Speech perception by budgerigars (*Melopsittacus undulatus*): Spoken vowels

ROBERT J. DOOLING and SUSAN D. BROWN
University of Maryland, College Park, Maryland

Discrimination of natural, sustained vowels was studied in 5 budgerigars. The birds were trained using operant conditioning procedures on a *same-different* task, which was structured so that response latencies would provide a measure of stimulus similarity. These response latencies were used to construct similarity matrices, which were then analyzed by multidimensional scaling (MDS) procedures. MDS produced spatial maps of these speech sounds where perceptual similarity was represented by spatial proximity. The results of the three experiments suggest that budgerigars perceive natural, spoken vowels according to phonetic categories, find the acoustic differences among different talkers less salient than the acoustic differences among vowel categories, and use formant frequencies in making these complex discriminations.

A number of recent psychophysical experiments have shown that budgerigars (*Melopsittacus undulatus*) have an unusual degree of spectral resolving power in the frequency region of 2–4 kHz—the frequency region where most of the energy in budgerigar vocalizations falls. The fact that frequency resolving power in budgerigars is roughly the inverse of the spectral energy distribution in their complex vocalizations has invited speculation regarding their specialized processing of species-specific vocal signals (Dooling, 1986). Spectral cues in the region of 2–4 kHz do underlie the perceptual categories for species-specific vocal signals in this species (Brown, Dooling, & O'Grady, 1988; Dooling, Park, Brown, Okanoya, & Soli, 1987). It would be of interest to know how budgerigars perceive complex acoustic signals that fall outside the region of enhanced spectral resolving power. Evidence of emergent perceptual categories for complex signals that fall outside of this region would argue for a more general—as opposed to special—auditory competence as the basis for complex sound and vocalization perception in this species.

Such a test would be interesting for yet another reason. Evidence from critical ratios, critical bands, and psychophysical tuning curves indicates that frequency resolving power improves in budgerigars from 250 Hz to about 3 kHz (Dooling & Saunders, 1975; Okanoya & Dooling, 1987; Saunders, Rintelmann, & Bock, 1979). This pattern of decreasing critical bandwidths with increasing frequency stands in marked contrast to the pattern found in humans and most other mammals (e.g., cats, chinchillas, and monkeys) and birds. In these other vertebrates, critical ratios and critical bandwidths increase

with frequency at a rate of about 3 dB/octave (see, e.g., Dooling, 1980; Fay, 1988; Okanoya & Dooling, 1987). Such differences in critical band functions suggest that budgerigars should not discriminate among human speech sounds very well and that humans should not discriminate among budgerigar vocalizations very well.

One approach to this problem is to test budgerigars directly on human speech sounds. Surprisingly, a recent study has shown that budgerigars can perceive the voicing distinction among the alveolar, bilabial, and velar plosive consonants of human speech (Dooling, Okanoya, & Brown, 1989). However, the acoustic basis for voice onset time perception in budgerigars probably involves temporal rather than spectral cues. Given the unusual shape of the budgerigar critical band function, a more interesting comparative test would be to use speech sounds that differ only in spectral features, such as vowels.

There have been several recent studies on the perception of vowels by animals. Dewson (1964; Dewson, Pribram, & Lynch, 1969) was able to train cats and monkeys to discriminate the vowel /a/ from the vowel /i/. Burdick and Miller (1975) extended these findings in an important way by demonstrating that chinchillas could be trained to classify the vowels /a/ and /i/ reliably, in spite of variations in talker, pitch contour, and intensity. Recent studies show that blackbirds, pigeons (Hienz, Sachs, & Sinnott, 1981), and several species of nonhuman primates can discriminate among different synthetic vowels (Hienz & Brady, 1988; Sinnott, 1989).

The similarities in vowel perception among mammals are perhaps not too surprising. Humans, cats, chinchillas, and nonhuman primates all show parallel critical band functions, which indicate that frequency is organized in a logarithmic fashion along the basilar membrane (Fay, 1988). The similar results from pigeons and blackbirds are somewhat more intriguing, but limited, because the birds were tested only with single exemplars of synthetic vowels.

This work was supported by NIH Grants NS19006 and HD00512 to R. J. Dooling. We thank K. Okanoya and T. J. Park for comments on an earlier draft and D. Reidell for care of the birds. Correspondence should be addressed to Robert J. Dooling, Psychology Department, University of Maryland, College Park, MD 20742-4411.

To the extent that the perception of vowels involves the relations among formant frequencies, one would predict similar results from animals with similar psychoacoustic abilities and whose cochleas are scale models of one another (Greenwood, 1961). The initial filtering processes in the cochlea are deemed so important that most psychoacoustic models of vowel perception in humans incorporate a critical bandwidth filter in the first stage of auditory analysis (Espinoza-Varas, 1987; Syrdal & Gopal, 1986; Zwicker, Terhardt, & Paulus, 1979).

Finally, much of the past work on speech perception among animals (including the work on birds) has relied on some version of a classification task resulting in comparisons of learning rates for different classes of stimuli or responses to intermediate and extreme forms of stimuli. By contrast, there are other procedures that rely on response latencies as a measure of stimulus similarity (Hienz et al., 1981; Sinnott, 1989). The present procedure involves a discrimination task structured so that shorter response latencies correlate well with greater perceptual dissimilarity—a relation that can be validated through the use of simple pure tones (Dooling, Brown, Park, Okanoya, & Soli, 1987) and complex sounds (Dooling, Park, et al., 1987). These procedures produce similarity measures suitable for analysis with multidimensional scaling (MDS) techniques.

MDS places points in multidimensional space such that interstimulus distances correspond as closely as possible to perceived similarities among stimuli (Borg & Lingoes, 1987; Shepard, 1980). Stimuli with similar perceptual properties are near each other in multidimensional space, whereas stimuli with different perceptual properties are far apart. MDS has proved to be a useful tool for investigating vowel perception in humans (Fox, 1983, 1985; Murry & Singh, 1980; Singh & Murry, 1978). In the following experiments, budgerigars were tested on natural spoken vowels from five different phonetic categories produced by different talkers.

EXPERIMENT 1

Method

Subjects. The subjects in this experiment were 2 adult female and 3 adult male budgerigars (*Melopsittacus undulatus*) housed in aviaries at the University of Maryland. All birds were well trained on the auditory *same-different* task and had participated in other psychoacoustic studies involving both simple and complex sounds.

Stimuli. The stimuli in Experiment 1 were the natural, spoken vowels from four different talkers. The stimuli were edited from sustained, steady-state vowels so that they were 200 msec in length with 5-msec rise and fall times. The stimuli were presented at a peak level of 72 dB SPL. Sound pressure level was measured by placing the microphone of a sound level meter just in front of the response panel in the location normally occupied by the bird's head during testing. The four vowels were /i/, /ε/, /a/, and /u/ as in the words *reed*, *red*, *rod*, and *rude*. Vowels are typically described by the perceptually salient frequencies of the first and second formants (Peterson & Barney, 1952). The formant frequencies F0, F1, F2, and F3 for these 16 stimuli were measured from the total power spectrum with the Kay Elemetrics Real Time DSP Model 5500 Sonograph.

Apparatus. The apparatus for training and testing the birds has been described earlier (Dooling, Brown, et al., 1987; Park, Okanoya, & Dooling, 1985). The birds were tested in wire cages mounted in sound-attenuated chambers. One wall of the wire test cage was modified by the addition of a custom-built response panel constructed of three sensitive microswitches with light-emitting diodes (LEDs) attached. A bird could trip the microswitch by pecking at the LED. The center microswitch and LED served as an observation key, and the left microswitch and LED served as a report key. Experimental events were controlled by an IBM AT microcomputer. All acoustic stimuli were stored in digital form with a 12-bit quantization scheme, output at a sampling rate of 20 kHz through D/A converters, and low-pass filtered at 10 kHz to prevent aliasing.

Procedure. Since all of the birds had already been trained to respond in an auditory *same-different* task, no additional training was necessary with the speech sounds used in these experiments. A trial began with the illumination of the observation LED. A response on the observation key resulted in the presentation of two stimuli separated by 300 msec. A response on the report key within 2 sec following the presentation of two different stimuli (measured from the end of the second stimulus) was rewarded with a 2-sec access to grain. A response on the report key within 2 sec following the presentation of two identical stimuli was punished with a 20-sec timeout period, during which the lights in the experimental chamber were extinguished. Each trial was followed by a 1-sec inter-trial interval and then by a new trial sequence, which started with the illumination of the observation LED. A new trial sequence was also initiated if the bird failed to respond during the 2-sec report interval. As in previous experiments, the bird's response latency on *different* trials was taken as an index of stimulus similarity and served as the dependent variable (Dooling, Brown, et al., 1987).

Since each trial consisted of the presentation of two stimuli, the data-collection procedures are best described in terms of cells in a matrix in which each stimulus is paired once with every other stimulus in the set (*different* trials) and the diagonal represents trials in which each stimulus is paired with itself (*same* trials). In all three experiments, the stimulus sets consisted of 10–16 different stimuli. A bird was tested in daily sessions until all possible pairwise combinations of the different stimuli were presented once, with the added constraint that there be equal numbers of *same* and *different* trials to fill a complete matrix of response latencies. This constraint was met by repeating trials involving the pairing of each stimulus with itself. The total number of trials required to satisfy this criterion was $2N(N-1)$, where N is the number of speech tokens in the test set. Once the bird completed a matrix, the data were stored on disk, and the bird was tested again on all possible pairwise combinations of the different stimuli in the same manner. The order of stimuli presented was randomized across trials. Each bird was tested until at least five complete matrices of response latencies were available for analysis.

Data reduction. Each cell in the response latency matrix from an individual bird contained the results from a single trial involving one pair of stimuli. At the end of testing, the latency values in each of the five matrices for each bird were subjected to a log transformation to accommodate the positively skewed distributions of response latencies, and a single matrix representing the average of the five matrices was computed (Dooling, Brown, et al., 1987). This average matrix was then folded about its main diagonal by averaging the corresponding cell entries in the lower and upper half matrices (i.e., $A-B$ with $B-A$ comparisons) and discarding the values in the diagonals (i.e., latencies from *same* trials). This resulted in a half matrix of response latencies required by the MDS algorithm.

Results and Discussion

The results from three-dimensional analysis by the MDS program SINDSCAL of the perceptual similarity among

Figure 1. (A) Two-dimensional spatial plot (I vs. II) of the four vowels /a/, /i/, /u/, and /ε/ produced by four male talkers. (B) Two-dimensional spatial plot (I vs. III) of these same vowel sounds.

the 16 vowel tokens for the 5 budgerigars are shown in Figures 1A and 1B. The subject weights for this three-dimensional solution are given in Table 1. The total variance in response latency accounted for by the MDS solution was 54.3%, with the first, second, and third dimensions accounting for 26.8%, 13.9%, and 13.6%, respectively.

The stimulus clusters in the spatial plot were confirmed with a hierarchical cluster analysis (Aldenderfer & Blashfield, 1984). The 16 vowel tokens are separated in perceptual space into four groups by phonetic category. Another way of demonstrating this point is to perform a *t* test comparing the log ($\times 100$) of the response latencies to all possible pairs of stimuli from within a phonetic category ($M = 146.4$ msec, $SD = 11.8$ msec) to all possible pairs of stimuli drawn from between phonetic categories ($M = 128.7$ msec, $SD = 10.0$ msec). A *t* test showed this difference to be significant [$t(118) = 7.43$, $p < .001$]. A similar test comparing the response laten-

cies to all possible pairs of stimuli drawn from within speakers ($M = 132.7$ msec, $SD = 12.8$ msec) to the response latencies to all possible pairs of stimuli drawn from between speakers ($M = 130.5$ msec, $SD = 11.5$ msec) did not reveal a significant difference [$t(118) = 0.74$, $p > .05$]. These results clearly show that budgerigars find the acoustic differences among vowel categories more salient than the acoustic differences among vowels produced by different talkers.

The relation between the three perceptual dimensions and the acoustic characteristics of speech was examined by correlating the vowel coordinates on each perceptual dimension with the following acoustic measures: F0, F1, F2, F3, F1/F0, F2/F1, and F3/F2. These simple correlations are given in Table 2.

These results show that formant frequency measures are correlated with the location of vowels in the budgerigars' perceptual space. In particular, F1, F2/F1, and F3/F2 are significantly correlated with stimulus coordinates on the first dimension—the dimension that separates /a/ from /i/. On the second dimension, which separates /ε/ from /u/, F1, F2, F3, F1/F0, F2/F1, and F3/F2 are all correlated with stimulus coordinates. These results suggest that budgerigars use formant frequencies in making phonetically relevant discriminations among natural vowels. But, other information, such as the gender and age of the talker, is typically coded in natural speech. Human listeners maintain phonetically relevant discrimination in spite of acoustic variability introduced by male and female talkers. Experiment 2 was designed to explore this phenomenon with budgerigars.

EXPERIMENT 2

Method

Subjects. The 5 budgerigars from Experiment 1 and 1 additional female budgerigar were used in Experiment 2.

Stimuli. The stimuli were the natural, sustained vowels /a/ and /i/ produced by four male and four female talkers. Each of the 16

Table 1
Subject Weights for Birds Tested on /a/, /ε/, /i/, and /u/

Bird	Dimension 1	Dimension 2	Dimension 3
P21	.528	.313	.324
P74	.533	.282	.343
P8G	.409	.449	.338
P34	.508	.484	.307
PRW	.544	.237	.432

Table 2
Correlation between Perceptual Dimensions and Formant Frequencies for the Four Vowels /i/, /ε/, /a/, and /u/

Formant	Dimension 1	Dimension 2	Dimension 3
F0	.426	-.337	-.470
F1	.752†	.582*	-.038
F2	-.376	-.901†	-.296
F3	.359	-.620*	-.147
F1/F0	.299	.742†	.225
F2/F1	-.621*	-.694†	.001
F3/F2	.534*	.775†	.173

* $p < .05$. † $p < .01$.

Figure 2. (A) Two-dimensional spatial plot (I vs. II) of the two vowels /a/ and /i/ produced by four male and four female talkers. (B) Two-dimensional spatial plot (I vs. III) of these same vowel sounds.

stimuli was recorded and prepared in the same manner as in Experiment 1. The formant frequencies of these stimuli were also measured in the same way as those in Experiment 1.

Apparatus and Procedure. The apparatus was identical to that in Experiment 1. The testing procedures were also identical to those used in Experiment 1.

Results and Discussion

The results from a three-dimensional SINDSCAL analysis of these stimuli accounted for a total of 57.3% of the variance in response latency with the first, second, and third dimensions accounting for 27.6%, 17.0%, and 12.9% of this total variance, respectively. These results are shown as two two-dimensional spatial plots in Figures 2A and 2B. The subject weights for this SINDSCAL solution are given in Table 3. These 16 stimuli are separated by phonetic category along the first dimension, with /i/ vowel tokens on the left and /a/ vowel tokens on the right. The third dimension separates these stimuli by

gender of the talker, with vowels produced by females near the top and vowels produced by males near the bottom.

The stimulus groupings seen in this spatial plot were confirmed by a cluster analysis. As in Experiment 1, these stimulus groupings can also be confirmed by a simple analysis of response latencies. The response latencies for pairs of stimuli drawn from within-vowel categories ($M = 141.5$ msec, $SD = 12.2$ msec) are significantly longer [$t(118) = 8.39$, $p < .001$] than the response latencies for pairs of stimuli drawn from between-vowel categories ($M = 126.1$ msec, $SD = 7.6$ msec). The response latencies for pairs of talkers of the same sex ($M = 137.3$ msec, $SD = 14.6$ msec) are significantly longer [$t(118) = 3.37$, $p < .001$] than the response latencies for pairs of stimuli spoken by talkers of different sexes ($M = 129.9$ msec, $SD = 9.4$ msec).

As in Experiment 1, the relation between the three perceptual dimensions and the acoustic characteristics of speech was examined by correlating the vowel coordinates on each perceptual dimension with the same acoustic measures. These simple correlations are given in Table 4.

These results demonstrate again that budgerigars probably rely on formant information in discriminating between vowels from the two phonetic categories. F1, F2, F1/F0, F2/F1, and F3/F2 are all correlated with stimulus coordinates on the first dimension—the dimension that separates /a/ from /i/ in the budgerigar's perceptual space. The birds are also sensitive to the acoustic differences among vowels produced by males and females, and the acoustic cues underlying this discrimination are probably related to F2 and F3 (the second dimension) and F0 (the third dimension).

Because these two vowels produced by males and females are so spectrally distinct, it is perhaps not too surprising that budgerigars can discriminate among vowel and talker categories. The task can be made somewhat more complicated by adding vowels produced by children

Table 3
Subject Weights for Birds Tested on Male and Female /a/ and /i/

Bird	Dimension 1	Dimension 2	Dimension 3
P21	.369	.390	.381
P2Y	.453	.467	.379
P74	.527	.317	.406
P8G	.604	.364	.293
PRW	.547	.489	.289

Table 4
Correlation between Perceptual Dimensions and Formant Frequencies for the Vowels /a/ and /i/ Spoken by Males and Females

Formant	Dimension 1	Dimension 2	Dimension 3
F0	-.159	-.425	-.859†
F1	.811†	.203	-.407
F2	-.774†	-.536*	.094
F3	-.399	-.568*	-.377
F1/F0	.821†	.326	.205
F2/F1	-.866†	-.192	.353
F3/F2	.756†	.315	.343

* $p < .05$. † $p < .01$.

Figure 3. Two-dimensional spatial plot of the similarity among these vowel sounds for four budgerigars.

to the stimulus set. In the next experiment, budgerigars were tested on two vowel categories with tokens produced by men, women, and children.

EXPERIMENT 3

Method

Subjects. Two female and 2 male birds from Experiment 2 were used in Experiment 3.

Stimuli. The stimuli were the natural, sustained vowels /a/ and /i/ produced by two adult male, two adult female, and four child talkers of ages 5–8 years (two male, two female). Each of the 16 stimuli was recorded, prepared, and analyzed in the same manner as in Experiments 1 and 2.

Apparatus and Procedure. The apparatus and the testing procedures were identical to those in Experiment 1.

Results and Discussion

The two-dimensional spatial representation of these stimuli generated by SINDSCAL accounted for a total of 48.7% of the variance in response latency, with the first and second dimensions accounting for 30.3% and 18.4%, respectively. A two-dimensional plot is shown in Figure 3. The subject weights for the spatial solution are given in Table 5.

The results of a hierarchical cluster analysis of the average data from the four birds confirmed the stimulus groupings shown in the spatial plot. In addition, the response latencies to pairs of stimuli drawn from within-vowel categories ($M = 151.4$ msec, $SD = 10.3$ msec) were significantly longer [$t(118) = 7.31$, $p < .001$] than the response latencies to pairs of stimuli drawn from between-vowel categories ($M = 140.8$ msec, $SD = 4.9$ msec). The response latencies to pairs of sounds drawn from within adult or child talkers ($M = 148.3$ msec, $SD = 10.6$ msec) were also significantly longer [$t(118) = 2.90$, $p < .01$] than the response latencies to pairs of sounds drawn from between child and adult talkers ($M = 143.4$ msec, $SD = 7.8$ msec).

We conclude that budgerigars hear the similarities among vowels drawn from the same phonetic category produced by men, women, and children. The correlations between stimulus coordinates in perceptual space and the acoustic measures described above are shown in Table 6. As in the previous experiments, these results suggest that budgerigars use information in formant frequencies to discriminate among these vowels. Formants F1, F2, F3, F1/F0, F2/F1, and F3/F2 are all correlated with stimulus coordinates on the first dimension—the dimension that separates /a/ from /i/. F0, F1, and F3 are correlated with stimulus coordinates along the second dimension—the dimension that separates male, female, and child talkers.

CONCLUSIONS

In this study, we sought to provide comparative data from birds on the perception of natural, spoken vowels. We trained budgerigars on an auditory *same-different* discrimination task involving these speech sounds and then analyzed their response latencies with multidimensional scaling and cluster analysis procedures. The correlation between stimulus coordinates in multidimensional perceptual space and various formant frequency measures provides an indication of what formant information budgerigars may be using in perceiving these vowels. The complete test of this hypothesis, of course, would require sets of synthetic vowels where stimuli could be altered in specific ways to move them to predicted locations in perceptual space. Nevertheless, we feel that the present experiments provide strong evidence that formant frequencies in natural speech are used by budgerigars in discriminating among categories of vowels and talkers.

In Experiment 1, in which the four different vowels were widely separated in two-dimensional space, the perceptual coordinates of these vowels were correlated with formant frequencies—especially F1 and F2. In Experi-

Table 5
Subject Weights for Birds Tested on /a/
versus /i/ Spoken by Adults and Children

Bird	Dimension 1	Dimension 2
P21	.479	.430
P2Y	.669	.377
P74	.535	.471
P8G	.466	.450

Table 6
Correlation between Perceptual Dimensions and Formant
Frequencies for the Vowels /a/ and /i/ Spoken by Adults and Children

Formant	Dimension 1	Dimension 2
F0	-.214	-.739†
F1	.748†	-.612*
F2	-.942†	.076
F3	-.478	-.493*
F1/F0	.741†	-.036
F2/F1	-.763†	.515*
F3/F2	.811†	-.331

* $p < .05$. † $p < .01$.

ments 2 and 3, involving the contrast /a/ versus /i/, the first dimension in perceptual space separated these vowels by phonetic category. Distances along this dimension accounted for the greatest amount of variance in response latencies. On the average, the subject weights for each bird were also highest on the first dimension. In both of these experiments, the perceptual coordinates of the vowels along the first dimension were correlated with formant frequencies—again predominantly F1 and F2. For several reasons, these findings are interesting for theories of vocal perception in both budgerigars and humans.

First, the close correspondence between vocalization spectra and absolute auditory sensitivity and spectral resolving power strongly suggests that the budgerigars may be specialized for the perception of vocal signals (Brown et al., 1988; Dooling, 1986; Dooling & Saunders, 1975). However, although this may be true, the present results demonstrate that budgerigars are also capable of perceiving complex speech sounds in a phonetically appropriate way—sounds that fall outside the narrow range of enhanced spectral resolving power. These results argue against a specialized, peripheral auditory system in budgerigars dedicated to the perception of species-specific vocal signals, and for a more flexible, sophisticated auditory perceptual system.

Second, it has long been known that spectral cues are critically involved in the human perception of vowels (Delattre, Liberman, Cooper, & Gerstman, 1952; Peterson & Barney, 1952). The formant frequencies of the natural vowels used in these experiments—particularly F0 and F1—fall well outside the range of best auditory sensitivity and frequency resolving power for budgerigars. Nevertheless, budgerigars do discriminate among these natural speech sounds, and they appear to rely on the same information that humans do in discriminating among vowels (i.e., the relation between F1 and F2 formant frequencies) and among talkers (i.e., F0 and F3 formant frequencies) (Murry & Singh, 1980; Singh & Murry, 1978).

Perhaps more importantly, these experiments also demonstrate that budgerigars hear the similarities among vowels of the same phonetic category in spite of acoustic variations in vowels produced by different talkers. Although exactly how humans accomplish this task so well remains somewhat a mystery, most current models of vowel perception and categorization incorporate an early, peripheral stage of spectral filtering based on the mammalian critical band function (Espinoza-Varas, 1987). It is on just this and other, closely related kinds of psychoacoustic measures that the human and budgerigar auditory systems are clearly different (Dooling & Saunders, 1975; Dooling & Searcy, 1985; Saunders et al., 1979). For this reason, the budgerigar, unlike other animal models, may provide a unique test of the role of critical band filtering mechanisms in vowel perception by humans.

It is also important to point out that in the procedures used in these experiments, the birds were not trained to a particular class of sounds. Thus, the tendency of bud-

gerigars to show stimulus clusters, groupings, or categories cannot have been due to asymmetries in reinforcement contingencies or training conditions. Furthermore, while multidimensional scaling and cluster analysis techniques provide useful tools for visualizing these stimulus clusters, the “proof” that a grouping or cluster exists in these experiments is provided by the results of *t* tests comparing latencies within and between stimulus categories.

Finally, it is well known that budgerigars as well as other psittacine birds are extremely versatile vocal learners and can be readily trained to mimic a variety of complex sounds, including speech. It goes without saying that to accomplish this task, budgerigars must not only have a remarkably flexible vocal apparatus but also must be able to hear the relevant details of speech. The present results concerning vowels and those of a previous study showing that budgerigars discriminate among synthetic VOT stimuli in a phonetically appropriate way (Dooling et al., 1989) provide more direct evidence. It will be interesting to probe the limits of this capability to determine the degree to which social, motor, and perceptual constraints are involved in the learning of complex sounds, such as speech, which fall outside the range of best hearing in this species.

REFERENCES

- ALDENDERFER, M. S., & BLASHFIELD, R. K. (1984). *Cluster analysis*. Newbury Park, CA: Sage.
- BORG, I., & LINGOES, J. (1987). Multidimensional similarity structure analysis. New York: Springer-Verlag.
- BROWN, S. D., DOOLING, R. J., & O'GRADY, K. (1988). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): III. Contact calls. *Journal of Comparative Psychology*, **102**, 236-247.
- BURDICK, C. K., & MILLER, J. D. (1975). Speech perception by the chinchilla: Discrimination of sustained /a/ and /i/. *Journal of the Acoustical Society of America*, **58**, 415-427.
- DELATTRE, P., LIBERMAN, A. M., COOPER, F. S., & GERSTMAN, L. J. (1952). An experimental study of the acoustic determinants of vowel color: Observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word*, **8**, 195-210.
- DEWSON, J. H. (1964). Speech sound discrimination by cats. *Science*, **144**, 555-556.
- DEWSON, J. H., PRIBRAM, K. H., & LYNCH, J. C. (1969). Effects of ablations and temporal cortex upon speech sound discrimination in the monkey. *Experimental Neurology*, **24**, 579-591.
- DOOLING, R. J. (1980). Behavior and psychophysics of hearing in birds. In A. N. Popper & R. R. Fay (Eds.), *Comparative studies of hearing in vertebrates* (pp. 261-288). New York: Springer-Verlag.
- DOOLING, R. J. (1986). Perception of vocal signals by the budgerigar (*Melopsittacus undulatus*). *Experimental Biology*, **45**, 195-218.
- DOOLING, R. J., BROWN, S. D., PARK, T. J., OKANOYA, K., & SOLI, S. D. (1987). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): I. Pure tones. *Journal of Comparative Psychology*, **101**, 139-149.
- DOOLING, R. J., OKANOYA, K., & BROWN, S. D. (1989). Speech perception by budgerigars (*Melopsittacus undulatus*): The voiced-voiceless distinction. *Perception & Psychophysics*, **46**, 65-71.
- DOOLING, R. J., PARK, T. J., BROWN, S. D., OKANOYA, K., & SOLI, S. D. (1987). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals. *Journal of Comparative Psychology*, **101**, 367-381.
- DOOLING, R. J., & SAUNDERS, J. C. (1975). Hearing and vocalizations

- in the parakeet (*Melopsittacus undulatus*): Absolute thresholds, critical ratios, frequency difference limens, and vocalizations. *Journal of Comparative & Physiological Psychology*, **88**, 1-20.
- DOOLING, R. J., & SEARCY, M. H. (1985). Nonsimultaneous auditory masking in the budgerigar (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **99**, 226-230.
- ESPINOZA-VARAS, B. (1987). Involvement of the critical band in identification, perceived distance, and discrimination of vowels. In M. E. H. Schouten (Ed.), *The psychophysics of speech perception*. Dordrecht: Martinus Nijhoff.
- FAY, R. R. (1988). *Hearing in vertebrates: A psychophysics databook*. Winnetka, IL: Hill-Fay.
- FOX, R. A. (1983). Perceptual structure among monothongs and diphthongs in English. *Language & Speech*, **26**, 21-61.
- FOX, R. A. (1985). Multidimensional scaling and perceptual features: Evidence of stimulus processing or memory prototypes. *Journal of Phonetics*, **13**, 205-217.
- GREENWOOD, D. D. (1961). Critical bandwidth and frequency coordinates of the basilar membrane. *Journal of the Acoustical Society of America*, **33**, 1344-1356.
- HIENZ, R. D., & BRADY, J. V. (1988). The acquisition of vowel discrimination by nonhuman primates. *Journal of the Acoustical Society of America*, **84**, 186-194.
- HIENZ, R. D., SACHS, M. B., & SINNOTT, J. M. (1981). Discrimination of steady-state vowels by blackbirds and pigeons. *Journal of the Acoustical Society of America*, **70**, 699-706.
- MURRY, T., & SINGH, S. (1980). Multidimensional analysis of male and female voices. *Journal of the Acoustical Society of America*, **68**, 1294-1300.
- OKANOYA, K., & DOOLING, R. J. (1987). Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds. *Journal of Comparative Psychology*, **101**, 7-15.
- PARK, T. J., OKANOYA, K., & DOOLING, R. J. (1985). Operant conditioning of small birds for acoustic discrimination. *Journal of Ethology (Japan)*, **3**, 5-9.
- PETERSON, G. E., & BARNEY, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, **24**, 175-184.
- SAUNDERS, J. C., RINTELMANN, W. F., & BOCK, G. R. (1979). Frequency selectivity in bird and man: A comparison among critical ratios, critical bands, and psychophysical tuning curves. *Hearing Research*, **1**, 303-323.
- SHEPARD, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, **210**, 390-398.
- SINGH, S., & MURRY, T. (1978). Multidimensional classification of normal voice qualities. *Journal of the Acoustical Society of America*, **64**, 81-87.
- SINNOTT, J. M. (1989). Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus*, *Macaca*) and humans. *Journal of the Acoustical Society of America*, **86**, 557-565.
- SYRDAL, A. K., & GOPAL, H. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, **79**, 1086-1100.
- ZWICKER, E., TERHARDT, E., & PAULUS, E. (1979). Automatic speech recognition using psychoacoustic models. *Journal of the Acoustical Society of America*, **65**, 487-498.

(Manuscript received May 19, 1989;
revision accepted for publication December 20, 1989.)