

Song-syllable perception in song sparrows (*Melospiza melodia*) and swamp sparrows (*Melospiza georgiana*): An approach from animal psychophysics

KAZUO OKANOYA and ROBERT J. DOOLING
University of Maryland, College Park, Maryland

Perception of song syllables by female song sparrows (*Melospiza melodia*) and swamp sparrows (*Melospiza georgiana*) was examined using operant conditioning techniques. Similarity matrices among song syllables were constructed from response latencies to detect changes in a repeating background of syllables. Multidimensional scaling (MDS) of these similarity matrices described the perceived relations among the syllable types for each bird. Both species distinguished in a similar way between song- and swamp-sparrow syllables extracted from natural songs. However, when selected features of the syllables were modified, song sparrows and swamp sparrows showed different patterns of perceptual confusion. These results illustrate the utility of MDS techniques for investigating the perception of complex vocal signals, as they demonstrate species differences in the perceptual significance of equivalent physical changes in such complex sounds. The findings also support the notion that swamp sparrows are more selective than song sparrows in the perception of species-specific song syllables.

The understanding of song learning in passerine song birds has been tremendously advanced through the comparative studies of two congeneric species of sparrows, the song sparrow (*Melospiza melodia*) and the swamp sparrow (*Melospiza georgiana*) (see, e.g., Marler, 1987; Marler & Peters, 1981). In these two species, innate preferences exist for the learning of species-specific songs (Marler & Peters, 1989).

To date, this phenomenon has been examined almost exclusively by the song-learning experiments of Marler and his colleagues. To summarize these findings, male swamp sparrows learned only swamp-sparrow syllables regardless of the temporal pattern with which syllables were presented (Marler & Peters, 1977). For male song sparrows, both temporal and phonological patterns were important (Peters, Searcy, & Marler, 1980). Both female song sparrows and female swamp sparrows were sensitive to species-specific phonology and temporal patterns (Searcy, Marler, & Peters, 1981). Moreover, male swamp sparrows could learn song-sparrow notes from hybrid syllables in which swamp-sparrow notes were inserted in song-sparrow syllables and presented in swamp-sparrow temporal patterns (Marler, 1987). This perceptual predisposition to learn conspecific songs is most likely innate in swamp sparrows and probably in song sparrows as well (Dooling & Searcy, 1980).

The above results, obtained primarily from field playback studies and laboratory song-learning studies, clearly point to an important role for perceptual processes in the

learning of species-specific songs. In the present study, we took a new approach to the study of perceptual specializations that might underlie song learning in these sparrows. We used operant conditioning techniques to measure how complex, natural vocalizations (i.e., song syllables) are perceived by birds.

Previous work has shown that these two species have very similar basic auditory sensitivity, with swamp sparrows hearing slightly better at higher frequencies than song sparrows. This difference, though small, corresponds to the differences in spectral energy distribution in the songs of swamp and song sparrows, and this could be one factor contributing to selective vocal learning by these birds (Okanoya & Dooling, 1988c). But thresholds measured under conditions of masking by broadband noise were not different for swamp and song sparrows (Okanoya & Dooling, 1988c).

In the present experiments, we used song syllables as stimuli to study directly the perceptual basis of selective song learning in the sparrows. By combining operant conditioning techniques with an analysis of response latencies by multidimensional scaling (MDS) procedures, we were able to compare the ability of swamp sparrows and song sparrows to differentiate among natural conspecific and heterospecific song syllables.

GENERAL METHOD

Subjects

The subjects of these experiments were 2 female song sparrows and 2 female swamp sparrows. The birds were taken from the field as nestlings (under Federal Permit PRT2-540 to P. Marler) and were subsequently maintained in the aviaries of the Rockefeller University Field Research Center, Millbrook, NY. At Rockefeller, the birds were ex-

Correspondence may be addressed to Robert J. Dooling, Department of Psychology, University of Maryland, College Park, MD 20742.

posed as nestlings to various combinations of natural syllables of both species. At the time of testing in the present experiments, the birds were between 1 and 6 years of age.

Apparatus

The apparatus used for training and testing the birds has been described previously (Okanoya & Dooling, 1988b). The birds were tested in a small wire cage (15 × 15 × 15 cm) placed in a sound-attenuation chamber (IAC-1; 56 × 62 × 78 cm). A standard pigeon grain hopper was mounted on one side of the cage. A response panel consisting of two sensitive microswitches with attached light-emitting diodes (LEDs) was mounted just above the hopper opening. The sparrow could trip the microswitch by striking the LED with its beak.

All experimental events, including stimulus presentation, response measurement, and experimental contingency, were controlled by an IBM AT computer. The song syllables were stored on the hard disk of the computer in digital form and output at a sampling rate of 20 kHz through a digital-to-analog converter (Data-Translation DT2801A) and low-pass filtered at 10 kHz for antialiasing. A loudspeaker mounted inside the sound-proof chamber transduced the stimuli.

Sound-field calibration was accomplished by placing the microphone of a General Radio 1982 sound-level meter in the position normally occupied by the bird's head during testing. Ambient noise level in the chamber was below measurement by the sound-level meter for frequencies higher than 250 Hz. Song syllables were presented at a peak sound-pressure level of 72 dBA.

Stimuli

The song syllables used in these experiments were taken from field recordings of full-length songs of both species made by Donald Kroodmsma while he was at Rockefeller University. Songs of swamp sparrows typically consist of a repetition of a single syllable, whereas songs of song sparrows are more complex, consisting of 3–6 different syllables and arranged in a characteristic syntactical pattern (Marler & Peters, 1989). Since swamp-sparrow song consists of simple repetition of a single syllable, we selected syllables from song-sparrow song that occurred in the repetitive sections of the song.

Ten song syllables of roughly equal length (100–160 msec) were selected from both song- and swamp-sparrow songs. One of the syllables from each species was modified. The broad-band, click-like notes, Type VI and Type I, were extracted from the swamp-sparrow syllable H, resulting in a new, modified syllable called H'. The Type VI note was then appended to the song-sparrow syllable C to make a new hybrid syllable called C'. The complete set of natural and hybrid syllables used in these experiments is shown in Figure 1.

Training and Testing Procedure

The training and testing procedures used in these experiments have also been described in detail elsewhere (Okanoya & Dooling, 1988b, 1988c) and thus are only briefly summarized here.

Mildly hungry birds were trained to peck one key (observation key) repeatedly during the repetitive presentation of one syllable (background), and to peck the other key (report key) when the repeating syllable was exchanged into another syllable (target). Syllables were presented at a rate of 5/sec to mimic the syntax of swamp-sparrow songs. A peck on the report key within 3 sec following a change in the repeating syllable was reinforced by a 4-sec access to food. Once the birds learned this response, they were tested on a set of natural syllables consisting of five song-sparrow syllables and five swamp-sparrow syllables (Figure 1). A matrix of stimuli (background × target) was constructed and one row of background was selected randomly from the matrix.

The procedure consisted of two phases: a habituation phase and a testing phase. During the initial habituation phase, the background sound was repeated for 10 sec. If the bird refrained from responding during this 10-sec habituation phase, the LEDs on both the observation and the report keys were illuminated and the testing phase began. If the bird pecked the report key during the habituation phase, a new habituation interval of 10 sec was initiated. This procedure was designed to decrease spontaneous responding.

During the test phase, a peck on the observation key initiated a random timing interval of 1–7 sec. Following this random interval, a peck on the observation key changed the background repeating syllable into

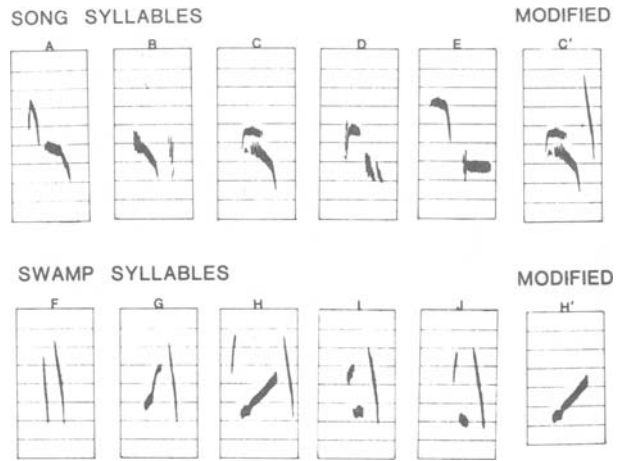


Figure 1. Sonograms of the set of natural and modified song-sparrow and swamp-sparrow syllables used in Experiments 1 and 2. In Experiment 2, two natural syllables, C and H, were replaced with their modified versions, C' and H'.

a target syllable. A response to the report key within 3 sec from the beginning of the repeating target stimuli was reinforced with a 4-sec access to food. About 20% of the trials were sham trials in which the target stimulus was the same as the background stimulus. A response on the report key during a sham trial or during a waiting interval resulted in a 16-sec timeout period during which the lights in the test chamber were extinguished but the repeating sound continued.

The testing phase continued until the background stimulus was paired with every other stimulus in the set three times. Testing proceeded a row at a time (i.e., same background stimulus) until all possible combinations within the row were exhausted. Another row was then randomly selected and a new habituation phase begun. This continued until all rows were tested. This procedure produced a total of three latency matrices for analysis.

Statistical Procedure

A median reaction-time matrix was constructed out of the three latency matrices. The response latency required to detect changes in the repeating background syllable was taken as an index of perceptual similarity between the background and the target syllables (Okanoya & Dooling, 1988b). The upper and lower halves of the matrix were averaged and log-transformed to compensate for positively skewed distribution of the response latencies. The transformed matrices were submitted to an MDS program, SINDSCAL (Shepard, 1980). MDS places the stimuli in multidimensional space so that interstimulus distances correspond to stimulus similarities. The latency data were scaled in two dimensions to permit the simplest interpretation.

EXPERIMENT 1 PERCEPTION OF NATURAL SONG SYLLABLES

In the first experiment, we sought to determine the degree to which swamp sparrows and song sparrows could differentiate among a set of song syllables made up of natural conspecific and heterospecific syllables, and whether they perceived the similarity among the syllables from each species. In other words, are conspecific syllables perceived as more similar to one another than to the syllables of the other species?

Results and Discussion

The two-dimensional spatial representations obtained by SINDSCAL are shown in Figure 2 for swamp sparrows (top

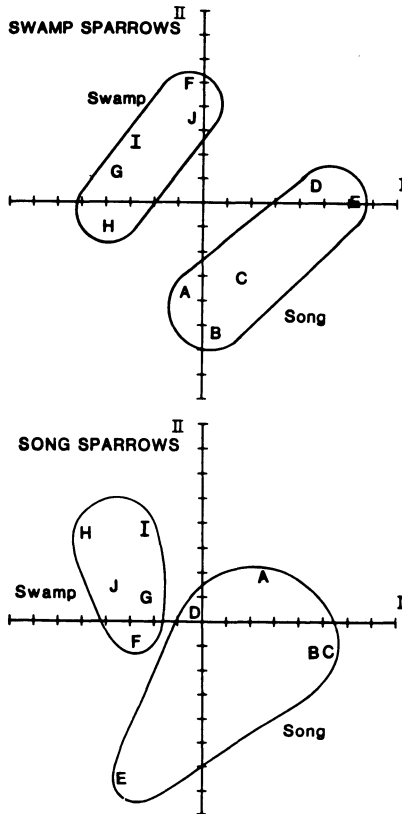


Figure 2. Two-dimensional spatial map of the 10 syllables obtained from a SINDSCAL analysis of the response latencies of the 2 swamp sparrows (top) and the 2 song sparrows (bottom).

panel) and song sparrows (bottom panel). Both species clearly discriminated between swamp- and song-sparrow syllables. The variance accounted for by the two-dimensional solution was 63.3% for the song sparrows, with the first and second dimensions accounting for 35.4% and 27.9%, respectively. For the swamp sparrows, the total variance accounted for was 52.5%, with the first and second dimensions accounting for 30.9% and 21.6%, respectively.

Syllables from each species were clearly separated by the first dimension for both sparrows, with the syllables falling neatly into two clusters. This grouping was also confirmed by a cluster analysis (Aldenderfer & Blashfield, 1984). These results indicate that for both species the most salient perceptual dimension was species identity. Since the same temporal pattern was used for both the swamp-sparrow syllables and the song-sparrow syllables in this experiment, we conclude that syllable type (phonology) alone is sufficient to convey species identity. The subject weights for both birds of each species were similar, indicating that the two individuals were behaving in a similar manner.

**EXPERIMENT 2
PERCEPTION OF MODIFIED SONG SYLLABLES**

The results of the first experiment showed that syllable type was sufficient to convey species identity for both swamp sparrows and song sparrows. Experiment 2 was

designed to further specify the features important for conveying species identity.

The most obvious differences between song- and swamp-sparrow syllables is the presence or absence of the Type VI note, which is found in nearly all swamp-sparrow syllables (Marler & Pickert, 1985). This note appears in the songs of sparrows reared in isolation with intact hearing (Marler & Sherman, 1985) and probably represents one of the critical defining elements of swamp-sparrow songs (Clark, Marler, & Beeman 1987). In this experiment, the importance of the Type VI note in determining species identity was examined using artificially modified syllables.

The same subjects and procedures used in Experiment 1 were used in this experiment. The test stimuli consisted of the two new altered syllables plus the eight remaining, unaltered stimuli from Experiment 1. The purpose of Experiment 2 was to test whether these alterations interfered with the perception of similarity among conspecific syllables.

Results and Discussion

The two-dimensional spatial representation by SINDSCAL based on response latencies to detect modified syllable sets are shown in Figure 3 for swamp sparrows (top panel) and song sparrows (bottom panel). Total variance accounted for was 51.5% for the song sparrows, with the first and second dimensions accounting for 26.3% and 25.2%, respectively. For the swamp sparrows, the total variance accounted

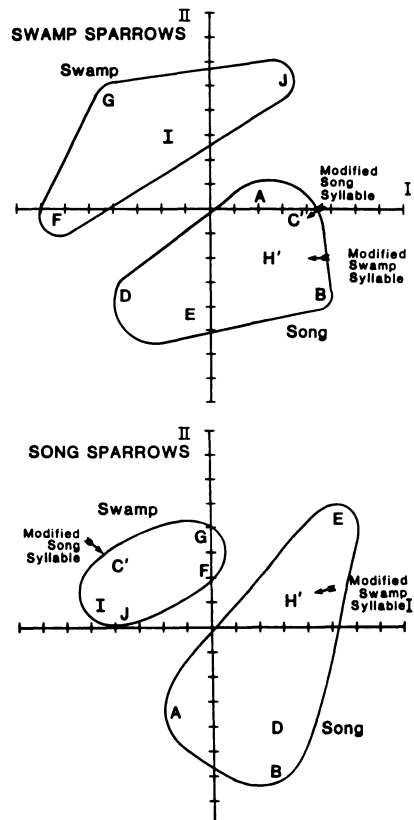


Figure 3. The two-dimensional spatial representation for the eight natural and two modified syllables by SINDSCAL for the 2 swamp sparrows (top) and for the 2 song sparrows (bottom).

for was 41.0%, with the first and second dimensions accounting for 22.2% and 18.9%, respectively.

As in Experiment 1, two clusters were obtained, with one cluster containing song-sparrow syllables and the other cluster containing swamp-sparrow syllables. Relative to these two clusters, however, the spatial location of the modified syllables was different for the two species. For the song sparrows, C' (the modified song-sparrow syllable) was grouped with the swamp-sparrow syllables. H' (the modified swamp-sparrow syllable) was grouped with the song-sparrow syllables.

From these results we conclude that, for female song sparrows, the presence of the Type VI swamp-sparrow note is of primary importance in determining whether a modified syllable is perceived as similar to other swamp-sparrow syllables. For female swamp sparrows, on the other hand, both C' and H' were grouped with the song-sparrow syllables. In other words, swamp sparrows perceived the swamp-sparrow syllable without the Type VI note as more similar to song-sparrow syllables than to swamp-sparrow syllables. Swamp sparrows also perceived the song-sparrow syllable with the Type VI note added as being more similar to song-sparrow syllables than to swamp-sparrow syllables. For swamp sparrows, then, the presence of a Type VI note in a syllable was necessary but not sufficient for the syllable to be perceived as similar to other conspecific syllables. We interpret these results as evidence for greater selectivity in syllable perception for swamp sparrows than for song sparrows.

GENERAL DISCUSSION

Most of what we know about the perceptual basis of selective vocal learning in sparrows comes from the song-learning experiments of Marler and his colleagues (see, e.g., Marler & Peters, 1989). But field playback studies with male sparrows (Peters et al., 1980), copulation/solicitation studies with female sparrows (Searcy et al., 1981), and heart-rate recordings (Dooling & Searcy, 1980) have also been quite useful in refining the understanding of perceptual selectivity for conspecific song in these sparrows.

The techniques used in the present experiments bring yet another tool to the investigation of song perception, but one that has a number of distinct advantages. First, like ethological procedures that typically rely on natural responses, the perceptual categories that emerge from the MDS of response latencies obtained in a detection task are not artificial. That is, because each sound is always paired with every other sound in the set, perceptual groupings are not induced by the training and testing procedures.

Another advantage of the present procedure is that both males and females can be tested with the same procedure. Ethological approaches to measuring song perception must often rely on different behavioral assays for males (e.g., responses to song playback) and females (e.g., copulation/solicitation display). One consequence of using a behavioral response in a natural setting as a perceptual assay is that the results are strongly influenced by motivational factors. The present techniques, by contrast, might be characterized as just the opposite. That is, they reveal how natural stimuli are perceptually organized independent of the natural context or motivational milieu in which the stimuli are normally perceived.

The present experiments illustrate the usefulness of these procedures for describing species differences in the perceptual organization of complex, natural vocal signals. Even in these rather simple experiments, the results are intriguing and unexpected. Equivalent physical changes in these stimuli are not perceived in the same way by the two species. For song sparrows, the presence of a Type VI note is sufficient for modified syllables to be perceived as more similar to swamp-sparrow syllables than to song-sparrow

syllables. Interestingly, preliminary data from canaries tested with these same stimuli show that canaries are similar to song sparrows in their responses.

For swamp sparrows, the story is more complicated and there are at least several possibilities. One is that swamp sparrows rely on the Type I note, which is lacking in both modified syllables. Another possibility is that, compared to song sparrows, swamp sparrows attend more to spectral differences among the syllables. Normal swamp-sparrow syllables have, on average, more energy at higher frequencies than do song-sparrow syllables. Removing the Type I and Type VI notes from a normal swamp-sparrow syllable effectively lowers its pitch. These alternatives could easily be tested with more elaborate modifications to these song syllables. From such an approach, it should be possible to construct "classification" rules by which each species perceives species-specific syllables (Okanoya & Dooling, 1988a).

In aggregate, these psychophysical results support results from other ethological studies on song learning in these two species, suggesting that swamp sparrows are more selective than song sparrows in the perception of conspecific syllables (Dooling & Searcy, 1980; Marler, 1987). A complete understanding of selective vocal learning may require both approaches. We already know from psychophysical procedures that these two species have similar auditory sensitivity (Okanoya & Dooling, 1988b). The fact that species differences in syllable perception emerge in spite of species similarities in basic auditory sensitivity indicates that the present techniques should prove useful for examining more integrative, top-down perceptual processes in these birds.

REFERENCES

- ALDENDERFER, M. S., & BLASHFIELD, R. K. (1984). *Cluster analysis*. Newbury Park, CA: Sage.
- CLARK, W. C., MARLER, P., & BEEMAN, K. (1987). Quantitative analysis of animal vocal phonology: An application to swamp sparrow song. *Ethology*, **76**, 101-115.
- DOOLING, R. J., & SEARCY, M. H. (1980). Early perceptual selectivity in the swamp sparrow. *Developmental Psychobiology*, **13**, 499-506.
- MARLER, P. (1987). Sensitive periods and the roles of specific and general sensory stimulation in birdsong learning. In J. P. Rauschecker & P. Marler (Eds.), *Imprinting and cortical plasticity* (pp. 99-136). New York: Wiley.
- MARLER, P., & PETERS, S. (1977). Selective vocal learning in a sparrow. *Science*, **198**, 519-521.
- MARLER, P., & PETERS, S. (1981). Subsong and plastic song: Their role in the vocal learning process. In D. E. Kroodsmas & E. H. Miller (Eds.), *Acoustic communication in birds* (pp. 25-50). New York: Academic Press.
- MARLER, P., & PETERS, S. (1989). Species differences in auditory responsiveness in early vocal learning. In R. J. Dooling & S. H. Hulse (Eds.), *The comparative psychology of audition: Perceiving complex sounds* (pp. 243-273). Hillsdale, NJ: Erlbaum.
- MARLER, P., & PICKERT, R. (1984). Species-universal micro-structure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Animal Behaviour*, **32**, 673-689.
- MARLER, P., & SHERMAN, V. (1985). Innate differences in singing behavior of sparrows reared in isolation from adult conspecific song. *Animal Behaviour*, **33**, 57-71.
- OKANOYA, K., & DOOLING, R. J. (1988a). Decision rules for perception of species-specific syllables in song and swamp sparrows. *Neural Networks: Abstracts of the First Annual INNS Meeting*, **1**, 314.
- OKANOYA, K., & DOOLING, R. J. (1988b). Hearing in the swamp sparrow, *Melospiza georgiana*, and the song sparrow, *Melospiza melodia*. *Animal Behaviour*, **36**, 726-732.
- OKANOYA, K., & DOOLING, R. J. (1988c). Obtaining acoustic similarity measures from animals: A method for species comparison. *Journal of the Acoustical Society of America*, **83**, 1690-1693.
- PETERS, S., SEARCY, W. A., & MARLER, P. (1980). Species song discrimination in choice experiments with territorial male swamp and song sparrows. *Animal Behaviour*, **28**, 393-404.
- SEARCY, S., MARLER, P., & PETERS, S. (1981). Species song discrimination in adult female song and swamp sparrows. *Animal Behaviour*, **29**, 997-1103.
- SHEPARD, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, **210**, 390-398.