Speech perception by budgerigars (*Melopsittacus undulatus*): The voiced-voiceless distinction

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

Discrimination of synthetic speech sounds from the bilabial, alveolar, and velar voice onset time (VOT) series was studied in 5 budgerigars. The birds were trained, using operant conditioning procedures, to detect changes in a repeating background of sound consisting of a synthetic speech token. Response latencies for detection were measured and were used to construct similarity matrices. Multidimensional scaling procedures were then used to produce spatial maps of these speech sounds, in which perceptual similarity was represented by spatial proximity. The results of these experiments suggest that budgerigars discriminate among synthetic speech sounds from these three VOT continua, especially between those from the bilabial and alveolar series, in a categorical fashion.

The study of speech perception by animals provides comparative data that is important for understanding human speech perception. Animal models permit testing of the generality of speech perception results obtained from humans. To date, there is an impressive accumulation of data on the perception of synthetic voice onset time (VOT) stimuli by humans and other mammals (e.g., Kuhl, 1979, 1986, 1987). The animal work, in general, aims to distinguish speech categorization effects that can be attributed to general auditory-processing mechanisms from those due to higher order, more integrative levels of processing (Kuhl, 1979, 1986, 1987).

Unfortunately, the comparative contribution to understanding the perception of VOT in speech has largely been restricted to mammals. In short, results show that mammals with auditory perceptual capabilities roughly similar to those of humans also show heightened discriminability in the short voicing-lag region of the VOT continuum (Kuhl & Miller, 1975, 1978; Kuhl & Padden, 1982, 1983; Waters & Wilson, 1976). One conclusion drawn from these results is that VOT categorization effects are due largely to general auditory processes rather than "phonetic" processes that require speech-specific mechanisms unique to humans.

Another conclusion, an extension of the first, is that general auditory processes characteristic of the mammalian auditory system guide the evolution and acquisition of a speech-sound repertoire in humans (Kuhl, 1979, 1981, 1986, 1987). The basis for this second conclusion is that the human (mammalian) auditory system appears to be well matched to the phonemically relevant acoustic

changes that are characteristic of speech (Kuhl, 1979, 1981, 1986). However, since nearly all work on speech perception in animals has been done with mammals, it is not clear that exclusively mammalian psychoacoustic capabilities are responsible for this phenomenon.

To our knowledge, there have been only three attempts to study speech perception by nonmammals, and all three of these studies have been done with birds. One study showed that redwing blackbirds, brown-headed cowbirds, and pigeons can be trained to discriminate among the synthetic, steady-state vowels /ε/, /æ/, /a/, and /ɔ/ (Hienz, Sachs, & Sinnott, 1981). In a study using natural consonant-vowel-consonant syllables, Japanese quail discriminated among syllable initial /b/, /d/, and /g/ across different vowel contexts. These results showed that quail are capable of appropriate phonetic classification in spite of a lack of acoustic invariance (Kluender, Diehl, & Killeen, 1987). On the whole, these data suggest that at least some speech sounds may be perceived in similar ways by birds and mammals.

In a third study, two budgerigars were trained on the end points of a synthetic /da/-/ta/ continuum and were then tested on intermediate VOTs (Dooling, Soli, et al., 1987). The two birds showed different but distinct boundaries along this continuum and, on average, these boundaries were at shorter VOTs than those reported for humans and chinchillas. The present experiment aims to extend these preliminary findings using more birds and a refined procedure for estimating perceptual categories from discrimination data.

With few exceptions, past work on speech perception by 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 of responses to intermediate and extreme forms of stimuli. We recently developed a new set of operant procedures for testing discrimination of complex sounds in birds that is more efficient and more useful for describing an

This work was supported by NIH Grants NS19006 and HD00512 to R. Dooling. We thank R. Diehl, P. Marler, J. D. Miller, T. Park, S. Soli, and C. S. Watson for comments on an earlier draft, and J. Downing and D. Reidell for care of the birds. Correspondence may be addressed to Robert J. Dooling, Psychology Department, University of Maryland, College Park, MD 20742.

animal's perceptual world (Okanoya & Dooling, 1988). This procedure utilizes a discrimination task structured so that shorter response latencies correlate well with greater perceptual dissimilarity; a relation that can be validated using simple pure tones (Dooling, Brown, Park, Okanoya, & Soli, 1987) and complex sounds (Dooling, Park, Brown, Okanoya, & Soli, 1987). These procedures produce similarity measures that are suitable for analysis using multidimensional scaling (MDS) techniques (Okanoya & Dooling, 1988).

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 that have similar perceptual properties are near each other in multidimensional space, whereas stimuli that have different perceptual properties are far apart. This makes MDS an ideal tool to analyze the phenomenon of perceptual categories. The following experiments tested 5 budgerigars on each of the synthetic VOT series (bilabial, alveolar, and velar).

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 5 birds had participated in other psychoacoustic studies involving both simple and complex sounds.

Stimuli

The stimuli were tokens from the bilabial /ba/-/pa/, alveolar /da/-/ta/, and velar /ga/-/ka/ VOT series. These tokens were synthesized with the Haskins Laboratories parallel resonance synthesizer using control parameters developed by Lisker and Abramson (1970). The 450-msec stimuli had VOTs ranging from 0 to +70 msec in 10-msec steps. These stimuli are the same as those schematically represented in Kuhl and Miller (1978, Figures 1 and 2, pp. 907-908). The birds were tested on one series at a time, and all 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.

Apparatus

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 two sensitive microswitches with light-emitting diodes (LEDs) attached. A bird could trip the microswitch by pecking at the LED. The left microswitch served as an observation key and the right microswitch served as a report key. An IBM AT microcomputer controlled all experimental events. The stimuli were stored on hard disk with a 12-bit quantization scheme, output at a sampling rate of 20 kHz, and low-pass filtered at 10 kHz to prevent aliasing.

Training and Testing Procedures

The procedures for training and testing birds were similar to those described previously (Okanoya & Dooling, 1988). We trained birds to peck one key (observation key) repeatedly during the repetitive presentation of one sound (background), and to peck the other key (report key) when a new sound (target) was presented alternately with the background sound. A peck on the report key during this alternating-stimulus pattern was defined as a correct response and was rewarded with a 2-sec access to food.

The procedure consisted of two distinct phases; a habituation phase (which was in effect each time a new background stimulus was selected) and a testing phase. During the habituation phase, the background sound repeated at the rate of 2/sec. This phase continued until the bird withheld responding on the report key for 10 sec or pecked the observation key four times. The purpose of this habituation phase was to reduce the rate of false responding. Immediately following the habituation phase was the testing phase, in which the LEDs on both the observation and the report keys were illuminated.

The same background stimulus that was presented during the habituation phase also continued during the testing phase; however, during the latter phase, a peck on the observation key initiated a random waiting interval of 1 to 7 sec. Following this interval, a peck on the observation key initiated an alternation of the target stimulus with the background stimulus. A response on the report key within 2 sec of the beginning of this alternating pattern was reinforced with a 2-sec access to food. If the bird did not peck either the observing key or the report key within 2 sec, the trial was ended and the response latency was recorded as 2 sec.

About 15% 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 the waiting interval was punished with a 16-sec timeout period. During this period, the lights in the test chamber were extinguished but the repeating sound continued. The variable waiting interval for the next trial began with the first peck during the observing phase (which began immediately after the end of the last trial). Thus, from the bird's point of view, sham trials in which no false alarms occurred were indistinguishable from the habituation phase.

A matrix of stimuli was constructed so that each stimulus could be paired with every other stimulus in the set. A row was randomly selected as a background stimulus for testing. Each stimulus in the row was then randomly selected as the target to be detected against the repeating background. The testing phase continued until the background stimulus was paired with every other stimulus in the set three times. All possible combinations within a row were tested three times before starting the next row.

A session proceeded one 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 began with a new background sound. This procedure continued until all rows were tested. Thus, the same sound served both as a background and as a target stimulus. Since each stimulus combination was tested three times, three response-latency matrices were available for analysis at the conclusion of testing.

Generally, the birds were tested in two daily sessions and about six of these sessions, each lasting about 30 min, were required to complete three matrices. In practice, each session lasted 30 min or until the bird stopped responding, whichever came first. If the bird failed to complete an entire matrix of stimulus combinations in a given session, the remaining trials were completed in the next session. A bird was tested until three entire matrices were available for analysis. Two additional criteria used for accepting a data matrix for subsequent analysis were that the overall correct detection for an entire matrix was above 75% with a false-alarm rate below 20%, and the bottom and top halves of the response-latency matrix were significantly correlated. This generally required between 1 and 2 weeks of daily testing.

For each bird, a median latency matrix was constructed from the three raw latency matrices. The upper and lower halves of the matrices were then averaged to produce a single latency half-matrix. This half-matrix was log-transformed to compensate for the positively skewed distribution of reaction times and was then analyzed by an MDS procedure, SINDSCAL (Arabie, Carroll, & Desarbo, 1987; Shepard, 1980). SINDSCAL produces a combined spatial map through a simultaneous analysis of several matrices. SIND-

SCAL has two advantages that make it especially useful in the present context. First, SINDSCAL imparts a preferred orientation to the dimensions of the final solution, with the first dimension accounting for most variance, the second dimension the next most, and so on. Second, SINDSCAL allows for an interpretation of individual differences, should they exist.

In the present experiment, a single, averaged, median half-matrix of response latencies for each bird was submitted for analysis, and a two-dimensional spatial map of the VOT stimuli was obtained as output. A two-dimensional solution was obtained because the second dimension accounts for a significant proportion of the variance. Furthermore, the purpose of this investigation was to determine whether changing a single component (VOT) within these stimulus complexes might result, for the avian ear, in more than the two perceptual categories typically observed for humans and chinchillas (Kuhl & Miller, 1978; Miller, Wier, Pastore, Kelly, & Dooling, 1976).

RESULTS

The analysis by SINDSCAL of the perceptual similarity among the eight bilabial VOT tokens for the 5 budgerigars is shown in Figure 1a. The subject weights for

Figure 1. (a) A two-dimensional spatial plot by SINDSCAL of the results from five budgerigars tested on the bilabial VOT stimuli. The two largest clusters by a cluster analysis are enclosed by dashed lines. (b) The results of a hierarchical cluster analysis on the average matrix of the 5 budgerigars shown as a dendrogram.

Table 1
Subject Weights for the Bilabial VOT Continuum

Bird	Dimension 1	Dimension 2
PLB	.695	.521
P28	.556	.446
P53	.801	.338
SPK	.842	.206
DFY	.662	.505

the SINDSCAL solution are given in Table 1. The total variance in response latency accounted for by the MDS solution was 73%, with the first and second dimensions accounting for 54% and 19%, respectively. The results of an average-linkage hierarchical cluster analysis (Alderfer & Blashfield, 1984) on the latency data averaged across birds are shown as a dendrogram in Figure 1b. Like MDS, cluster analysis also describes the structure of similarity data by grouping stimuli into subsets, each of which should correspond to a meaningful feature of the stimuli. Hierarchical cluster analysis groups, or links, stimuli by a predefined rule according to a Euclidean distance metric, with more similar calls clustered at the less aggregated levels of the hierarchy.

The eight bilabial VOT tokens were split into two groups along the first dimension in the stimulus space, with shorter VOTs on the right and longer VOTs on the left. The first dimension accounts for the greatest variance in response latencies. The subject weights indicate that all the birds were consistent in finding the acoustic changes along the first dimension more salient than those along the second dimension. We conclude from these results that bilabial VOT is a salient acoustic dimension for budgerigars, and that a perceptual boundary occurs in this continuum for budgerigars.

Identical analyses were conducted for both the alveolar and the velar VOT series. The two-dimensional spatial representation generated by SINDSCAL for the eight alveolar VOT tokens for the 5 budgerigars is shown in Figure 2a. The results of a hierarchical cluster analysis of the same data averaged across all five birds are shown as a dendrogram in Figure 2b. The subject weights for this solution are given in Table 2. The total variance in response latency accounted for by this MDS solution was 72%, with the first and second dimensions accounting for 64% and 8%, respectively.

As for the bilabial VOT series, the eight stimuli from the alveolar VOT series were also split into two groups along the first dimension. Short VOTs are on the right and long VOTs are on the left. As shown in Table 2, all 5 subjects were consistent in finding the first dimension the most salient. We conclude, therefore, that alveolar VOT is a salient acoustic dimension for budgerigars, and that a perceptual boundary occurs in this continuum for the birds.

For the velar VOT series, a similar analysis was conducted. The total variance in response latency accounted for by the MDS solution was 71%, with the first and second dimensions accounting for 51% and 20% of this variance, respectively. The stimulus space is shown in

Figure 2. (a) A two-dimensional spatial plot by SINDSCAL of the results from 5 budgerigars tested on the alveolar VOT stimuli. The two largest clusters by a cluster analysis are enclosed by dashed lines. (b) The results of a hierarchical cluster analysis on the average matrix of the 5 budgerigars shown as a dendrogram.

Figure 3a, and the results of a hierarchical cluster analysis on the data from the 5 birds are shown as a dendrogram in Figure 3b. The subject weights for this solution are given in Table 3. The velar stimuli in this experiment were generally more spread out than the bilabial and alveolar stimuli. Furthermore, the spatial map shows that the differences among the long VOT stimuli appear to be especially salient. This pattern of results at long VOTs for the budgerigar is different from those observed for humans and chinchillas (Kuhl & Miller, 1978). It may be that the enhanced discriminability for long VOT stimuli means that there are more than two boundaries for the budgerigar along this synthetic speech continuum.

These results provide strong evidence that budgerigars hear at least the bilabial and alveolar synthetic speech stimuli much like humans and chinchillas do. However,

the testing procedures used in this experiment—especially the analysis of response latencies by SINDSCAL—are quite different from the procedures normally used in examining speech perception in animals. Results from both animal and human experiments using these stimuli are typically presented in the form of labeling and discrimination functions. The 50% point (or “phonetic boundary”) for humans and chinchillas listening to the bilabial, alveolar, and velar continua are roughly 25, 35, and 42 msec, respectively (Kuhl & Miller, 1978). In addition, discrimination functions typically show an increase in discriminability between adjacent stimuli at this same phonetic boundary (see, e.g., Kuhl & Miller, 1975, 1978; Kuhl & Padden, 1982, 1983).

The latency data from our budgerigars are discrimination data, not labeling data. A function similar to the traditional discrimination function can be obtained by plotting response latencies for only adjacent stimulus pairs. Since longer response latencies indicate poorer discriminability and shorter response latencies indicate better discriminability, we would expect to find a trough in a plot of response latency versus VOT. Figure 4 shows the average relative response latencies for the five birds for adjacent stimuli from all three continua. These latency functions show a trough between 20 and 30 msec for the bilabial VOT series, between 30 and 40 msec for the alveolar series, and between 30 and 40 msec and 50 and 60 msec for the velar VOT series.

It is also possible, from a one dimensional MD-SCAL (Kruskal & Wish, 1978) solution, to determine the pair of adjacent VOT tokens that are most discriminable for each bird. However, to determine more precisely a VOT value that should correspond to the familiar “boundary” location obtained from labeling experiments, we also plotted the relative distances between stimuli (scale values from MD-SCAL) as a function of VOT. A third-order polynomial function was fit to these data and the point of steepest slope (which corresponds to the region of greatest discriminability) was taken as the category boundary. The polynomial function provided an excellent fit to the scale values for the bilabial and alveolar series, but a somewhat poorer fit to scale values for the velar series, as might be expected from the spatial plots. The average point of the steepest slope for the 5 birds is at 25.8 msec for the bilabial VOT series, 34.0 msec for the alveolar VOT series, and 41.0 msec for the velar VOT series. The location of the steepest slope for each bird for each series is given in Table 4.

Table 2
Subject Weights for the Alveolar VOT Continuum

Bird	Dimension 1	Dimension 2
PLB	.882	.128
P28	.669	.422
P53	.787	.405
SPK	.869	.170
DFY	.718	.105

Figure 3. (a) A two-dimensional spatial plot by SINDSCAL of the results from five budgerigars tested on the velar VOT stimuli. The two largest clusters by a cluster analysis are enclosed by dashed lines. (b) The results of a hierarchical cluster analysis on the average matrix of the 5 budgerigars shown as a dendrogram.

DISCUSSION

The purpose of this study was to provide comparative data from birds on the perception of synthetic speech sounds from three continua that ranged perceptually from voiced to voiceless to humans. We trained budgerigars on an auditory-discrimination task involving these speech sounds and then analyzed their response latencies with MDS and cluster analysis procedures. The results show that in the range of VOTs from 0 to +70 msec, budgerigars hear a perceptual change in the bilabial, alveolar, and velar VOT continua roughly at points at which humans and other mammals also hear an abrupt perceptual change. For the velar VOT continuum, there is also the hint that budgerigars may differ from humans and chinchillas in the ability to discriminate among the tokens at long VOTs.

These data are relevant for both methodological and theoretical reasons. MDS techniques have proven useful in

understanding human perception of complex sounds, such as speech and music, in which the physical correlates of the perceptual experience are often unclear (see, e.g., Murry & Singh, 1980; Shepard, 1980). The present experiment demonstrates that MDS techniques are also useful for studying the perception of complex speech sounds by animals. One distinct advantage of the present procedure, compared with other studies of speech perception in animals, is that the birds were not trained to a particular class of sounds, since each stimulus serves as both a background and a target. Thus, the perceptual categories revealed by these techniques cannot be due to asymmetries in reinforcement contingencies or training conditions.

The results from the present experiment are also of theoretical interest. Theories of categorical perception of speech differ primarily in the roles attributed to general auditory processing and phonetic processing (for a review, see Kewley-Port, Watson, & Foyle, 1988; Repp, 1984). The present results with budgerigars are relevant to these theories of speech perception.

Both temporal and spectral cues have been implicated in human perception of VOT boundaries (see, e.g., Miller et al., 1976; Pisoni, 1977; Soli, 1983; Stevens & Klatt, 1974; Summerfield & Haggard, 1977). Thus, for general processing theories to account for the present results, it would have to be shown that the temporal and spectral resolving powers of the human and budgerigar auditory systems are similar. Fortunately, such comparative data are available. The temporal resolving power of budgerigars measured psychoacoustically appears to be roughly similar to that reported for humans (Dooling & Haskell, 1978; Dooling & Searcy, 1981, 1985b). Thus, to the extent that temporal cues underlie these speech-sound categories, budgerigars and humans would be expected to show similar categories and boundaries.

The spectral resolving abilities of budgerigars and humans as measured psychoacoustically (e.g., frequency-difference limens, critical ratios, critical bands), however, are clearly different (Dooling & Saunders, 1975; Dooling & Searcy, 1985a; Saunders, Rintelmann, & Bock, 1979). Budgerigars show smaller critical bands than do humans between 2 and 3 kHz and larger critical bands than do humans below 2 kHz (Okanoya & Dooling, 1987). Such differences in spectral resolving power result in a differential coding of spectral information in complex sounds by the two species (Dooling, 1986; Dooling, Soli, et al., 1987). Since the present results show that budgerigars can resolve the phonemically relevant acoustic information contained in synthetic speech sounds, this find-

Table 3
Subject Weights for the Velar VOT Continuum

Bird	Dimension 1	Dimension 2
PLB	.719	.559
P28	.620	.508
P53	.671	.452
SPK	.777	.347
DFY	.714	.332

Figure 4. The relative response latency for adjacent stimulus pairs as a function of VOT for all three VOT series. The shortest latency occurred for the stimulus pairs that were most discriminable.

ing fails to support a large role for spectral features in the categorical perception of these stimuli.

Yet another possibility is that perceptual categories for speech are primarily dependent on more central processes (see Kewley-Port et al., 1988, for a review). To explain similar results across species, one must assume that the different species have in common the minimum required temporal and spectral resolving powers to encode the VOT continua in sufficient and similar detail. Categorical boundaries are then imposed by central mechanisms operating on the incoming, noncategorical information provided by the peripheral auditory system. Similar category boundaries in budgerigars, humans, and chinchillas would imply similar attentional and memory-storage mechanisms. This presents no major problem unless these central nervous system mechanisms are also taken to include higher order, motor-articulation processes unique to speech-sound production (Kuhl, 1986; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Studdert-Kennedy, Liberman, Harris, & Cooper, 1970).

Finally, the present experiment was in part motivated by results from an earlier attempt to train 2 budgerigars to discriminate among tokens in the /da/-/ta/ synthetic VOT series (Dooling, Soli, et al., 1987). This earlier experiment used a classification task in which the load on memory was high, as in previous VOT and vowel studies

with chinchillas (Burdick & Miller, 1975; Kuhl & Miller, 1975). Though it was possible to obtain boundaries for both birds (at a slightly shorter VOT value than is typically found for humans and chinchillas), the most telling result from this first study with budgerigars was the difficulty encountered in training the birds to discriminate between the endpoints of the /da/-/ta/ series. This difficulty stands in marked contrast to their ability to learn a similar discrimination problem involving pure tones or bird calls (Dooling, Brown, et al., 1987; Park & Dooling, 1985).

These two procedures differ in a number of ways. The most obvious difference is the length of time the birds had to "remember" the stimuli. In the earlier classification procedure, the birds had to remember which stimulus (i.e., 0 VOT or +70 VOT) was associated with which response (i.e., "GO" or "NOGO") for periods of tens of seconds. In the present discrimination procedure, the birds had only to retain relevant acoustic information for a few hundred milliseconds (i.e., a 450-msec stimulus duration plus a 50-msec interstimulus interval). The difference in the behavior of budgerigars in these two studies points to the involvement of central processes, since the primary difference between the procedures is one of memory load. We conclude, therefore, that budgerigars are different from chinchillas and humans in their ability to remember—rather than to discriminate—the salient acoustic features required for categorizing the /da/-/ta/ continuum.

In conclusion, the fact that chinchillas and humans discriminate among tokens from the VOT continua in similar ways has been taken as evidence for a close, and perhaps even special, relationship between the acoustic characteristics of speech and natural psychophysical or perceptual boundaries of the mammalian auditory system (Kuhl, 1986). The present results on speech perception

Table 4
Voice Onset Time (in msec) at Point of Steepest Slope

Bird	/ba/-/pa/	/da/-/ta/	/ga/-/ka/
PLB	10	25	44
P28	35	35	40
P53	20	41	44
SPK	30	32	35
DFY	34	37	42
Mean	25.8	34.0	41.0

in birds do not discount the idea that such a close relationship between production and perception exists, or that it plays a role in the evolution or acquisition of a speech-sound repertoire in humans. The present results do demonstrate, however, that, at least for VOT, the perceptual boundaries involved are not unique to humans or even to mammals.

REFERENCES

- ALDENDERFER, M. S., & BLASHFIELD, R. K. (1984). *Cluster analysis*. Newbury Park, CA: Sage Publications.
- ARABIE, P., CARROLL, J. D., & DESARBO, W. S. (1987). *Three-way scaling and clustering*. Newbury Park, CA: Sage Publications.
- BORG, I., & LINGOES, J. (1987). *Multidimensional similarity structure analysis*. New York: Springer-Verlag.
- 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.
- DOOLING, R. J. (1986). Perception of vocal signals by budgerigars (*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., & HASKELL, R. J. (1978). Auditory duration discrimination in the parakeet (*Melopsittacus undulatus*). *Journal of the Acoustical Society of America*, **63**, 1640-1643.
- 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. (1981). Amplitude modulation thresholds for the parakeet (*Melopsittacus undulatus*). *Journal of Comparative Physiology*, **143**, 383-388.
- DOOLING, R. J., & SEARCY, M. H. (1985a). Non-simultaneous auditory masking in the budgerigar (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **99**, 226-230.
- DOOLING, R. J., & SEARCY, M. H. (1985b). Temporal integration of acoustic signals by the budgerigar (*Melopsittacus undulatus*). *Journal of the Acoustical Society of America*, **77**, 1917-1920.
- DOOLING, R. J., SOLI, S. D., KLINE, R. M., PARK, T. J., HUE, C., & BUNNELL, T. (1987). Perception of synthetic speech sounds by the budgerigar (*Melopsittacus undulatus*). *Bulletin of the Psychonomic Society*, **25**, 139-142.
- 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.
- KEWLEY-PORT, D., WATSON, C. S., & FOYLE, D. C. (1988). Auditory temporal acuity in relation to category boundaries: Speech and non-speech stimuli. *Journal of the Acoustical Society of America*, **83**, 1133-1145.
- KLUENDER, K. R., DIEHL, R. L., & KILLEEN, P. R. (1987). Japanese quail can learn phonetic categories. *Science*, **237**, 1195-1197.
- KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage Publications.
- KUHL, P. K. (1979). Models and mechanisms in speech perception: Species comparisons provide further contributions. *Brain, Behavior, & Evolution*, **16**, 374-408.
- KUHL, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *Journal of the Acoustical Society of America*, **70**, 340-349.
- KUHL, P. K. (1986). The special-mechanisms debate in speech: Contributions of tests on animals (and the relation of these tests to studies using non-speech signals). *Experimental Biology: Sensory & Perceptual Processes*, **45**, 233-265.
- KUHL, P. K. (1987). The special-mechanisms debate in speech research: Categorization tests on animals and infants. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 355-386). Cambridge MA: Cambridge University Press.
- KUHL, P. K., & MILLER, J. D. (1975). Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science*, **190**, 69-72.
- KUHL, P. K., & MILLER, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America*, **63**, 905-917.
- KUHL, P. K., & PADDEN, D. M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception & Psychophysics*, **32**, 542-550.
- KUHL, P. K., & PADDEN, D. M. (1983). Enhanced discriminability at the phonetic boundaries for the place feature in macaques. *Journal of the Acoustical Society of America*, **73**, 1003-1010.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431-461.
- LISKER, L., & ABRAMSON, A. S. (1970). The voicing dimension: Some experiments in comparative phonetics. In *Proceedings of the 6th International Congress of Phonetic Sciences* (pp. 563-567). Prague: Academia.
- MILLER, J. D., WIER, C. C., PASTORE, R. E., KELLY, W. M., & DOOLING, R. J. (1976). Discrimination and labeling of noise-buzz sequences with varying noise lead times: An example of categorical perception. *Journal of the Acoustical Society of America*, **60**, 410-417.
- 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.
- OKANOYA, K., & DOOLING, R. (1988). Obtaining acoustic similarity measures from animals: A method for species comparisons. *Journal of the Acoustical Society of America*, **83**, 1690-1693.
- PARK, T. J., & DOOLING, R. J. (1985). Perception of species-specific contact calls by budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, **99**, 391-402.
- PISONI, D. B. (1977). Identification and discrimination of the relative onset of two component tones: Implications for the perception of voicing stops. *Journal of the Acoustical Society of America*, **61**, 1352-1361.
- REPP, B. H. (1984). Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (pp. 99-131). New York: Academic Press.
- 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.
- SOLI, S. D. (1983). The role of spectral cues in discrimination of voice onset time differences. *Journal of the Acoustical Society of America*, **73**, 2150-2165.
- STEVENS, K. N., & KLATT, D. H. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, **55**, 653-659.
- STUDDERT-KENNEDY, M., LIBERMAN, A., HARRIS, K. S., & COOPER, F. S. (1970). Motor theory of speech perception. *Psychological Review*, **77**, 234-249.
- SUMMERFIELD, Q., & HAGGARD, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. *Journal of the Acoustical Society of America*, **62**, 436-448.
- WATERS, R. S., & WILSON, W. A., JR. (1976). Speech perception by rhesus monkeys: The voicing distinction in synthesized labial and velar stop consonants. *Perception & Psychophysics*, **19**, 285-289.