing four natural songs: a song from the test location (local dialect), a song of the same theme from a neighbouring dialect, a song from a dialect of the other theme, and a song of another subspecies of white-crowned sparrow (Fig. 3). No dialect was used more than once as a neighbouring dialect and none was used more than once to represent the other theme. For the other subspecies, we used songs of four Nuttalls' white-crowned sparrow (*Z. l. nuttalli*) males and five Gambel's white-crowned sparrow (*Z. l. gambelii*) males. *Zonotrichia l. pugetensis* winters with *Z. l. nuttalli* and
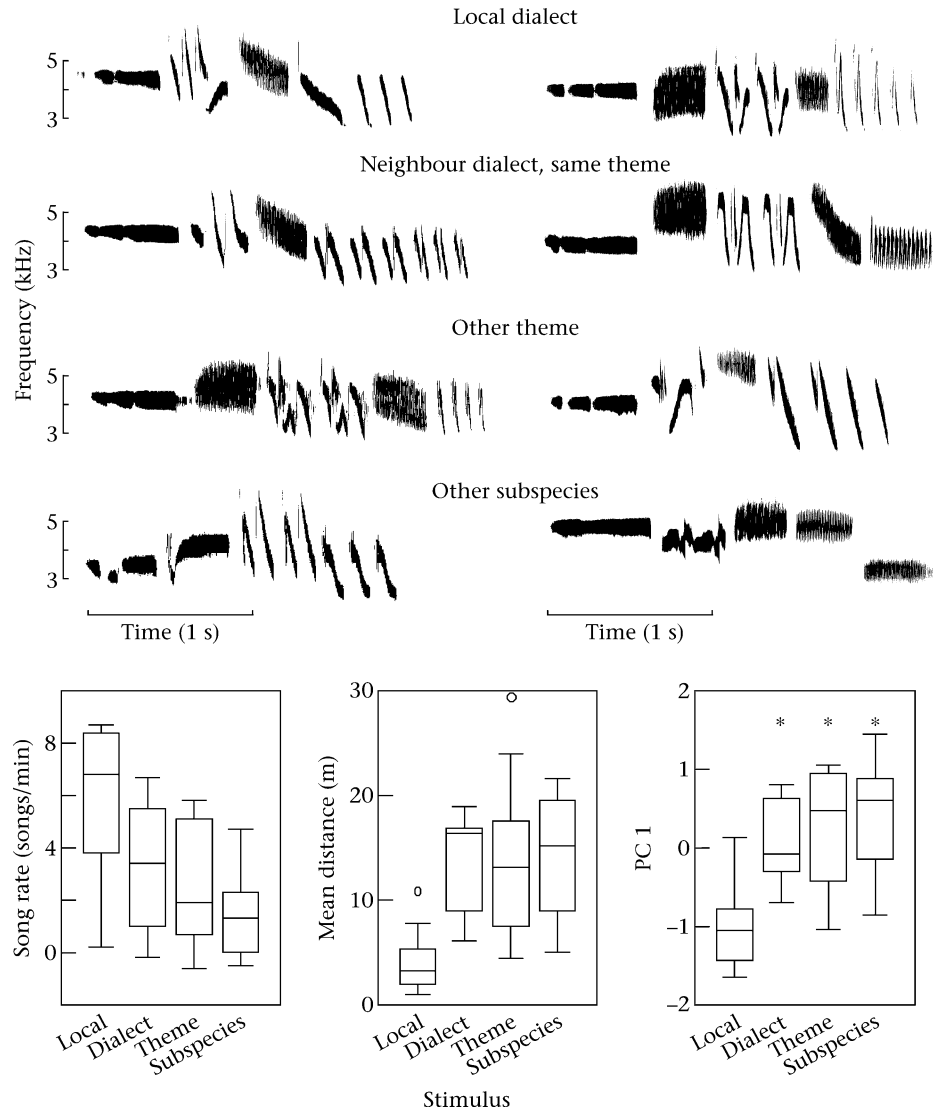
Figure 3. Experiment 1 stimuli and results. The four stimuli used at Bandon (local is dialect 1, a southern theme) are shown in the left column, and the four stimuli used at Friday Harbor (local is dialect 6, a northern theme) are shown on the right. Box plots showing two of three original response variables (change in song rate and mean approach distance) along with principal component 1 are shown below. Results for close approach distance are not shown because this variable was highly correlated with mean approach distance (see text). The lower and upper edges of the boxes represent the first and third quartiles; the median divides each box. The vertical lines ('whiskers') include the range of values within 1.5 times the interquartile range. Outliers are individual circles. Asterisks in the PC1 graph indicate significant differences relative to the local song.

Z. l. gambelli, males of all three subspecies sing together in late winter, and *Z. l. gambelli* males sing on their northward migration through *Z. l. pugetensis* habitat (Chilton & Lein 1996; personal observation).

Results

Mean approach distance (Pearson's correlation: $r_7 = 0.87$), closest approach distance ($r_7 = 0.70$) and song rate ($r_7 = -0.74$) had strong correlations with PC1. Males gave significantly stronger responses, as shown by smaller PC scores (closer approach and an increased song rate), to the local dialect than they did to either a neighbouring dialect of the same theme (Mann–Whitney U test:

$T = 52$, $N_1 = N_2 = 9$), the other theme ($T = 54$), or another subspecies ($T = 50$, Bonferroni $P < 0.05$; Fig. 3). Responses did not differ between the neighbouring dialect of the same theme and a different theme ($T = 82$, $N_1 = N_2 = 9$, Bonferroni NS). We performed a power analysis to evaluate this negative result.

We found three playback studies in the literature that allowed us to calculate the effect size, d, for the comparison between responses to two different dialects. In the study most similar to ours methodologically, Thompson & Baker (1993) converted three response variables of *Z. l. nuttalli* males into a single discriminant function variable for use in significance testing. For the comparison between a local and neighbouring dialect, d equalled 1.46, and yielded an adjusted power of 0.79.

Two other studies that measured single response variables for the same comparison (number of songs or approach latency), produced estimates of d and adjusted power, respectively, of 1.24 and 0.66 (northern cardinal, *Cardinalis cardinalis*, Lemon 1967), and 1.07 and 0.56 (corn bunting, *Miliaria calandra*, McGregor 1983). We assume that our technique of measuring three response variables would also produce an effect size of 1.46, yielding a moderate degree of power to detect response differences between two dialects.

EXPERIMENT 2: RELATIVE IMPORTANCE OF INTRODUCTION AND TRILL IN ONE DIALECT

Methods

Experiment 1 established that males give weaker responses to foreign songs, but did not identify which feature(s) of a song make it foreign or familiar. In experiment 2 we used computer-modified songs composed of the introductions and trills from the local dialect at one site and from one neighbouring dialect to test the relative importance of variation in the introduction and trill in dialect 'recognition' by males. Based on Baptista's hypothesis that the trill defines local dialects in this subspecies, we expected that substitution of a foreign trill into the song would have a greater effect on male responsiveness than would substitution of a foreign introduction.

Subjects

D.A.N. tested 24 males in Fort Stevens State Park, Oregon at the mouth of the Columbia River (dialect 2) in April and May of 1997 and 1998.

Stimuli

We created eight sets of stimuli, each containing three computer-modified 'hybrid' songs, by splicing together the introduction or trill from the songs of different local males, and the introduction or trill from the songs of different males in dialect 7, recorded at Newport or Mary's Peak, Oregon. A control stimulus was created by splicing the introduction from one local male's song to the trill from another local male's song. A stimulus with a foreign introduction and local trill was created by splicing the local trill used in the control after the introduction of the neighbouring dialect. A stimulus with a foreign trill and local introduction contained the same local introduction as in the control, followed by the trill from the same neighbouring dialect song used in the foreign introduction/local trill stimulus (Fig. 4).

Results

Mean approach distance (Pearson's correlation: $r_6 = 0.88$), closest approach distance ($r_6 = 0.84$) and song rate ($r_6 = -0.65$) had strong correlations with PC1. Males gave significantly stronger responses, as measured by smaller PC1 scores (close approach and high song rate) to playback of the local stimulus than they did to the local

introduction/foreign trill stimulus (Mann–Whitney U test: $T = 43$, $N_1 = N_2 = 8$, Bonferroni $P < 0.05$). Responses to the foreign introduction/local trill song did not differ from the control ($T = 64$, $N_1 = N_2 = 8$, NS). PC1 scores and two of the original response variables are shown in Fig. 4. On the assumption that substituting a phrase from a foreign dialect produced the same effect size as an entire foreign dialect ($d = 1.46$), the adjusted power estimate was 0.74. This seemed a reasonable assumption given that several of the natural dialects differed primarily in the trill.

EXPERIMENT 3: RELATIVE IMPORTANCE OF THE INTRODUCTION AND TRILL, DIALECTS REPLICATED

Methods

Experiment 2 used computer-modified 'hybrid' songs to examine the relative importance of the introduction and the trill in influencing male responses. The results suggest that the responses of males in the dialect 2 area were influenced more by variation in the song's trill than by variation in the introduction, when variation was introduced by phrases from the neighbouring dialect 7. To test whether this effect was general across dialects in this subspecies, we replicated test locations, as in experiment 1, and created stimuli using phrases from the local dialect and a nearby dialect for each test location.

Subjects

Three subjects were tested at each of the nine locations used in experiment 1, except that we substituted Nehalem Bay for Pacific City in dialect 2 ($N = 9$ dialects). Trials were run between late April and early June of 1999–2001, with all three males tested within a few days of one another at each location. The males at Brookings were tested on 28 June 2000. D.A.N. performed all trials except for the three males at Fort Canby and two of three at Bullard's Beach.

Stimuli

Nine sets of three stimuli each were assembled as in experiment 2 by splicing the introduction from one song to the trill from another. In assembling stimuli for experiment 3, we used two songs chosen from each of the nine test locations as the local songs and one song from a neighbouring dialect (at seven locations) or from two dialects away (two locations) relative to each location (Fig. 5). We used non-neighbouring dialects at two locations to ensure that no dialect was used twice as the neighbouring dialect in the experiment. Thus, no dialect was used more than once as either the local dialect or neighbouring dialect in stimulus synthesis. In every case, the song chosen from the neighbouring dialect (hereafter used to include the two non-neighbouring dialects also) had a different complex syllable type in the introduction
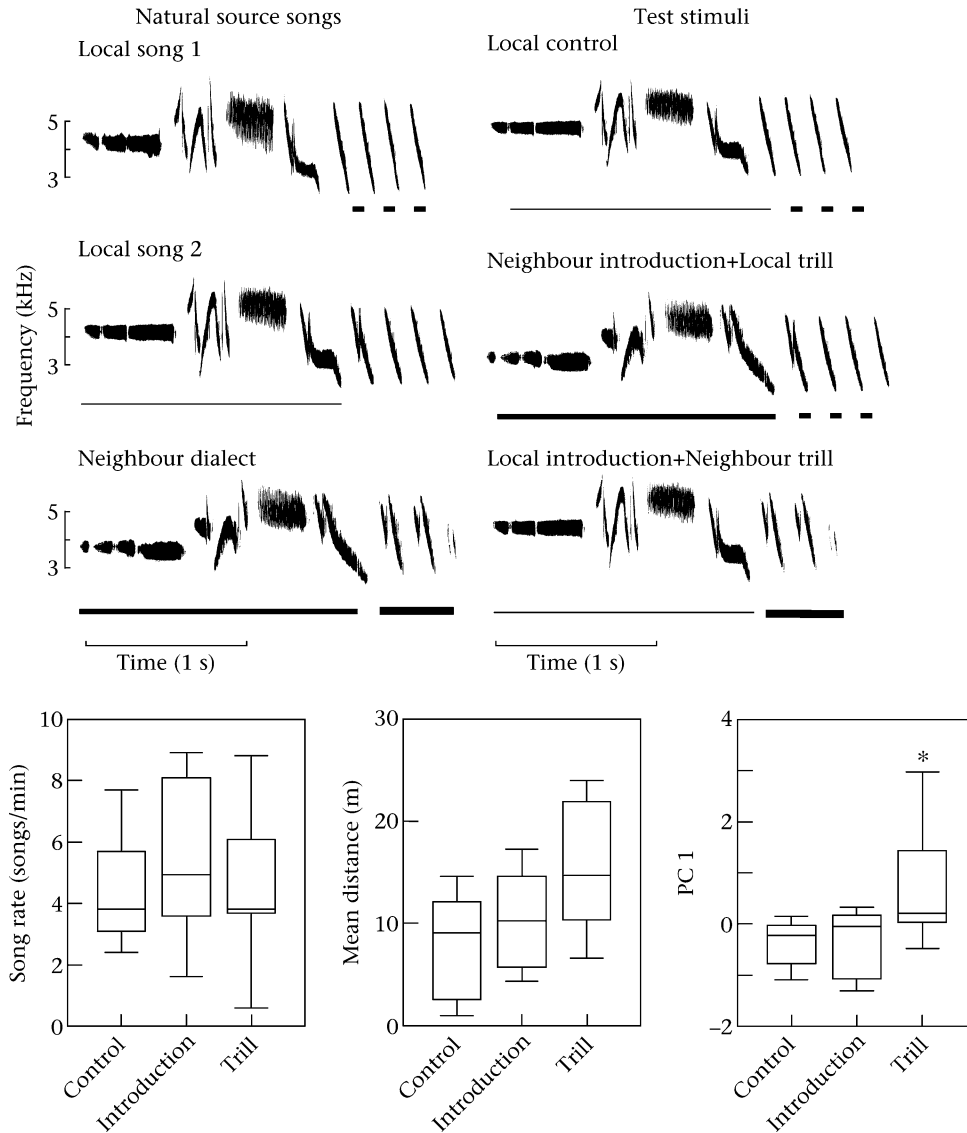
Figure 4. Experiment 2 stimuli and results. The three stimuli are shown in the right column, and the three natural songs that served as the sources for the phrases used in stimulus synthesis are shown on the left. Different styles of underlining identify phrases spliced from the source songs to create the synthetic songs. Results below are presented as in Fig. 3. 'Introduction' and 'trill' refer to the stimuli with those phrases, respectively, taken from a neighbouring dialect. An asterisk indicates a significant difference relative to the control.

and a different trill type than the local songs used for stimulus synthesis (see Baptista 1977 for a catalogue of syllable types). As a consequence, the complex syllable and the trill from a neighbouring dialect differed from their counterparts in the control song (Fig. 5). Had we chosen the song from the neighbouring dialect at random, the common sharing of complex syllables among dialects would have likely led to the production of 'neighbour introduction' stimuli that were nearly identical to local songs. The complex syllable type chosen from the neighbouring dialect did not occur at the test location in six cases ($N = 15-25$ males recorded at each location), was sung by fewer than 10% of males in two cases, and was common (46%) in one case, at Fort Casey. Our procedure of choosing an acoustically distinctive complex syllable from the neighbouring dialect should have increased the

likelihood that we would detect a different response to this stimulus if such a response occurs.

Results

The three behavioural response measures reduced to two PCs. Mean approach distance (Pearson's correlation: $r_7 = 0.80$) and close approach distance ($r_7 = 0.73$) correlated with PC1, whereas song rate had a strong correlation with PC2 ($r_7 = 0.89$) and a weak correlation with PC1 ($r_7 = -0.10$). Males approached significantly closer, as measured by PC1, to playback of the local control song than they did to the song with the local introduction/ neighbouring trill (Mann–Whitney U test: $T = 58$, $N_1 = N_2 = 9$, $P < 0.05$). The stimulus with the neighbouring introduction/local trill also elicited a weaker response
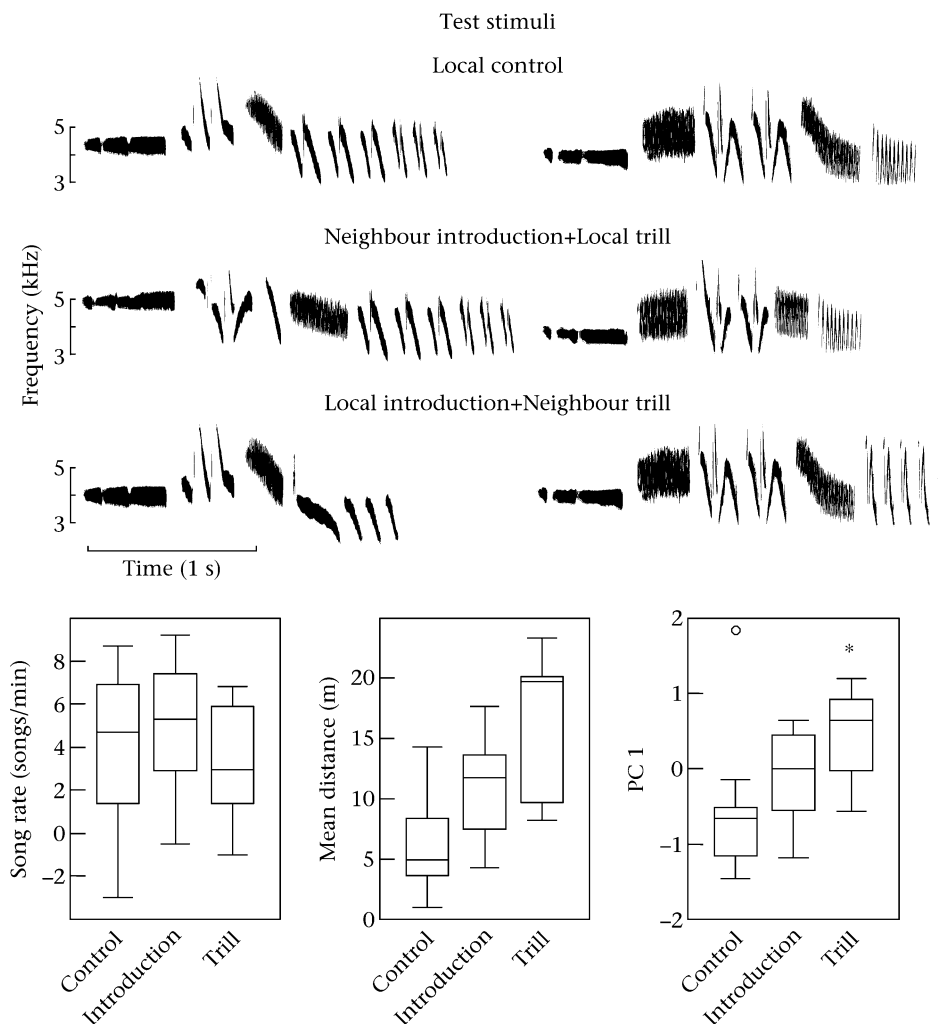
Figure 5. Experiment 3 stimuli and results. Above are shown trios of stimuli, created as described in Fig. 4, for two locations. In the left column are stimuli used at Florence (dialect 7) with dialect 1 as the foreign dialect. The right column shows the stimuli used at Westport (dialect 5) with dialect 6 as the foreign dialect. Results below are presented as in Fig. 3. 'Introduction' and 'trill' refer to the stimuli with those phrases, respectively, taken from a neighbouring dialect. An asterisk indicates a significant difference relative to the control.

than the control, but not significantly so ($T = 66$, $N_1 = N_2 = 9$, NS). Making the same assumption as in experiment 2, and given an effect size of 1.46, the adjusted power estimate was 0.79. PC2 did not differ between stimuli. We conclude that across dialects, substitution of the trill from a neighbouring dialect elicits weaker responses from males than does substitution of a neighbouring dialect's introduction.

DISCUSSION

As has been shown in many other species of birds, and in other subspecies of white-crowned sparrow, Puget Sound white-crowned sparrow males give stronger responses, as measured by approach to the loudspeaker and singing rate, to songs of their breeding location than they do to songs recorded elsewhere (Catchpole & Slater 1995). Although humans can easily discern a hierarchical organization of song dialects nested within two themes,

northern and southern, in the songs of this subspecies (Fig. 1), the results of experiment 1 provide no indication that a foreign theme is perceived differently than is a neighbouring dialect of the same theme. Because we replicated dialect areas of test subjects in experiment 1, we can infer that this conclusion applies to the subspecies as a whole, and not just to birds breeding in one dialect area. White-crowned sparrow males appear to lump all foreign dialects into one 'nonlocal' category, and as far as our methods can detect, do not distinguish between nonlocal songs. Our method had a reasonable degree of power, approximately 0.79, although differences might emerge in a larger sample. In our design, however, we could not increase N much further without pseudoreplicating dialects. The hierarchical organization of song structure that we detect may provide clues to the cultural history of these songs, in that the northern and southern theme dialects probably have different cultural ancestors, but this distinction in song structure does not appear to be important to male birds.

Experiments 2 and 3 used computer-modified stimuli to identify the acoustic features of song that males attend to. These experiments were motivated by Baptista's (1977) hypothesis that the terminal trill is the defining feature of a vocal dialect in this subspecies. In some other subspecies of white-crowned sparrow as well, the terminal part of the song, a trill phrase (*Z. l. nuttalli*: Marler & Tamura 1962) or other phrase (*Z. l. gambelii*: Chilton et al. 2002), appears to vary geographically. In contrast, in *Z. l. nuttalli* on the Point Reyes National Seashore, two to six dialects can be defined within the same region, depending upon what parts of the song are used by humans to sort songs into categories (Baker & Thompson 1985). On Point Reyes, the complex syllables, probably homologous to the complex syllables in *Z. l. pugetensis*, rather than the trill syllables, distinguish the six smallest dialects.

Our experiments tested two hypotheses pertaining to the general hypothesis that the trill is an important cue to males in song recognition. In experiment 2 the local and neighbouring dialects were not replicated while in experiment 3 they were. By replicating test locations and neighbouring dialects, experiment 3 provides a more general test of the hypothesis that the trill is the primary focus of a male's attention. Both experiments 2 and 3 led to the same conclusion: substitution of a trill phrase from the neighbouring dialect produces relatively weaker responses than does substitution of a neighbouring introductory phrase. The convergence on the same conclusion provides some confidence that, in this case at least, results obtained on one population of subjects may generalize to other populations. This conclusion concerning the importance of the trill may not generalize to other situations. In a single pair of neighbouring *Z. l. nuttalli* dialects on Point Reyes, male responses are most influenced by substitution of phrases in the song's introduction, the part of the song that appears to differ most (to us) between the Limantour and Drake dialects (Thompson & Baker 1993).

By comparing responses to a single pair of dialects, experiment 2 resembles most previous studies of the perception of geographical song variation, and we performed it so that we could compare results obtained with differing degrees of replication. One difference between our experiment and most previous tests is that we used different stimulus exemplars for each subject, thereby sampling the 'stimulus space' more thoroughly (Kroodsma 1990). The design used in experiment 2, or preferably, a reciprocal design (Baker et al. 1984; Leader et al. 2002), addresses the narrower, but important hypothesis of whether two specific dialects elicit different responses from the birds. Most of the dialects recognized by researchers seem fairly distinct from one another, including dialects 2 and 7 that were contrasted in experiment 2, but a few are rather similar, perhaps reflecting recent common cultural ancestry (dialect 7 resembles 11, and dialects 6, 8 and 9 form a trio of similar dialects, with the latter two being very small, in Puget Sound). In addition, there is subtle variation between sites within some trill-defined dialects (Chilton & Lein 1996; personal observation). Additional experiments are needed to test whether these subtle forms of song variation are meaningful to the birds.

We do not claim that males are inattentive to the song's introduction, even though statistical power to detect a response difference was good (0.74−0.79). In a larger sample, statistically significant differences in responses to songs with a foreign introduction relative to the control might be obtained. We only claim that the trill is weighted more heavily than is the introduction, just as some acoustic features, such as song frequency, are weighted more heavily than others in the context of species recognition by song (Nelson 1988). A similar result was obtained in another study in which young 'bilingual' males at the boundary between dialects 2 and 3 were presented with playback of either dialect at different times (Nelson 2000). Dialects 2 and 3 represent different themes (Fig. 1), and differ most noticeably to human observers in their overall sequence of buzzes and notes in the introduction, while sharing the trill. Dialect 3 is probably a cultural hybrid between the introduction from dialect 5 to the north and dialect 2's trill to the south. Individual bilingual males selectively increased their frequency of performance of the dialect that was played back to them for up to an hour afterward. This demonstrates that males do attend to information in the song's introduction, at least when these phrases are present in the male's own repertoire.

Our playback results suggest that males pay close attention to the song's trill, the part of the song that provides the most precise 'address' to the song's geographical origin. We cannot conclude that males derive information about location per se from the trill. It may be that a trill, a precisely repeated series of syllables, is particularly well suited to convey information about the singer's motivation, condition, learning ability, or some other measure of competence (Podos 1997; Vallet et al. 1998). In the nightingale, *Luscinia megarhynchos*, trilled and nontrilled syllables appear to be learned by different processes (Hughes et al. 2002), suggesting the possibility that these song structures may have different functions. Evidence that trilled and nontrilled syllables may differ functionally comes from an experiment conducted within the *Z. l. nuttalli* Bodega Bay dialect (Soha & Whaling 2002). When presented with isolated whistle, buzz, or trill phrases from the local dialect, males gave the strongest responses to the trill; the males' responses to the other phrases did not differ from those made to song sparrow, *Melospiza melodia*, song. Males may assess each other during countersinging duels based on their trill performances and preferentially converge on matching trills during the song-learning process (Nelson 2000). This hypothesis, with its assumption that trill sharing facilitates assessment, requires testing.

The complex syllable types in the introduction are distributed more widely in space than are terminal trill types, and may span two or more trill-defined dialects (Baptista 1977; Chilton & Lein 1996; personal observation). A song's introduction may therefore provide less specific information about where the song was learned and may instead encode other types of information. As hypothesized by Marler (1960), different kinds of information (e.g. species and individual identity) may be best encoded in separate parts of a song. The introductory

whistle, which does not display marked geographical variation in this subspecies (D. A. Nelson, unpublished data) or in the mountain white-crowned sparrow, *Z. l. oriantha* (Harbison et al. 1999), may serve as a reliable species-identifying cue in young birds, but does not appear to serve this function in adults (Soha 2000; Soha & Whaling 2002). To our ears, the long duration and stable frequency of the whistle make it ideal for distinguishing pitch differences between the songs of individuals, a function possibly shared with the highly variable complex syllables. Song frequency has been shown to be an important cue to individual recognition in two other sparrow species (Brooks & Falls 1975; Nelson 1989).

In conclusion, our results support part of the general scheme of identifying song dialects in the Puget Sound white-crowned sparrow first hypothesized by Baptista (1977). Males appear to be more attentive to variation in the song's terminal trill than in the introduction. However, we could find no evidence that males attend to higher-order differences in song structure: songs from different themes elicited the same responses as those from nearby dialects within the same theme. Because we tested males in nine dialects throughout most of this subspecies' geographical range, we have some confidence that these conclusions apply generally to this subspecies.

Finally, these results are also relevant to the evolutionary question of whether learned dialects influence the genetic structure of populations. In a survey of genetic variation at four microsatellite loci among the locations studied here along with additional locations, we found equivocal evidence that dialects differ genetically (Soha et al., 2004). To fully address this question, we also need information on how females respond to dialect variation because the sexes often differ in song perception (Ratcliffe & Otter 1996). The results of our playback experiments on males indicate that our classification of dialects based on differences in the terminal trill is not arbitrary but is based on a first approximation of the song features used by the males themselves.

References

Baker, M. C. 1987. Intergradation of song between two subspecies of white-crowned sparrows on the west coast of North America. *Ornis Scandinavica*, **18**, 265–268.

Baker, M. C. 2001. Bird song research: the past 100 years. *Bird Behavior*, **14**, 3–50.

Baker, M. C. & Cunningham, M. A. 1985. The biology of bird-song dialects. *Behavioral and Brain Sciences*, **8**, 85–133.

Baker, M. C. & Thompson, D. B. 1985. Song dialects of white-crowned sparrows: historical processes inferred from patterns of geographic variation. *Condor*, **87**, 127–141.

Baker, M. C., Tomback, D. F., Thompson, D. B., Theimer, T. C. & Bradley, D. C. 1984. Behavioral consequences of song-learning: discrimination of song types by male white-crowned sparrows. *Learning and Motivation*, **15**, 428–440.

Baker, M. C., Spitler-Nabors, K. J., Thompson, A. D., Jr & Cunningham, M. A. 1987. Reproductive behaviour of female white-crowned sparrows: effects of dialects and synthetic hybrid songs. *Animal Behaviour*, **35**, 1766–1774.

Balaban, E. 1988. Cultural and genetic variation in swamp sparrows (*Melospiza georgiana*). II. Behavioral salience of geographic song variants. *Behaviour*, **105**, 292–322.

Baptista, L. F. 1977. Geographic variation in song and dialects of the Puget Sound white-crowned sparrow. *Condor*, **79**, 356–370.

Bradbury, J. W., Cortopassi, K. A. & Clemmons, J. R. 2001. Geographical variation in the contact calls of orange-fronted parakeets. *Auk*, **118**, 958–972.

Brooks, R. J. & Falls, J. B. 1975. Individual recognition by song in white-throated sparrows. III. Song features used in individual recognition. *Canadian Journal of Zoology*, **53**, 1749–1761.

Catchpole, C. K. & Slater, P. J. B. 1995. *Bird Song. Biological Themes and Variations*. Cambridge: Cambridge University Press.

Chilton, G. & Lein, M. R. 1996. Long-term changes in songs and dialect boundaries of Puget Sound white-crowned sparrows. *Condor*, **98**, 567–580.

Chilton, G., Wiebe, M. O. & Handford, P. 2002. Large-scale geographic variation in songs of Gambel's white-crowned sparrows. *Condor*, **104**, 378–386.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, New Jersey: L. Erlbaum.

DeWolfe, B. B. & Baptista, L. F. 1995. Singing behavior, song types on their wintering grounds and the question of leap-frog migration in Puget Sound white-crowned sparrows. *Condor*, **97**, 376–389.

Dixon, W. J. & Massey, F. J., Jr. 1969. *Introduction to Statistical Analysis*. New York: McGraw-Hill.

Engineering Design. 1999. *Signal*. Belmont, Massachusetts: Engineering Design.

Harbison, H., Nelson, D. A. & Hahn, T. P. 1999. Long-term persistence of song dialects in the mountain white-crowned sparrow. *Condor*, **101**, 133–148.

Heinemann, D. 1981. Song dialects, migration, and population structure of Puget Sound white-crowned sparrows. *Auk*, **98**, 512–521.

Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, **75**, 800–802.

Hughes, M., Hultsch, H. & Todt, D. 2002. Imitation and invention in song learning in nightingales (*Luscinia megarhynchos* B., Turdidae). *Ethology*, **108**, 97–113.

Krebs, J. R. & Kroodsma, D. E. 1980. Repertoires and geographical variation in bird song. In: *Advances in the Study of Behavior vol. 11* (Ed. by J. S. Rosenblatt, R. A. Hinde, C. Beer & M. -C. Busnel), pp. 143–177. New York: Academic Press.

Kroodsma, D. E. 1989. Suggested experimental designs for song playbacks. *Animal Behaviour*, **37**, 600–609.

Kroodsma, D. E. 1990. How the mismatch between experimental design and the intended hypothesis limits confidence in knowledge, as illustrated by an example from bird-song

dialects. In: *Interpretation and Explanation in the Study of Animal Behaviour* (Ed. by M. Bekoff & D. Jamieson), pp. 226–245. Boulder, Colorado: Westview Press.

Kroodsma, D. E., Baker, M. C., Baptista, L. F. & Petrinovich, L. 1984. Vocal 'dialects' in Nuttall's white-crowned sparrow. *Current Ornithology*, **2**, 103–133.

Leader, N., Wright, J. & Yom-Tov, Y. 2002. Dialect discrimination by male orange-tufted sunbirds (*Nectarinia osea*): reactions to own vs. neighbor dialects. *Ethology*, **108**, 367–376.

Lemon, R. E. 1967. The response of cardinals to songs of different dialects. *Animal Behaviour*, **15**, 538–545.

Lemon, R. E. 1975. How birds develop song dialects. *Condor*, **77**, 385–406.

McGregor, P. K. 1983. The response of corn buntings to playback of different dialects. *Zeitschrift für Tierpsychologie*, **62**, 256–260.

McGregor, P. K. 1992a. Design of playback experiments: the Thornbirdge Hall NATO ARW consensus. In: *Playback and Studies of Animal Communication: Problems and Prospects* (Ed. by P. K. McGregor), pp. 1–9. New York: Plenum.

McGregor, P. K. 1992b. Quantifying responses to playback: one, many, or composite multivariate measures? In: *Playback and Studies of Animal Communication* (Ed. by P. K. McGregor), pp. 79–96. New York: Plenum.

Marler, P. 1960. Bird songs and mate selection. In: *Animal Sounds and Communication* (Ed. by W. N. Tavolga), pp. 348–367. Washington, D.C.: A.I.B.S. Symposium Proceedings.

Marler, P. & Tamura, M. 1962. Song 'dialects' in three populations of white-crowned sparrows. *Condor*, **64**, 368–377.

Mundinger, P. C. 1982. Microgeographic and macrogeographic variation in acquired vocalizations of birds. In: *Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 147–208. New York: Academic Press.

Mundry, R. & Fischer, J. 1998. Use of statistical programs for nonparametric tests of small samples often leads to incorrect *P* values: examples from *Animal Behaviour*. *Animal Behaviour*, **56**, 256–259.

Nelson, D. A. 1988. Feature weighting in species song recognition by the field sparrow, *Spizella pusilla*. *Behaviour*, **106**, 158–182.

Nelson, D. A. 1989. Song frequency as a cue for recognition of species and individuals in the field sparrow (*Spizella pusilla*). *Journal of Comparative Psychology*, **103**, 171–176.

Nelson, D. A. 2000. Song overproduction, selective attrition, and vocal dialects in the white-crowned sparrow. *Animal Behaviour*, **60**, 887–898.

Pellerin, M. 1982. The role of silences and elements in the recognition of a dialect in the corn bunting. *Behaviour*, **81**, 287–295.

Podos, J. 1997. A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution*, **51**, 537–551.

Ratcliffe, L. & Otter, K. 1996. Sex differences in song recognition. In: *Ecology and Evolution of Acoustic Communication in Birds* (Ed. by D. E. Kroodsma & E. H. Miller), pp. 339–355. Ithaca, New York: Comstock.

Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Searcy, W. A., Nowicki, S., Hughes, M. & Peters, S. 2002. Geographic song discrimination in relation to dispersal distances in song sparrows. *American Naturalist*, **159**, 221–230.

Soha, J. A. & Marler, P. 2000. A species-specific acoustic cue for selective song learning in the white-crowned sparrow. *Animal Behaviour*, **60**, 297–306.

Soha, J. A. & Whaling, C. 2002. Responses of adult white-crowned sparrows to playback of song phrases: implications for the ontogeny of song recognition. *Condor*, **104**, 848–854.

Soha, J. A., Nelson, D. A. & Parker, P. G. 2004. Genetic analysis of song dialect populations in Puget Sound white-crowned sparrows. *Behavioral Ecology*, **15**, 636–646.

Thompson, A. D., Jr & Baker, M. C. 1993. Song dialect recognition by male white-crowned sparrows: effects of manipulated song components. *Condor*, **95**, 414–421.

Vallet, E., Berne, I. & Kreutzer, M. 1998. Two-note syllables in canary songs elicit high levels of sexual display. *Animal Behaviour*, **55**, 291–297.

Warren, P. S. 2002. Geographic variation and dialects in songs of the bronzed cowbird (*Molothrus aeneus*). *Auk*, **119**, 349–361.