

# Motor Stereotypy and Diversity in Songs of Mimic Thrushes

Roderick A. Suthers,<sup>1,2,3,4,\*</sup> Franz Goller,<sup>1,3</sup> and Rebecca S. Hartley<sup>4,†</sup>

<sup>1</sup>Medical Sciences Program, <sup>2</sup>Program for Neural Science, <sup>3</sup>Center for the Integrative Study of Animal Behavior, and <sup>4</sup>Department of Biology, Indiana University, Bloomington, Indiana 47405

## SUMMARY

The relationship between the motor and acoustic similarity of song was examined in brown thrashers (*Toxostoma rufum*) and grey catbirds (*Dumetella carolinensis*) (family Mimidae), which have very large song repertoires and sometimes mimic other species. Motor similarity was assessed by cross correlation of syringeal airflows and air sac pressures that accompany sound production. Although most syllables were sung only once in the song analyzed, some were repeated, either immediately forming a couplet, or after a period of intervening song, as a distant repetition. Both couplets and distant repetitions are produced by distinctive, stereotyped motor patterns. Their motor similarity does not decrease as the time interval between repetitions increases, suggesting that repeated syllables are stored in memory as fixed motor programs. The acoustic similarity between nonrepeated syllables, as indicated by

correlation of their spectrograms, has a significant positive correlation with their motor similarity. This correlation is weak, however, suggesting that there is no simple linear relationship between motor action and acoustic output and that similar sounds may sometimes be produced by different motor mechanisms. When compared without regard to the sequence in which they are sung, syllables paired for maximum spectral similarity form a continuum with repeated syllables in terms of their acoustic and motor similarity. The prominence of couplets in the "syntax" of normal song is enhanced by the dissimilarity of successive nonrepeated syllables that make up the remainder of the song. © 1996 John Wiley & Sons, Inc.

**Keywords:** motor control, syrinx, bird song, vocalization, respiration.

## INTRODUCTION

The production of bird song is controlled by a system of central song control nuclei that generate the motor output to the avian vocal organ, the syrinx (Nottebohm et al., 1976; Nottebohm, 1991). The manner in which song is represented in these nuclei is still poorly understood. The adult song is acquired through a process of trial and error motor learning that requires auditory feedback and previous exposure to conspecific song during a sensitive

period in development (Konishi, 1985; Marler, 1991). The size and plasticity of the adult song repertoire varies greatly according to the species, as must the corresponding demands on the motor organization and control of song. An indication of how motor organization is influenced by repertoire size may be obtained by comparing the motor basis of syllable production in species with small versus those with large repertoires. Two such species are among the few in which the peripheral motor aspects of song have been directly investigated.

The brown-headed cowbird (*Molothrus ater ater*) has a limited, stereotyped song repertoire. An adult cowbird's song consists of about half a dozen different song types (King et al., 1980; West et al., 1983). The motor basis of song production in cowbirds was studied by Allan and Suthers (1994) who showed, by monitoring syringeal airflow and respi-

Received February 22, 1995; accepted December 12, 1995  
Journal of Neurobiology, Vol. 30, No. 2, pp. 231–245 (1996)  
© 1996 John Wiley & Sons, Inc.  
CCC 0022-3034/96/020231-15

\* To whom correspondence should be addressed.

† Present address: Department of Biology, Seattle University, Seattle, WA 98122.

ratory pressure during song, that each song type is produced by its own stereotyped motor pattern to which each side of the syrinx contributes certain notes.

Mimic thrushes (Mimidae), on the other hand, have very large and highly variable repertoires. The brown thrasher and grey catbird may sing two or three syllables per second continuously for many minutes, or even hours, occasionally mimicking other species (Kroodsmma and Baylis, 1982). Borror (1964) identified 117 phrases from an individual catbird; Thompson and Jane (1969) estimated their repertoire at 180 syllable types; and Lohr (1989) estimated between 70 and 130 syllable patterns. Syllables are infrequently repeated, except for single immediate repetitions to form "couplets," which are much more common in thrasher than in catbird song. Playback experiments suggest that syllable morphology (Fletcher and Smith, 1978) and overall temporal patterning (Boughey and Thompson, 1976) may be important in identifying conspecifics. The songs of individual catbirds contain unique syllable patterns that enable one to identify the singer from sound spectrograms of its song (Lohr, 1989). Brown thrashers appear to have even larger repertoires than catbirds. Kroodsmma and Parker (1977) estimated an individual thrasher may sing at least 1,800 song units (some of which include more than one syllable) and Boughey and Thompson (1981) estimated 2400 syllables (= "figure types"). Little is known about the ontogenetic development of mimid song or the extent to which vocal learning continues throughout life.

The respiratory and syringeal mechanisms underlying thrasher and catbird song were studied with the aid of bronchial thermistors and air sac pressure transducers that permit the respiratory dynamics and sound generated on each side of the syrinx to be monitored. As in other songbirds studied, each side of the syrinx is controlled independently of the other by the ipsilateral tracheosyringeal branch of the hypoglossal nerve. During song these birds rapidly switch phonation from side to side, even within syllables, and often use both sides of the syrinx simultaneously to produce independently modulated "two voice" syllables. Both sides of the syrinx make more or less equal overall contributions to the total song; but often one or the other side is closed through adductive action of syringeal muscles, and during this time sound is generated only on the contralateral side (Suthers, 1990; Suthers et al., 1994; Goller and Suthers, 1995).

The vocal and motor versatility of thrashers and catbirds make them especially interesting subjects in which to study the motor correlates responsible for the phonetic diversity of song. Here we examine the relationship between motor and vocal diversity in these birds to better understand the motor organization of their very large song repertoires. Because most syllables are repeated infrequently, if at all, it is not known if each of the hundreds of different syllable types is stored in the brain as a discrete motor program or if they are improvised during song. Does the bird "reinvent" the same motor gesture each time it repeats a syllable or is it able to produce essentially the same sound in different ways? When a syllable is repeated does the stereotypy of the motor pattern decrease as a function of the amount of time that has elapsed between repetitions, suggesting it is stored only in short-term memory, or is there no degradation of motor stereotypy with the passage of time, suggesting storage in long-term memory? A better understanding of the rules that govern the relationship between motor and vocal versatility may also indicate if motor constraints affect the sequence in which syllables are sung and would be of value to behavioral investigations concerning various aspects of vocal communication.

## MATERIALS AND METHODS

### General Methods

Experiments were conducted on three adult male grey catbirds and six adult male brown thrashers that were hand-reared in the laboratory from the age of about 1 week and tutored with recordings of conspecific song. Individuals were not acoustically or visually isolated from one another and were also exposed to songs of each other and of northern cardinals (*Cardinalis cardinalis*), rufous-sided towhees (*Pipilo erythrophthalmus*), and canaries (*Serinus canaria*). They were maintained on a photoperiod that simulated the seasonal changes in day length. Prior to an experiment a 3-week timed release pellet of testosterone propionate (10 mg, catbird; 15 mg, thrasher; Innovative Research of America) was implanted subcutaneously to increase the frequency of song. The reader is referred to Suthers et al. (1994) for a detailed account of the methods used in this study. Briefly, surgical anesthesia was achieved with either isoflurane, halothane, or chloropent (4–5  $\mu$ L/g, i.m.; Fort Dodge Laboratory). The pressure driving air through the two sides of the syrinx was monitored via a piezoresistive pressure transducer (Endevco, model 8507B-5 or Fujikura, model FPM-02PG), mounted on the bird's back, attached to a silicone cannula in the cra-

nial thoracic air sac. The rate of airflow through the left and right sides of the syrinx and the flow oscillation associated with the sound generated on each side were recorded by implanting a 5 mil (0.13 mm) microbead thermistor near the center of each primary bronchial lumen. The rate of airflow cannot be calibrated because the cross-sectional area of the bronchial lumen is likely to vary during song. Nonlinearity in the output of the thermistor should not affect the comparisons between different syllables of the same bird significantly. The leads from each thermistor were routed under the bird's skin to a microconnector on its back from which they traveled, together with leads from the pressure transducer, out through the top of the cage to electronic conditioning and recording instruments. These leads formed an adjustable leash that allowed the bird to move freely within its cage. An electronic feedback circuit heated the thermistors and maintained them at a constant temperature. Song was recorded on a microphone (Panasonic, WM-2298) mounted in front of the cage.

Tape recorded data on sound, rates of airflow, and air sac pressure were displayed using an eight channel thermal array recorder (Graptect model WR3600) together with spectrographs of the emitted song and bronchial sounds (Uniscan II, Multigon, Mt. Vernon, NY). Data were also analyzed on a microcomputer using *Signal software v. 2.2 (Engineering Design, Belmont, MA)*. Tape recorded data were reproduced at half speed through an antialias filter (TTE Inc, series J87) with a high cutoff at 8 kHz and digitized at 20 kHz/channel (Data Translation, DT-2821-G).

### Identification of Repeated Syllables

Segments of song containing between about 125 and 300 syllables from each of four thrashers and three catbirds, were selected on the basis of a high quality sound recording and good signal-to-noise ratio. The song unit that we refer to as a syllable is produced during one expiration and is equivalent to a "figure" as defined by Boughey and Thompson (1976, 1981). Sound spectrograms of all syllables in each song sample were compared visually and the syllables that were repeated within the sample were identified on the basis of their similar spectrograms. The high degree of acoustic similarity between repeated syllables contrasts with the variability of other syllables in mimic thrush song, making repeated syllables easy to identify by visual comparison of spectrograms. No information on the patterns of airflow or pressure was available when repeated syllables were identified.

### Computation of Motor and Spectrographic Correlates

Not all birds gave good recordings of air sac pressure and bilateral airflow. In some experiments, for example, mucus or other foreign material became deposited on a thermistor, degrading its time constant, or the fine wires

between the thermistor and the backpack broke. In the case of thrasher 1 we used a prototype "septal" tracheal flow probe (see Suthers et al., 1994) that gave an inferior measure of bronchial airflow. Experiments on three thrashers and two catbirds that also yielded consistently high quality measures of air sac pressure and bilateral syringeal airflow were further analyzed to assess the degree to which similar vocalizations are generated by similar motor activity.

Pearson correlation coefficients (XCR) were calculated (Signal v. 2.2) for matrices of 20 syllables at a time, produced by the same individual. The XCR provides a measure of similarity where 0 equals no similarity and 1 equals total similarity. In the case of airflows and pressure, the two signals being cross correlated were adjusted to the same length by extending the shorter buffers with trailing zeros. Prior to cross correlation, the members of each syllable pair were aligned so that corresponding rates of airflow or pressure were cross correlated at the same alignment, namely that at which the rectified envelopes of the two vocalizations had the highest correlation coefficient.

Spectrograms were cross correlated to assess the similarity of the vocalizations. The spectrogram correlation coefficient was calculated in Signal as a two-dimensional cross correlation between two digital spectrograms. Both spectrograms always had the same frequency range and were made of equal length by adding trailing zeros to the shorter buffer. The fast Fourier transform step size was 0.5 ms or 512 points, time resolution was 10.2 ms, and frequency resolution was 97.7 Hz. The upper and lower intensity thresholds were adjusted to minimize background noise and were the same for all comparisons.

For all possible syllable pairs, the XCRs for the left and right flow rates and subsyringeal pressure were averaged to calculate a mean motor XCR, or *motor correlate*, reflecting these motor indices of syllable production. The 20 syllables in each correlation matrix were paired and ranked either according to the best match of their spectrographic XCR or the best match of their motor XCR.

### Assumptions Underlying Cross-Correlation Analysis

The spectrographic and motor correlates, upon which much of our analysis is based, have the advantage of providing an overall index for the properties of the vocal output and its underlying motor activity. Cross correlation of time-frequency spectrograms ensures that both the temporal and frequency structure of each syllable are included.

The measure of motor activity based on subsyringeal pressure and the rate of airflow through each side of the syrinx, provides an index of motor action that includes the effect of both respiratory and syringeal muscle activity during vocalization (Suthers, 1990; Suthers et al., 1994; Goller and Suthers, 1995, 1996). Knowledge of these two variables enables one to monitor changes in

syringeal resistance, which is determined by the activity of syringeal adductor and abductor muscles and respiratory muscles. Resistance increases during phonation as adductor muscles move the lateral labium and/or medial tympaniform membranes into the syringeal lumen. It is also possible to determine the side or sides of the syrinx producing sound, because unilateral sound production is accomplished by full adduction (zero flow despite positive pressure) on the contralateral side that is thus muted due to the absence of airflow across the syringeal membranes. Respiratory muscle activity is coordinated with that of syringeal muscles and equally important in song production (Hartley, 1990; Simpson and Vicario, 1990; Vicario, 1991; Williams et al., 1992). In addition to setting the overall temporal pattern of song, respiratory muscles modulate its acoustic properties. The interaction between respiratory and syringeal muscles in vocal control can be deduced from the relationship between airflow and pressure because in the absence of syringeal adduction or abduction, airflow is proportional to the air sac pressure generated by expiratory muscles (Suthers et al., 1994).

There are, however, limitations in this measure of acoustic and motor similarity. Spectrograms can be distorted by background noise or echoes from nearby objects and even small differences in fundamental frequency are amplified by harmonics if they are present. The motor correlate assumes that similar changes in syringeal airflow and pressure are produced by similar motor patterns. It omits the activity of syringeal muscles that may regulate sound frequency by varying the tension on the tympaniform membranes, as well as that of other muscles that may affect the filter properties of the supra-syringeal vocal tract (Nowicki, 1987; Westneat et al., 1993), unless this activity also affects syringeal airflow or pressure.

Although a more direct measure of motor similarity might be achieved by a detailed comparison of the electromyograms of muscles involved in song production, only a few of these muscles can be recorded at one time. This practical limitation of electromyography is especially troublesome in assessing the motor stereotypy of mimic thrushes because many syllables are seldom repeated. The motor correlate based on syringeal pressure and airflow that we use here, despite its shortcomings, does provide an overall index of the summed mechanical action of both respiratory and syringeal muscles. Additional electromyographic data would provide a more sensitive measure of additional differences between syllables but would not negate the differences detected by our current technique.

## RESULTS

### Occurrence of Syllable Repetition

Syllables from catbirds and brown thrashers were classified, by visual inspection of spectrograms,

**Table 1** Occurrence of Repeated Syllables

Bird	Syllable Types		
	Total	Appearing in Couplets	Appearing as Distant Repetitions
TH NR	126	5 (4.0%)	3 (2.4%)
TH 1	172	18 (10.5%)	5 (2.9%)
TH 6	293	17 (5.8%)	22 (7.5%)
TH 18	119	10 (8.4%)	8 (6.7%)
CT 1	135	2 (1.5%)	4 (3.0%)
CT 2	201	5 (2.5%)	11 (5.5%)
CT 5	125	1 (0.8%)	1 (0.8%)

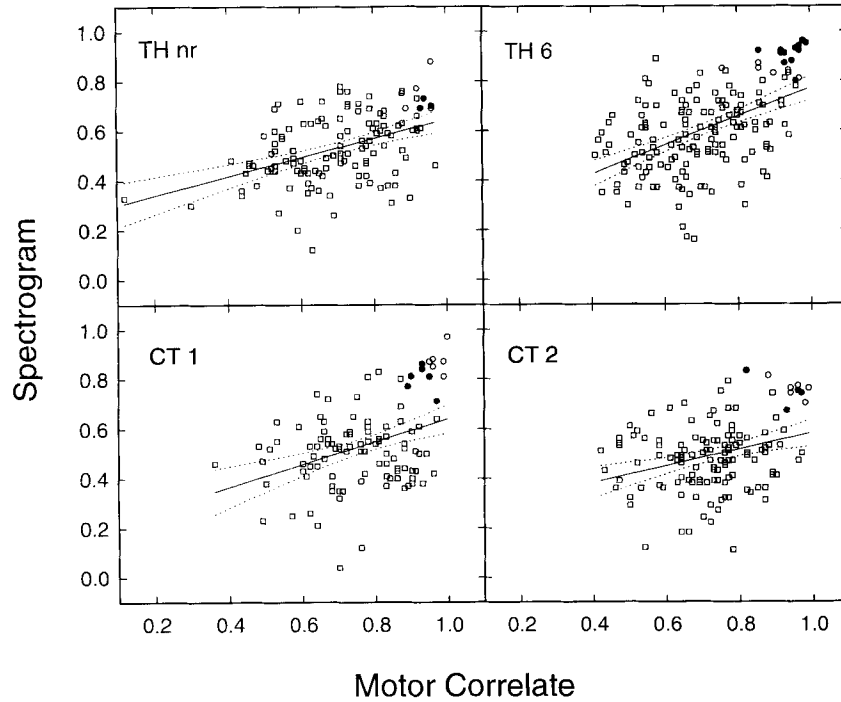
into one of three groups according to the occurrence and timing of their repetition in the song sample: nonrepeated syllables that were sung only once in the sample of song analyzed; couplets resulting from the immediate repetition of a syllable to produce a pair of almost identical vocalizations that are not separated from each other by any intervening song; and distant repetitions in which a syllable is repeated after an intervening segment of song made up of other syllables.

Repeated syllables, consisting of couplets and distant repetitions, comprise only 2–15% of the syllables sampled from three catbirds and four thrashers (Table 1). All individuals sang both couplets and distant repetitions, although most couplets were sung by brown thrashers. Birds of both species sometimes repeated a syllable after an intervening period of song composed of dissimilar syllables.

### Relationship between Vocal and Motor Similarity

To assess the relationship between vocal and motor variability, the acoustic and motor data for each syllable in the song sample were cross correlated with that of every other syllable in each matrix of 20 syllables, regardless of the sequence in which they were sung, and the spectrographic and motor correlates were computed for each of these comparisons. Syllables were then paired with whichever other syllable gave the highest spectrographic correlate (Fig. 1) to determine if there is a positive correlation between the acoustic and motor similarity, consistent with the hypothesis that a particular motor pattern gives rise to a unique, predictable acoustic output or phonetic structure.

When syllables were paired by this criterion, the degree of spectrogram and motor similarity has a significant, although relatively weak, positive cor-



**Figure 1** Motor and acoustic similarity of song syllables paired according to most similar spectrograms. (□) Syllables not repeated in song sample; (○) couplets; (●) distant matches. Broken lines indicate 95% confidence intervals. Values of  $r^2$  are: TH nr = 0.18; TH 6 = 0.26; CT 1 = 0.14; CT 2 = 0.095.  $p < 0.001$  for each bird. TH, brown thrasher; CT, grey catbird.

relation (Fig. 1). In five birds, mean motor XCRs of nonrepeated syllable pairs ranged from 0.68 to 0.78 and mean spectrogram XCRs varied from 0.47 to 0.58. Both the mean motor and spectrographic XCRs of nonrepeated syllables are significantly ( $p < 0.001$ ) lower than those of repeated syllables (Table 2). The positive slope of data in Fig-

ure 1 indicates a tendency for spectrographically similar syllables to also share similar patterns of air-flow and pressure, but the coefficients of determination are low and the divergent values of motor and spectrogram XCRs in a number of syllables suggest that a bird may sometimes generate spectrographically similar sounds in different ways (Fig. 1).

**Table 2** Cross-Correlation Coefficients of Matched Syllable Pairs

Bird	Motor Correlate			Spectrographic Correlate		
	Nonrepeated	Couplet	Distant Repetitions	Nonrepeated	Couplet	Distant Repetitions
TH NR	0.69 ± 0.15* (118)	0.92 ± 0.03 (5)	0.94 ± 0.01 (3)	0.52 ± 0.13*	0.76 ± 0.07	0.71 ± 0.02
TH 6	0.68 ± 0.13* (142)	0.89 ± 0.06† (7)	0.94 ± 0.03 (12)	0.56 ± 0.14*	0.84 ± 0.02*	0.91 ± 0.04
TH 18	0.78 ± 0.14* (94)	0.92 ± 0.06 (11)	0.93 ± 0.02 (3)	0.58 ± 0.12*	0.75 ± 0.08	0.72 ± 0.1
CT 1	0.74 ± 0.13* (98)	0.97 ± 0.02† (6)	0.93 ± 0.03 (6)	0.50 ± 0.13*	0.87 ± 0.05†	0.80 ± 0.05
CT 2	0.71 ± 0.12* (120)	0.95 ± 0.03 (6)	0.92 ± 0.75 (4)	0.47 ± 0.12*	0.76 ± 0.03	0.75 ± 0.06

Syllables paired with best match for spectrographic similarity. Mean ± standard deviation. Number of pairs in parenthesis.

\*  $p < 0.01$ .

†  $0.05 > p > 0.01$ , two-tailed  $t$  test of significance between neighboring columns.

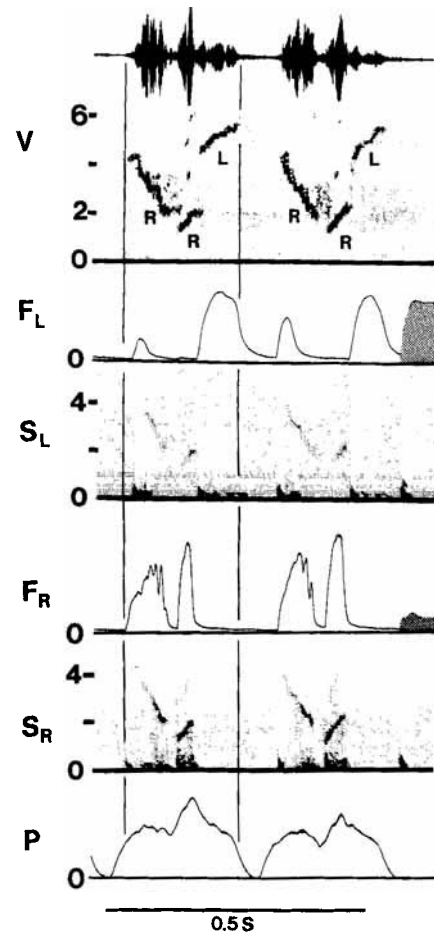
When these syllables were paired according to the best match of their motor, rather than spectrographic, correlate, there was still a significant positive correlation between the spectral and motor correlates in all birds ( $p < 0.001$ ; except CT 2,  $p < 0.01$ ). Mean correlation coefficients for individual birds ranged from 0.224 to 0.387, somewhat lower than when syllables were paired for best spectrogram match.

### Motor Similarity of Repeated Syllables

Syllable repetition is always accompanied by the repetition of the particular distinctive, stereotyped motor act responsible for producing that syllable type. Distant repetitions and couplets have a mean motor XCR of about 0.9 or higher and a mean spectrogram XCR of at least 0.7, significantly higher than the corresponding values of syllables that are not repeated (Table 2). In most cases, the XCR of distant repetitions is not significantly different from that of couplets. The only exceptions are TH 6 for which both the motor and spectrographic correlates of distant repetitions are significantly higher than those of couplets, and CT 1 for which both of these correlates are significantly lower for distant repetitions compared to that of couplets (Table 2).

**Couplets.** The pairs of syllables that form couplets are very similar in their phonology and in terms of the patterns of syringeal airflow and pressure that accompany their production (Fig. 1, Table 2). They are easily identified by listening to the song or by visual inspection of spectrograms. A couplet containing steep frequency modulated (FM) notes contributed first by the right and then by the left side of the syrinx is shown in Figure 2. Although there is no inspiration between the first and second rendition of the syllable in this example (air sac pressure drops but both sides of the syrinx remain closed), an inspiration occurs between the members of some couplets. In Figure 3 the syllables of four couplets having less FM, but containing two-voice components, have been separated and aligned vertically to more easily compare their airflow and pressure waveforms.

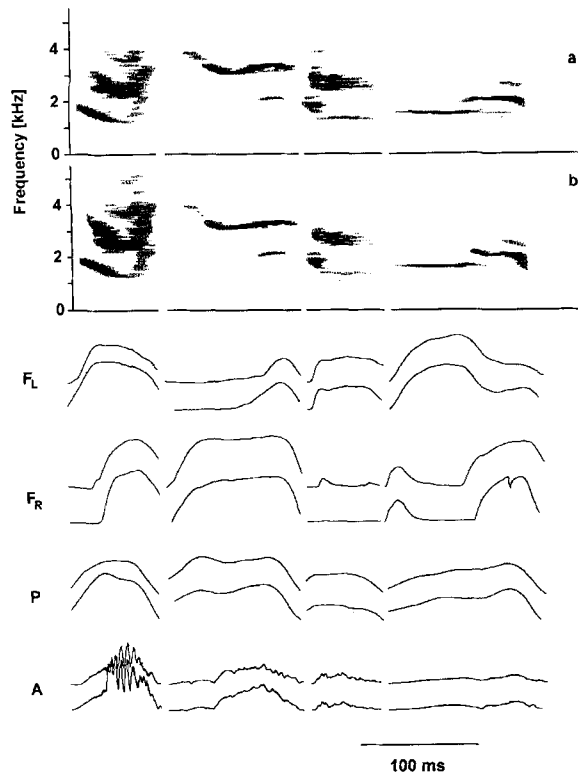
Slight differences in the temporal pattern of airflow or pressure correspond to small differences in the emitted sounds. For example, differences in the linearity in the first and second rendition of the first downward FM note in the couplet shown in Figure 2 coincide with differing inflections in the rate of airflow through the right syrinx. In the last couplet



**Figure 2** Couplet from brown thrasher 6. Inspiration and expiration both cause upward deflection of airflow trace. Inspiration (stippled) occurs during negative air sac pressure. The first two elements of each syllable are produced on the right side of the syrinx, as indicated by ipsilateral bronchial sound and airflow. The third element is generated on the left side; its sound frequency is too high for the time constant of the bronchial thermistor but the right syrinx is closed (positive pressure but no airflow). V, oscillograph and sonograph of emitted vocalization; R, elements produced by right side of syrinx; L, elements produced by left side of syrinx;  $F_L$  and  $F_R$ , rate of airflow through left and right side of syrinx, respectively;  $S_L$  and  $S_R$ , sound in left and right primary bronchus, respectively; P, pressure in the cranial thoracic air sac. Sound frequency in kilohertz. Vertical lines align events at beginning and end of first syllable.

of Figure 3, the second note, which is produced on the right side at an almost constant frequency of about 2 kHz, begins a little earlier in the first rendition of this syllable, corresponding to an earlier onset of airflow through the right syrinx.

Within the syntax of the song, couplets stand out as matched pairs of syllables that are clearly dis-



**Figure 3** Four couplets shown as vertically offset pairs. (a) Upper spectrograph and trace are first member of the couplet; (b) lower spectrograph and trace are second member. Spectrograms are of emitted vocalization. First and second couplets are TH 22; third couplet CT 2; fourth couplet TH 10. A, rectified amplitude of emitted syllables; other symbols as in legend of Figure 2. For each couplet, the correlation coefficients for  $F_L$ ,  $F_R$ ,  $P$ , and  $A$  are 0.99 or higher.

tinct from the acoustically diverse nonrepeated syllables that comprise a major portion of the songs of these birds. The couplet's prominence as a distinctive feature of the song arises not only from the stereotypical repetition of its syllable, but also from the fact that noncouplet syllables tend to maintain a high degree of acoustic distinctness (i.e., low spectrographic XCR) with respect to their immediate neighbors in the song.

The importance of syntax is evident when one compares the spectral and motor correlates of the two syllables that make up the couplet, with those of neighboring noncouplet syllables in the remainder of the song. Couplets have a mean motor correlate of  $0.95 \pm 0.03$  and mean spectral correlate of  $0.82 \pm 0.06$ . These values are significantly ( $p < 0.01$ ) higher than those of neighboring noncouplet syllables that have mean motor correlates between 0.60 and 0.72 and mean spectral correlates be-

tween 0.18 and 0.28 (Table 3). The spectrogram correlates of best and worst matches are all significantly different from those of neighboring syllables. There is no significant difference between the motor correlates of neighboring syllables and those of the worst spectrographic matches, but the motor correlates of the best spectrographic matches are significantly higher than those of the neighboring syllables for all birds except catbird 1 (Table 3). The distinctive nature of the couplet is graphically illustrated in Figure 4 that also indicates the consistently low correlates of sequentially paired noncouplet syllables.

**Distant Repetitions.** Fifty-one distant repetitions were visually identified on the basis of their spectrograms. As in the case of couplets, the accompanying pattern of syringeal pressure and airflow are very similar (Table 2). Many of these syllables are repeated only once in the song sample. A two-voice syllable, repeated after singing 38 other syllables, shows almost identical patterns of syringeal airflow and pressure to produce a frequency modulated left voice superimposed on a nearly constant frequency right voice (Fig. 5). Even cyclical amplitude modulated (AM) syllables (Suthers et al., 1994), during which syringeal muscles cyclically change the aperture or resistance of the right side of the syrinx, are reproduced with considerable precision, although the number of adduction cycles in the right syrinx sometimes varies slightly (Fig. 6, syllables a, b).

Some syllables were repeated more than once, up to a maximum of four distant repetitions in the sample of song studied. One syllable with a complex spectral pattern was reproduced three different times during a recording period; the second and third repetitions of this syllable were followed immediately by another distantly repeated syllable (Fig. 6).

### Disparities between Motor and Spectrographic Similarity

It is informative to examine those pairs of syllables represented by outlying points in Figure 1 for which either the motor or spectrographic XCR is high but the other measure of similarity is low. Disparate values for spectrographic and motor XCRs are sometimes due to extraneous factors not directly related to the motor control of the vocalization. These include echoes or noise in the spectrograms and timing differences in the opening or closing of one side of the syrinx during inspiration

**Table 3** Motor and Spectral Similarity of Neighboring Syllables and Those Paired for Best or Worst Spectrographic Match

Bird	Best Spectrographic Match		Neighbors		Worst Spectrographic Match	
	Motor Correlate	Spectrogram Correlate	Motor Correlate	Spectrogram Correlate	Motor Correlate	Spectrogram Correlate
TH NR	0.69 ± 0.15*	0.52 ± 0.13*	0.62 ± 0.15	0.28 ± 0.16	0.59 ± 0.14	0.05 ± 0.04*
TH 6	0.68 ± 0.13*	0.56 ± 0.14*	0.60 ± 0.12	0.19 ± 0.21	0.58 ± 0.12	0.00 ± 0.01*
TH 18	0.78 ± 0.14†	0.58 ± 0.12*	0.72 ± 0.16	0.18 ± 0.19	0.71 ± 0.13	0.01 ± 0.03*
CT 1	0.74 ± 0.13	0.50 ± 0.13*	0.72 ± 0.13	0.24 ± 0.17	0.70 ± 0.14	0.02 ± 0.03*
CT 2	0.71 ± 0.12†	0.47 ± 0.12*	0.67 ± 0.13	0.19 ± 0.19	0.66 ± 0.12	0.00 ± 0.01*

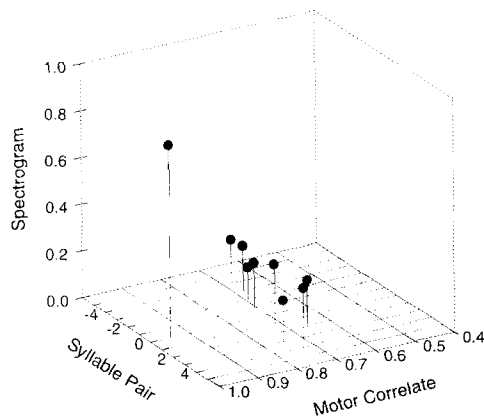
Mean motor and spectrogram correlates ± standard deviation.  $n = 94-142$ .

\*  $p < 0.01$ .

†  $0.05 > p > 0.01$  compared to neighbors (two tailed  $t$  test couplet). When the motor correlates of best and worst spectrographic matches are compared,  $p < 0.01$ .

that can significantly affect the XCR, while having a minimal effect on sound production. In other instances, differences in the motor or acoustic correlates clearly reflect different phonatory mechanisms.

A pair of syllables having a high spectrographic, but low motor, XCR is shown in Figure 7. Despite



**Figure 4** The spectrographic and motor similarity of couplets (syllable pair 0) and of the adjacent preceding (-1 to -4) and succeeding (1 to 4) syllable pairs. Each syllable was compared with its neighbors for motor and acoustic similarity. The mean motor and spectral correlates of the couplets is  $0.95 \pm 0.03$  and  $0.82 \pm 0.06$ , respectively. The mean motor ( $0.70 \pm 0.03$ ) and spectral ( $0.20 \pm 0.04$ ) correlates of neighboring syllables are not significantly different except for a member of the couplet pair and its noncouplet neighbor ( $p < 0.0001$ ). Syllable pair -1 is comprised of the first member of the couplet and the preceding syllable; pair -2 is comprised of the two syllables immediately preceding the couplet, etc. Each point indicates the average XCRs for a pair of adjacent syllables in song sequences centered on 13 couplets from five thrashers and three couplets from two catbirds.

a general similarity in the vocal output of these two-voice syllables, there is an obvious difference in the onset time and duration of the component at about 2.2 kHz. This acoustic difference is the result of different patterns of motor action in both the respiratory and the syringeal muscles. In the case of syllable "a" (upper traces) the expiratory muscles maintain an air sac pressure that rises to a peak during the first third of the vocalization and then gradually declines. The pressure accompanying syllable "b" (lower traces) has two peaks separated by a dip in the middle of the vocalization. This bimodal pressure pattern causes a parallel change in sound intensity and the rate of airflow through the right syrinx, but only a slight variation in sound frequency that is not enough to degrade the spectral correlation coefficient. The motor pattern to the syringeal muscles also differs. In syllable "a" the lowest frequency component (1.7 kHz) is generated on the left side of the syrinx and the 2.2-kHz component begins about 60 ms later when the right side of the syrinx opens. In the case of syllable "b," the left side of the syrinx is closed during most of the vocalization. Here the 1.7-kHz component is generated on the right side of the syrinx and a brief overlapping element at 2.2 kHz is produced by the left side when it opens just before the end of the syllable. Thus both the syringeal muscles that control the timing of sound production by opening or closing the syrinx and those controlling the frequency of the vocalization, receive different temporal patterns of activation in these two syllables. It is particularly interesting that there is a reversal of the side producing the two major, nearly constant frequency elements in these syllables. The bird produces two similar syllables but the corresponding frequency elements are generated on the opposite sides of the syrinx.



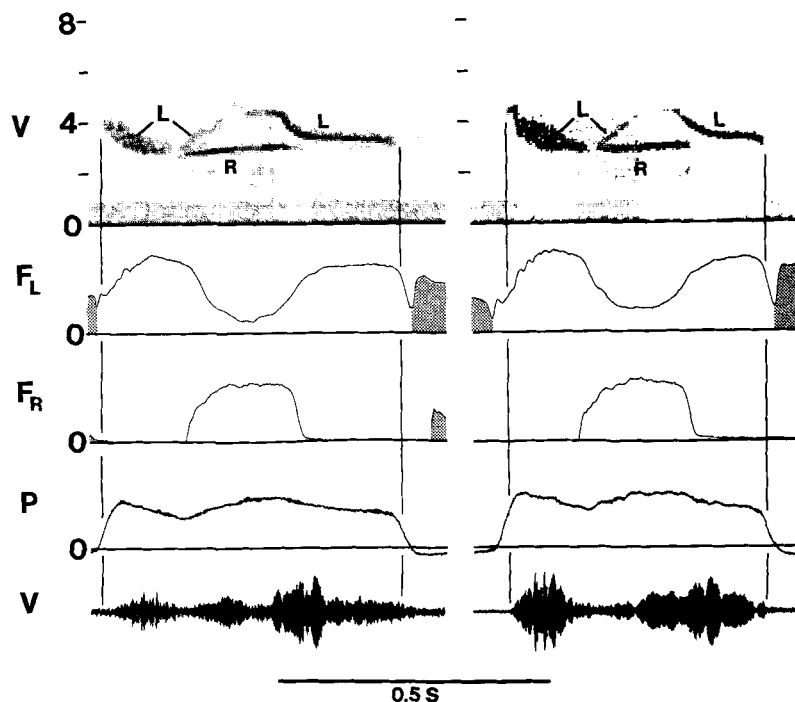


Figure 5 Distant repetition of a two-voice syllable separated by 38 intervening syllables. Thrasher 6.

Three examples of syllables that have high scores for motor similarity but low scores for spectrographic similarity are shown in Figure 8. In each case, syllables "a" and "b" were produced with very similar patterns of subsyringeal pressure, temporal patterns of gating syringeal airflow, and relative changes in the rate of airflow through each side of the syrinx. The acoustic structure of the syllables comprising each pair differ, however, particularly with regard to their spectral content. The spectral range of syllable "b" is shifted upward in the first two pairs and downward in the third pair of syllables. This frequency shift is probably due to differences in the activity of syringeal muscles controlling tension on the medial tympaniform membranes and is therefore not included in our measure of motor similarity.

## DISCUSSION

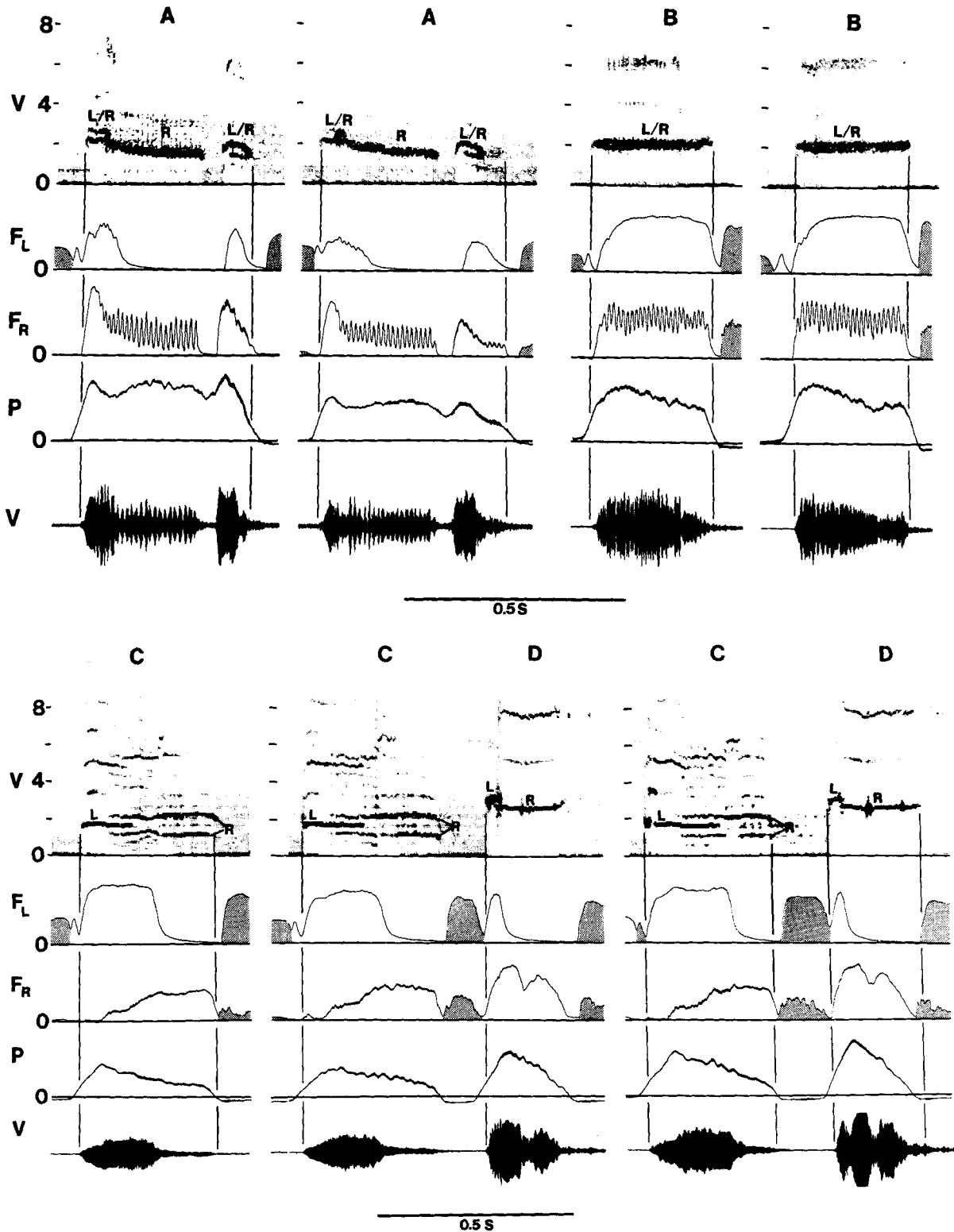
### Motor and Acoustic Variability in Nonrepeated Syllables

The phonetic variety that characterizes the songs of mimic thrushes makes them well suited to investigate the relationship between motor action and acoustic output. How many different ways, for ex-

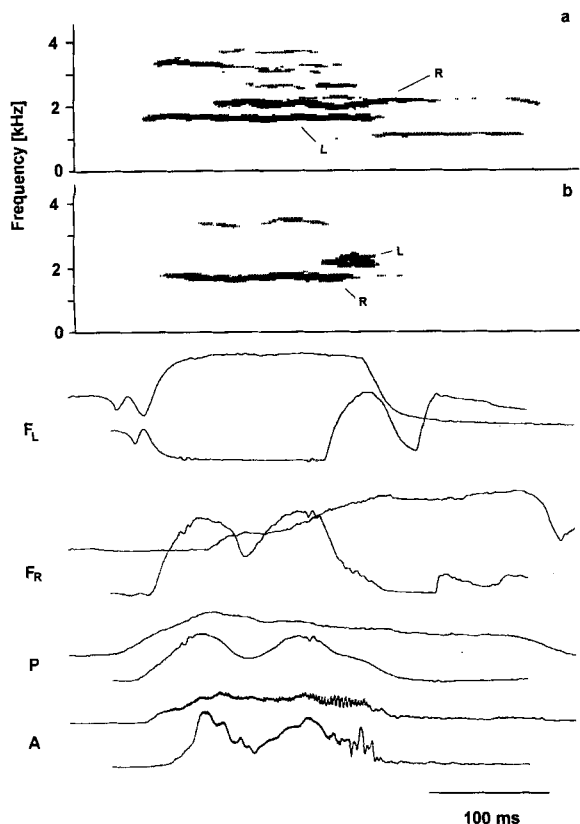
ample, can the oscine syrinx and its associated vocal system generate a particular syllable? Because the phonetic diversity of mimic thrush song provides a wide range of acoustic structure in an ever changing context, a search for correlations between acoustic and motor similarity may shed light on the relationship between motor action and vocal output. If there is a simple set of motor rules according to which each acoustic feature of song is produced by a single specific motor action, then acoustically similar syllables should be generated by correspondingly similar patterns of motor activity. Conversely, if there are multiple motor solutions for producing each acoustic feature or phonetic element of song, then motor pattern similarity should not be correlated with acoustic similarity.

The finding that different syllable types in a song repertoire are each produced by a distinctive stereotyped motor pattern does not exclude the possibility that some other motor pattern(s) could be used to produce the same syllable. Even if an individual adult bird uses a consistent, stereotyped motor pattern for each syllable type it sings, these motor patterns might simply be the first of several possible alternative motor gestures, all producing the same vocal output, that the bird "discovered" by trial and error motor learning during song acquisition.

Williams and Nottebohm (1985) found that



**Figure 6** Syllable A distant repetitions of a syllable pair containing rapid AM, generated on the right side of the syrinx, followed immediately by a brief two-voice syllable. This pair was repeated after 34 intervening syllables. Rapid AM syllable B was repeated after 94 intervening syllables. Syllable C was repeated after 29 syllables and again after 17 syllables. Both sides of the syrinx contribute to the entire vocalization. In each case the first part of the syllable is produced on the left and the last part is generated on the right side with considerable overlap producing a two-voice portion in the middle. The second and third repetitions were followed immediately by syllable D. Thrasher 6.

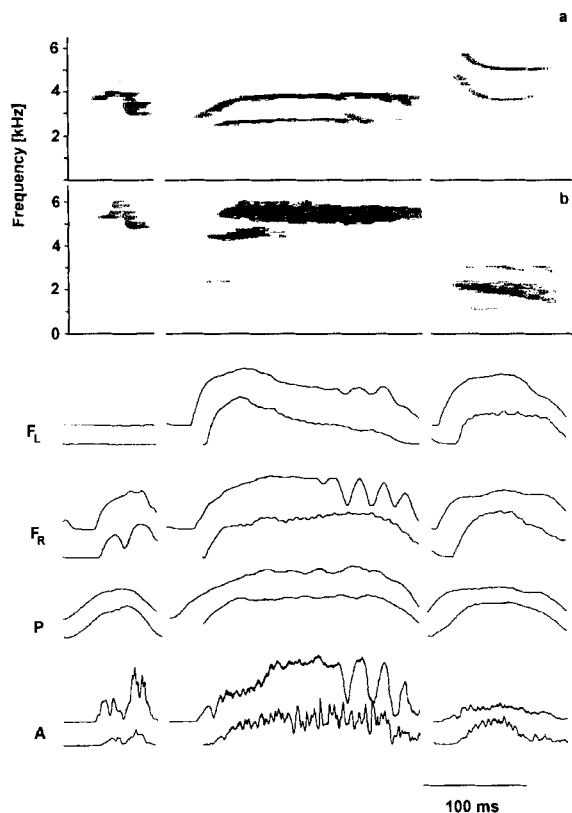


**Figure 7** Syllables “a” and “b” have relatively high spectrographic correlates ( $r = 0.81$ ) but low motor correlates ( $r = 0.65$ ). The low frequency component at about 1.7 kHz is generated on the left side of the syrinx in example “a” but on the right side of the syrinx in “b.” Thrasher 6. Symbols as in Figure 2.

different hypoglossal motor neurons innervating the syrinx in male zebra finches (*Taeniopygia guttata*) showed selective long latency responses to various elements of conspecific song. They hypothesized that songbirds may perceive song as a series of articulatory gestures in a manner analogous to that suggested by Liberman et al. (1967) in their motor theory of speech perception, put forth to explain the fact that acoustically different elements of speech may sometimes be perceived as a single phoneme.

If songbirds are perceiving song as articulatory gestures, this implies that a unique vocal gesture exists for each phoneme. The relationship between vocal and motor similarity of nonrepeated syllables, which we have found in catbirds and thrashers, suggests that, at least in the case of mimic thrushes, this may not be true. If there is a single distinctive motor action by which each syllable can be produced, one would expect a stronger correlation between vocal and motor correlates of similar

syllables. If, at the other extreme, similar syllables can be generated by an essentially unlimited number of different motor patterns, there should be no significant correlation between vocal and motor correlates. The weak correlation between the motor and spectrographic correlates of nonrepeated syllables matched for either acoustic (Fig. 1) or motor similarity, is consistent with the hypothesis that acoustically similar syllables are sometimes generated using dissimilar motor patterns. We focused on syllables matched for acoustic similarity because the spectrographic correlation coefficients vary over a wider range than do motor correlates and so provide a more sensitive index of syllable similarity. There may be only one motor pattern by which some syllables or components of syllables can be produced but two or more alternative motor patterns capable of generating other syllables, depending on the anatomical and physical constraints of the vocal system. A better understanding



**Figure 8** Three syllables “a” and “b” have relatively high motor correlates ( $r = 0.94$ – $0.96$ ) but low spectrographic correlates ( $r = 0.0$ ). The frequency structure of each pair is similar but syllables in “a” are shifted downward (first two syllables) or upward (third syllable) compared to counterpart in “b.” First syllable from catbird 2; second and third syllables from thrasher 10.

of this relationship awaits a more sensitive measure of motor and acoustic similarity.

The motor dynamics of song production in mimic thrushes suggest some ways in which different motor actions might yield similar vocalizations (Suthers, 1990; Suthers et al., 1994). During many of their song syllables, thrashers and catbirds switch phonation from one side of the syrinx to the other or from one side to both sides. It is conceivable, for example, that the same syllable type, or certain acoustic components of a syllable, might sometimes be produced on one side of the syrinx and sometimes on the other, resulting in a low motor correlate when these syllables are cross correlated. The fact that in thrashers the right side tends to produce a higher fundamental frequency than the left (Suthers et al., 1994) must limit the extent to which the left and right sides are interchangeable, but there is considerable overlap in the frequency ranges of the two sides. Such a left–right reversal of otherwise basically similar motor patterns may account for the combination of a high spectrographic correlate and low motor correlate in the two syllables shown in Figure 7. The predominant high and low relatively constant frequency elements generated on the right and left sides, respectively, of the syrinx in syllable “a” are each produced on the opposite side in syllable “b.” If in syllable “b” the left side had begun to open about 80 ms earlier, the flow and pressure patterns for the two syllables would be almost left–right mirror images of each other.

Studies on the syringeal mechanisms of mimid song production also suggest other ways in which similar acoustic effects can be produced by different motor mechanisms that do not require a left–right reversal of the motor output. There is more than one motor mechanism by which rapid, cyclical AM can be generated, for example (Suthers et al., 1994). Also, in regard to the fundamental frequency, the ventral syringeal muscles have a major role in controlling the frequency of the ipsilateral sound, but other respiratory and syringeal muscles also contribute to frequency control in varying degrees (Goller and Suthers, 1995, 1996).

Comparison of the motor patterns used by different individuals singing shared syllables could provide further insight into the degree of motor flexibility that is possible for a given vocal output. The diversity of song in mimic thrushes makes shared syllables difficult to find and we did not identify any in our catbirds or thrashers. Circumstantial evidence that separate individuals may use different motor patterns to produce the same syllable

has been reported in sparrows. Notes shared by different swamp sparrows (*Melospiza georgiana*) are altered differently in different individuals after paralysis of one side of the syrinx by unilateral section of the tracheosyringeal branch of the hypoglossal nerve (Nowicki and Peters, 1991; Nowicki and Podos, 1993). Nowicki and his colleagues argue that this implies individual sparrows used different motor patterns, before nerve section, to produce the same note, although the possibility of individual differences in the ability to compensate for unilateral denervation cannot be excluded.

### Motor Stereotypy of Repeated Syllables

Our data clearly indicate that individual thrashers and catbirds use a stereotyped motor pattern each time they repeat a syllable. However, despite the obvious similarity of the motor action underlying repeated syllables, the patterns of airflow and pressure are seldom identical each time a syllable type is delivered. When, for example, syllable “b” (Fig. 6) was repeated after 94 intervening syllables, the motor output to muscles of the right side of the syrinx produced only 23 instead of 24 cyclical changes in syringeal resistance to airflow. Minor motor variation between successive repetitions of a syllable might occur for several reasons: the bird may be attempting, through auditory feedback, to exactly replicate the acoustic output by compensating for small differences in its physiological state (e.g., differences in the volume of air available for phonation in the respiratory system); slight variability might be “deliberately” introduced by the bird to create some acoustic variation or “microplasticity” within the framework of an otherwise stereotyped phonetic gesture; or motor variation may reflect imprecision or “noise” in the motor control and effector system.

The first alternative is unlikely because slight differences in airflow or pressure patterns are normally accompanied by small variations in the vocal output. We have no basis for evaluating the second and third alternatives but favor the third. The accuracy of motor control in song production has presumably evolved in response to the bird’s perceptual and cognitive needs. The motor (and resulting acoustic) noise within a given syllable type is probably unimportant providing it does not obscure the behaviorally important features of the syllable.

### Syllable Similarity and Diversity in Syntax of Song

The prominence of couplets in thrasher song suggests that they have a special behavioral significance, perhaps a special role in species recognition as suggested by the playback experiments of Boughy and Thompson (1976). Couplets may occur more frequently in our catbirds, which were hand-reared with thrashers, than in the wild catbird population. Their prominence to the listener depends not only on their acoustic similarity, described above, but equally on the dissimilarity of successive nonrepeated syllables that make up the rest of the song. When syllables are compared out of the sequence or syntax in which they were sung, the distinctive nature of couplets is less apparent (Fig. 1). Except for couplets, thrashers and catbirds seldom sing syllables that are acoustically very similar in immediate succession. Spectrogram correlation coefficients of neighboring syllables average about midway between those expected if the bird tried to maximize the difference between successive syllables (worst spectrographic match) and those predicted if the bird tried to minimize that difference (best spectrographic match) (Table 3). In song, where each syllable has two neighbors, there are fewer degrees of freedom in matching syllables than when only pairs of syllables are considered. For this reason alone, it is unlikely that a singing bird could achieve the best or worst correlates of matched pairs. Even so, it is apparent that the syntax of mimic thrush song is not determined by ordering syllables in the sequence that maximizes the acoustical differences between noncouplet neighbors, even though such a sequence would further enhance the acoustic prominence of couplets. There may be no advantage in decreasing the similarity of neighboring syllables beyond that necessary for reliably distinguishing between couplets and noncouplets. A more stringent requirement for acoustic dissimilarity may unnecessarily restrict the phonetic structure of neighboring syllables, reducing the variety of syllable types in the song. The phonetic diversity of neighboring syllables in mimids may be determined by a balance between the competing needs to enhance the perceptual salience of couplets, on the one hand, and to maintain phonetic versatility on the other. The fact that the motor correlate of neighboring syllables is not significantly different from that of syllables paired for their worst spectrographic match may be due to the limited sensitivity of our index of motor similarity or it may reflect physical con-

straints on the extent to which the syrinx can be reconfigured in the short time (averaging 170 ms; Suthers et al., 1994) between syllables.

### Couplets, Distant Repetitions, and Central Representation of Song

The vocal repertoire of these mimic thrushes has not been accurately determined but is estimated to contain between over 100 up to nearly 2000 different syllable types (e.g., Thompson and Jane, 1969; Kroodsma and Parker, 1977). How is such a large and seemingly plastic song repertoire represented in the brain? Two possibilities are that either the bird improvises the motor patterns required to produce the sounds as it sings them or that the pattern of muscle activation required to produce each syllable is stored as a central motor program that can be activated to produce the syllable, as hypothesized by Konishi (1985). Vu et al. (1994) obtained experimental evidence that zebra finch song is controlled by a central motor program in the telencephalic song control nucleus HVC.

Although our experiments do not directly address this question, they provide some pertinent circumstantial evidence. The immediate repetition of the same motor pattern, such as occurs in a couplet, might only require that the motor pattern used to produce the first syllable of the couplet is stored in short-term memory at least long enough to accurately reproduce the syllable a second time. If only short-term memory is involved, then the motor correlates of syllables that are repeated after singing many other syllables, the distant repetitions, should have lower XCR values and be less accurate copies of each other than are couplets. The fact that there is no degradation in the motor or spectrogram correlates of distant repetitions, compared to those of immediate repetitions, suggests that distantly repeated syllables are stored as motor programs in long-term memory. Couplets themselves are also often repeated from time to time throughout song, indicating that they too are stored in long-term memory.

The number of syllable types that are repeated increases with the amount of song analyzed, but it is not clear what proportion of its repertoire an adult mimid eventually repeats. It may be that a subset of syllable types is stored in long-term memory and the remainder are improvised while singing. If so, it would be very interesting to know what determines whether a particular syllable is stored in memory.

## Comparisons with Cowbirds

Motor stereotypy of different song types has recently been studied in cowbirds (Allan and Suthers, 1994) using a similar method of correlating spectrographs, syringeal airflow, and respiratory pressure. In contrast to mimic thrushes, adult male cowbirds learn a repertoire of about four to seven different song types during their first year that they sing for the rest of their life (King et al., 1980; West et al., 1983). In a given male, each song type, consisting of two or three introductory note clusters followed by a final whistle, has a characteristic, stereotyped acoustic or phonetic structure and is accompanied by highly stereotyped patterns of syringeal airflow and pressure. Repetitions of the same song type had mean spectrographic correlates of about 0.8 and motor correlates between about 0.95 and 1.0 (Allan and Suthers, 1994). It is especially interesting that the motor and spectrographic stereotypy of repeated syllables in catbirds and thrashers (Table 2, Fig. 4) is similar to that of song types in cowbirds. The vocal diversity in mimic thrushes is thus an additional attribute of their vocal ability, and does not simply reflect an inability to produce stereotyped vocalizations.

We thank H. Robert Lezotte for assistance in some of the experiments and Carolyn Pytte for comments on manuscript. Supported by NSF IBN 9411191 and PHS NS29467 to R.A.S. and training grant NSF DIR-9014276 from the Center for the Integrative Study of Animal Behavior, Indiana University to F.G.

## REFERENCES

- ALLAN, S. E. and SUTHERS, R. A. (1994). Lateralization and motor stereotypy of song production in the brown-headed cowbird. *J. Neurobiol.* **25**:1154–1166.
- BORROR, D. J. (1964). Songs of the thrushes (Turdidae), wrens (Troglodytidae), and mockingbirds (Mimidae) of eastern North America. *Ohio J. Sci.* **64**:195–207.
- BOUGHEY, M. J. and THOMPSON, N. S. (1976). Species specificity and individual variation in the songs of the brown thrasher (*Toxostoma rufum*) and catbird (*Dumetella carolinensis*). *Behaviour* **57**:64–90.
- BOUGHEY, M. J. and THOMPSON, N. S. (1981). Song variety in the brown thrasher (*Toxostoma rufum*). *Z. Tierpsychol.* **56**:47–58.
- FLETCHER, L. E. and SMITH, D. G. (1978). Some parameters of song important in conspecific recognition by gray catbirds. *Auk* **95**:338–347.
- GOLLER, F. and SUTHERS, R. A. (1995). Implications for lateralization of bird song from unilateral gating of bilateral motor patterns. *Nature* **373**:63–66.
- GOLLER, F. and SUTHERS, R. A. (1996). The role of syringeal muscles in gating airflow and sound production in singing brown thrashers. *J. Neurophysiol.* **75**:867–876.
- HARTLEY, R. S. (1990). Expiratory muscle activity during song production in the canary. *Respir. Physiol.* **81**:177–188.
- KING, A., WEST, M., and EASTZER, D. (1980). Song structure and song development as potential contributors to reproductive isolation in cowbirds (*Molothrus ater*). *J. Comp. Physiol. Psychol.* **94**:1028–1039.
- KONISHI, M. (1985). Birdsong: from behavior to neuron. *Annu. Rev. Neurosci.* **8**:125–170.
- KROODSMA, D. E. and BAYLIS, J. R. (1982). A world survey of evidence for vocal learning in birds [Appendix]. In: *Acoustic Communication in Birds*, Vol. 2. D. E. Kroodsma and E. H. Miller, Eds. Academic Press, New York, pp. 311–337.
- KROODSMA, D. E. and PARKER, L. D. (1977). Vocal virtuosity in the brown thrasher. *Auk* **94**:783–785.
- LIBERMAN, A. M., COOPER, F., SHANKWEILER, D., and STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychol. Rev.* **74**:431–461.
- LOHR, B. S. (1989). The organization of song elements in the gray catbird. M.S. thesis, University of Wisconsin–Milwaukee.
- MARLER, P. (1991). Song-learning behavior: the interface with neuroethology. *Trends Neurosci.* **14**:199–206.
- NOTTEBOHM, F. (1991). Reassessing the mechanisms and origins of vocal learning in birds. *Trends Neurosci.* **14**:206–211.
- NOTTEBOHM, F., STOKES, T. M., and LEONARD, C. M. (1976). Central control of song in the canary (*Serinus canarius*). *J. Comp. Neurol.* **165**:457–486.
- NOWICKI, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* **325**:53–55.
- NOWICKI, S. and PETERS, S. (1991). Motor bases of behavioral stereotypy in swamp sparrow song note production. *Soc. Neurosci. Abst.* **17**:1051.
- NOWICKI, S. and PODOS, J. (1993). Complexity, coupling and contingency in the production of birdsong. In: *Perspectives in Ethology, Vol. 10: Behavior and Evolution*. P. P. G. Bateson, P. Klopfer, and N. Thompson, Eds. Plenum Press, New York, pp. 159–186.
- SIMPSON, H. B. and VICARIO, D. S. (1990). Brain pathways for learned and unlearned vocalizations differ for zebra finches. *J. Neurosci.* **10**:1541–1556.
- SUTHERS, R. A. (1990). Contributions to birdsong from the left and right sides of the intact syrinx. *Nature* **347**:473–477.
- SUTHERS, R. A., GOLLER, F., and HARTLEY, R. S.

- (1994). Motor dynamics of song production by mimic thrushes. *J. Neurobiol.* **25**:917–936.
- THOMPSON, W. L. and JANE, P. L. (1969). An analysis of catbird song. *Jack-Pine Warbler* **47**:115–125.
- VICARIO, D. S. (1991). Contributions of syringeal muscles to respiration and vocalization in the zebra finch. *J. Neurobiol.* **22**:63–73.
- VU, E. T., MAZUREK, M. E., and KUO, Y.-C. (1994). Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14**:6924–6934.
- WEST, M., KING, A., and HARROCKS, T. J. (1983). Cultural transmission of cowbird song (*Molothrus ater*): measuring its development and outcome. *J. Comp. Psychol.* **97**:327–337.
- WESTNEAT, M. W., LONG, J. H., Jr., HOESE, W., and NOWICKI, S. (1993). Kinematics of birdsong: functional correlation of cranial movements and acoustic features in sparrows. *J. Exp. Biol.* **182**:147–171.
- WILLIAMS, H. and NOTTEBOHM, F. (1985). Auditory responses in avian vocal motor neurons: a motor theory for song perception in birds. *Science* **229**:279–282.
- WILLIAMS, H., CRANE, L. A., HALE, T. K., ESPOSITO, M. A., and NOTTEBOHM, F. (1992). Right-sided dominance for song control in the zebra finch. *J. Neurobiol.* **23**:1006–1020.