

# Perception of synthetic /ba-/wa/ speech continuum by budgerigars (*Melopsittacus undulatus*)

Micheal L. Dent, Elizabeth F. Brittan-Powell, Robert J. Dooling, and Alisa Pierce  
*Department of Psychology, University of Maryland, College Park, Maryland 20742*

(Received 15 March 1996; accepted for publication 1 May 1997)

Other than humans, extensive vocal learning has only been widely demonstrated in birds. Moreover, there are only a handful of avian species that are known to be good mimics of human speech. One such species is the budgerigar (*Melopsittacus undulatus*), which is a popular mimic of human speech and learns new vocalizations throughout adult life. Using operant conditioning procedures with a repeating background task, we tested budgerigars on the discrimination of tokens from two synthetic /ba-/wa/ speech continua that differed in syllable, but not transition, duration. Budgerigars showed a significant improvement in discrimination performance on both continua near the phonetic boundary for humans. Budgerigars also showed a shift in the location of the phonetic boundary with a change in syllable length, similar to what has been described for humans and other primates. These results on a nonmammalian species provide support for the operation of a general, nonphonetic, auditory process as one mechanism which can lead to the well-known stimulus-length effect in humans. © 1997 Acoustical Society of America. [S0001-4966(97)03409-7]

PACS numbers: 43.80.Lb, 43.71.Es, 43.66.Gf [FD]

## INTRODUCTION

Budgerigars (*Melopsittacus undulatus*) are small Australian parrots that have recently been shown to produce a variety of human speech phonemes either singly, during warble-song, or in response to presented objects (Banta and Pepperberg, 1995). These birds have also proven to be excellent subjects for psychoacoustic studies. Recent studies have shown that budgerigars, like humans, perceive the vowel tokens /i/, /a/, /e/, and /u/ in phonetically appropriate categories in spite of variation in talker, pitch contour, and gender (Dooling and Brown, 1990; Dooling, 1992). In other studies on the perception of consonants by budgerigars (Dooling *et al.*, 1989), perceptual boundaries were near the human boundaries for voice-onset-time (VOT) pairs of /ba-/pa/ (bilabial), /da-/ta/ (alveolar), and /ga-/ka/ (velar) continua. Perception of speech sound categories and discrimination among phonetically relevant speech tokens are not unique to budgerigars. Other studies have shown that zebra finches, starlings, quail, blackbirds, and pigeons also discriminate and categorize speech sounds similar to humans (Hienz *et al.*, 1981; Kluender, 1991; Dooling, 1992; Dooling *et al.*, 1995).

There are several reasons why it is interesting to study how birds discriminate among speech sounds. First, some species of birds such as starlings, mynahs, and some parrots modify their vocalizations throughout their lives and are so flexible in what they will accept as acoustic models for learning that they even mimic human speech (e.g., Thorpe, 1959; Greenewalt, 1968; Klatt and Stefanski, 1974; Dooling, 1986; Pepperberg, 1990; Farabaugh *et al.*, 1994; Patterson and Pepperberg, 1994; Warren *et al.*, 1996). These well-known instances of human speech production show that birds can extract the important acoustic features of speech, despite different tutors. Evidence from a variety of studies shows they can imitate many phonemes, including all vowel sounds (Patterson and Pepperberg, 1994) and almost every conso-

nant sound in the English language (Turney *et al.*, 1994). Second, birds that can produce speech potentially provide the only known opportunity for testing the relation between production and perception of human speech in a nonhuman organism. Third, it is well known that budgerigars and other birds have the remarkable ability to regenerate their hair cells following acoustic overexposure or treatment with ototoxic drugs (Cotanche, 1987; Corwin and Cotanche, 1988; Ryals and Rubel, 1988). In the case of budgerigars, both the perception of species-specific vocal signals and the precision in production of learned contact calls are affected when hearing is lost from ototoxic drugs. However, both production and perception of these vocal signals recover as the auditory periphery becomes repopulated with new hair cells (Dooling *et al.*, in press). Because of this regenerative capability, birds provide a unique model for studying speech perception (and production in the case of budgerigars) after destruction and subsequent repair of the peripheral auditory system. Fourth, experiments with various speech categories and speech continua have shown that mammals such as chinchillas and monkeys with auditory capabilities similar to humans tend to show phonetic boundaries and categories similar to the human phonetic boundaries (Burdick and Miller, 1975; Kuhl and Miller, 1975; Kuhl and Padden, 1982; Kuhl, 1987). Birds have both peripheral and central auditory systems that are profoundly different than the mammalian auditory system (for a review, see Manley, 1990; Carr, 1992; Manley and Gleich, 1992), so they contribute a different perspective on the role of mammalian auditory processing in speech perception.

Thus there are numerous reasons for studying speech perception in birds in addition to the general comparative strategy of using animals to test whether a particular perceptual performance exhibited by humans listening to speech is a uniquely human phenomenon. A recent study by Sinnott and her colleagues (Sinnott *et al.*, submitted) is intriguing in

this respect and provides some of the motivation for the present study. They show that a nonhuman primate exhibits a boundary along a human speech continuum that is usually explained, in humans, by normalization for articulation rates. These results are also interesting because nonhuman mammals lack a supralaryngeal cavity and are incapable of producing human speech (for a review see Ploog, 1992). Birds, on the other hand, do not use their larynx for sound production. Instead, the syrinx acts as the sound source and is positioned such that birds have a suprasyringeal cavity. Several species have been shown to mimic human speech sounds and phrases extremely well.

Many experiments with humans have investigated the consonant–vowel (CV) /ba/–/wa/ continuum and have shown that information occurring later in the speech stream affects the perception of an earlier occurring cue (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989). As syllable duration increases, the /b/–/w/ perceptual boundary for humans moves toward transitions of longer duration, resulting in a phonetic boundary shift. Though it is difficult to conceive of a purely psychoacoustic explanation for this stimulus-length effect (see, for example, Pisoni *et al.*, 1983; Diehl and Walsh, 1989), comparative work tends to support psychoacoustic explanations for perceptual categories, in general.

One psychoacoustic explanation that has been suggested as playing a role in the shift of boundary location in the stimulus-length effect is backward masking (Jamieson, 1987). When the steady-state portion of the CV syllable increases, the transition may be more effectively masked, leading to a boundary shift. By showing that reducing the vowel intensity level also shifts boundary location, Jamieson provided evidence that some form of backward masking may provide the basis for such an effect.

In earlier experiments, quail performed similar to humans in their discrimination of several phonetic categories, and the authors argue that phonetic categories are natural auditory groupings, even though speech has a different functional significance for humans and nonhumans (Diehl and Kluender, 1989). These authors (Diehl and Kluender, 1989) suggest an auditory enhancement hypothesis for speech perception, where human listeners are particularly sensitive to the auditory cues that define phonetic categories and may typically discriminate changes at phonetic boundaries more efficiently than animals.

A more prevalent explanation for these results is the phonetic one that suggests listeners are exhibiting perceptual normalization for speaking. Miller and Liberman (1979) suggest that the listener interprets a set of acoustic cues in running speech in relation to the speaker's rate of articulation rather than by reference to some absolute value, a theory supported by many other studies on speech perception (e.g., Dorman *et al.*, 1977; Gay, 1978; Repp *et al.*, 1978; Repp, 1982). When Miller and Liberman added a stop consonant to the end of their syllables (/bad/–/wad/), the phonetic boundary shift occurred in the opposite direction, suggesting that listeners normalize for articulation rate.

Because we know that budgerigars can produce /b/ and /w/, the present experiment sought to examine how budgeri-

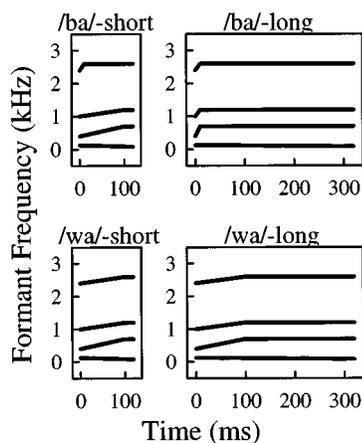


FIG. 1. Schematic of the spectrogram of the end points of the /ba/ and /wa/ continua for the short (120 ms) and long (320 ms) stimulus sets.

gars discriminated among tokens on a synthetic /ba/–/wa/ speech continuum. For these tests, we used a standard discrimination (low-uncertainty) task that has been used in a variety of other experiments on the discrimination of both simple and complex sounds in these birds (Dooling *et al.*, 1987). While budgerigars have demonstrated perceptual boundaries for other consonant contrasts using discrimination procedures, the /ba/–/wa/ contrast affords the opportunity to test the phenomenon of the boundary shift with increases in syllable duration.

## I. METHODS

### A. Subjects

Three adult budgerigars (two females and one male) were used as subjects. All of the birds were housed in a vivarium at the University of Maryland and were kept on a day/night cycle corresponding to the season. The birds were either purchased from a local pet store or bred in the vivarium. They were kept at approximately 90% of their free-feeding weight during the course of the experiment. In addition, four adult humans who were native speakers of English were used as subjects. All four were female students, ranging in age from 24 to 25 yr, working in the laboratory at the time of the experiment but who had no previous testing experience with these sounds. None of the subjects reported a history of speech or hearing disorders or spoke other languages fluently.

### B. Stimuli

The stimuli used in this experiment were two full-formant /ba/–/wa/ speech continua differing in length (Fig. 1). These speech sounds were generated by the Canadian speech research environment (CSRE) program (using a 10-kHz sampling rate) according to the parameters of Sinnott *et al.* (submitted). The duration of the initial formant transition changed in 10-ms steps to yield a set of ten stimuli ranging perceptually from /ba/ to /wa/. Figure 1 shows a schematic representation of the end-point stimuli for both the short and long continua. For both continua, F1 began at 400 Hz and moved to 700 Hz over a variable time period ranging

from a 10-ms /ba/ to a 100-ms /wa/.  $F_2$  moved from 1000 to 1200 Hz, and  $F_3$  moved from 2400 to 2600 Hz over the same time periods.  $F_4$  and  $F_5$  had no formant transitions.  $F_0$  fell linearly over the duration of the syllable from 125 to 80 Hz. All stimuli were gated externally with a 10-ms rise time. The duration of the steady-state vowel was manipulated so that the overall duration of the short syllables was 120 ms and the long syllables were 320 ms.

We chose to use these stimuli instead of the other stimulus continua on which humans have been tested (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989) to facilitate comparison with the only other animal data available on this speech continuum. All stimuli were presented at a peak sound-pressure level of 68 dB.

### C. Apparatus and procedure

The apparatus and procedure for testing the birds has been described previously (Okanoya and Dooling, 1987, 1991). The response panel and procedures are shown schematically in Fig. 2. The birds were tested in a wire cage (23×25×16 cm) that was placed in a small, foam-lined, sound isolation booth (IAC model IAC-2; 57×60×78 cm). A response panel consisting of two sensitive microswitches with light-emitting diodes (LEDs) was mounted on the wall of the test cage just above the food hopper. The microswitch was tripped by the bird pecking the LED. The left microswitch and LED served as the observation key, and the right microswitch and LED served as the report key. The speech stimuli were delivered from a JBL loudspeaker (model 2105H) mounted above the test cage. The experiment was controlled by an IBM 486 microcomputer operating Tucker–Davis Technology electronic and DSP modules. The behavior of the animals during test sessions was monitored continuously by a video camera system.

The birds were trained by a standard operant autoshaping program to peck the observation key during a continuously repeating background (interstimulus-interval of 380 ms) and to peck the report key when a new sound (target) was presented alternately with the background sound. A peck to the report key within 2 s following onset of an alternating sound pattern was considered a hit. A hit was reinforced with a 2-s access to food on a schedule of 80%–100%. Stimuli from the long CV continuum were tested first, and stimuli from the short CV stimuli were tested second.

During the testing phase, a peck on the observation key began a random interval of 1–6 s. Following this interval, the next peck on the observation key initiated a trial (defined as an alternation of the target stimulus with the background stimulus). The dependent variable in these experiments was the response latency on each trial involving a target stimulus (i.e., the alternating sound pattern with the target stimulus alternating with the background stimulus). Previous work on the perception of both simple and complex sounds (including speech sounds) has shown that response latency is a valid measure of stimulus similarity for these birds (Dooling *et al.*, 1995). If the bird failed to respond (miss), a response latency of 2000 ms was recorded, and a new trial sequence began with the observation interval for the next trial. Approxi-

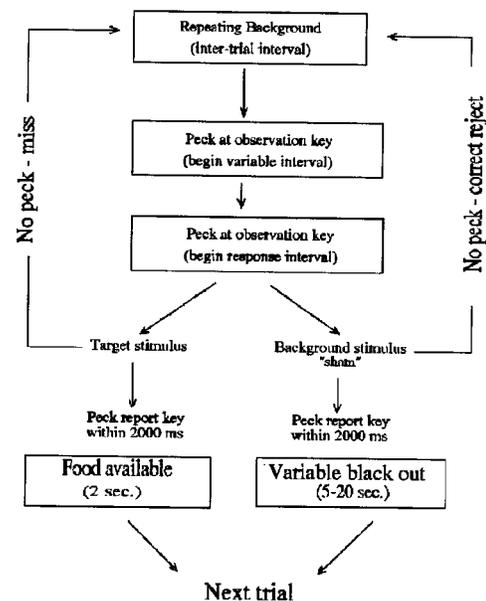
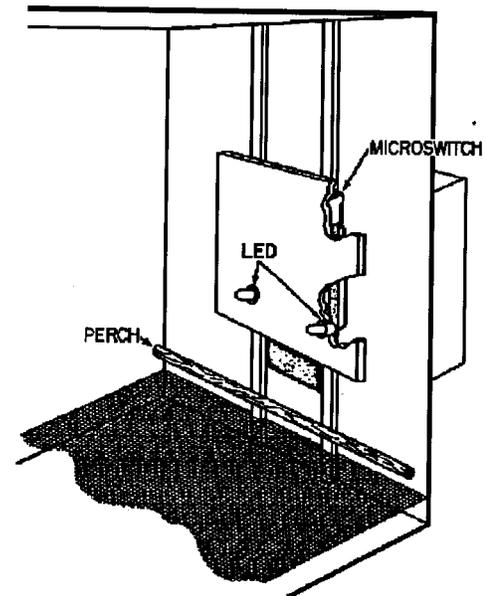


FIG. 2. Schematic of testing apparatus showing the arrangement of LEDs, food hopper, and speaker (top) and a diagram of the alternating sound task (bottom). After the variable interval begins, the next peck on the observation key starts the response interval. During the response interval, one of two trial types occurs: (1) if a target stimulus is presented to the bird, pecking the report key yields a food reward or (2) if a sham stimulus is presented, a peck to the report key causes a blackout interval. Failure to peck the report key during each of the trials starts a new intertrial interval.

mately 10% 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 variable blackout period during which the lights in the chamber were extinguished while the repeating background continued. The blackout time ranged from 5–20 s depending on the bird's recent history for false responses.

The strategy in this experiment was to test each stimulus in the continuum against every other stimulus in the con-

tinuum. The order of stimulus testing was determined prior to the experiment. A matrix of stimuli (background×target) was constructed, and one row was selected at random from this matrix for testing. The first stimulus in the row then served as the background (the continuously repeating stimulus), and the remaining stimuli in the row served as targets (selected randomly on a trial-by-trial basis). Testing continued until each target had been tested ten times and all rows were tested (i.e., until each sound served both as a background and as a target stimulus ten times). At the conclusion of testing (usually ten sessions), a 10×10 matrix of response latencies was available for analysis where each cell in the matrix contained a mean response latency based on the ten trials of each stimulus pair.

To ensure that these stimuli were satisfactory replicas of natural speech, we also tested human subjects on the same stimuli using two procedures: the discrimination task the birds were tested on and a labeling task. In the discrimination task, humans listened to the stimuli through headphones (Realistic model Nova 67) while pressing response keys on a hand-held panel. The human subjects viewed a video screen showing the inside of the animal operant chamber to provide feedback as to whether their response was correct (a hit, hopper activation) or incorrect (false alarm, lights extinguished). In addition, human subjects were also tested on an identification task in which they were required to label each stimulus played to them as either /ba/ or /wa/. In this test, the subjects pressed a button on a computer keyboard to randomly play one of the ten tokens, and then wrote down their response. In all, the ten tokens from each continuum were played ten times in random order, for a total of 100 trials for each continuum. Results were analyzed as the percent of responses labeled /ba/ for each subject, and the boundary along each continuum was defined as the 50% point.

#### D. Analysis

At the conclusion of testing on the repeating background discrimination task, matrices of percent correct values and response latencies were obtained for each subject (budgerigars and humans) for each stimulus contrast along the continua (Dooling *et al.*, 1995). The 10×10 matrices described above were folded about the diagonal, and the corresponding cells in the upper and lower halves were averaged to obtain triangular matrices that contained an average value from 20 trials for each stimulus contrast for each subject (10 vs 20 combined with 20 vs 10). From these matrices, we obtained both percent correct data and response latency data for each subject for all possible pairwise contrasts. A two-way analysis of variance (ANOVA) was used to test whether response latencies between the short and long continua differed, and whether the specific stimulus pairs along the continua differed from each other. The stimulus comparisons were done for both one-step contrasts (e.g., 10/20) and two-step contrasts (e.g., 10/30).

## II. RESULTS

The one-step stimulus comparisons were almost impossible, with each bird averaging less than 20% correct. There-

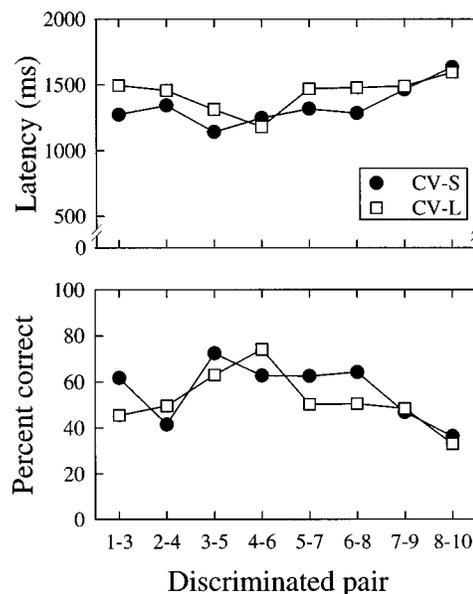


FIG. 3. Discrimination performance for the budgerigars for the two step stimuli comparison. Both response latency (top panel) and percent correct (bottom panel) are plotted as a function of length and show a peak in the discrimination function that shifts to the right with the increase in syllable duration.

fore we analyzed only two-step stimulus comparisons, which the birds discriminated among more easily for both the short and long stimulus conditions (Fig. 3). The two-step discrimination task response latency was inversely correlated with percent correct for all three birds both on the short ( $r = -0.98, -0.90, -0.98$ ) and long ( $r = -0.99, -0.99, -0.89$ ) stimulus continua. For the short stimulus (CV-S) continuum, the three budgerigars showed the best discrimination for the 30/50 stimulus pair (percent correct=72, response latency=1140 ms) while for the long stimulus (CV-L), the 40/60 pair was discriminated best (percent correct=74, response latency=1180 ms).

Across both continua, budgerigars were significantly better at discriminating the CV-S continuum than the CV-L continuum [ $F(1,7) = 5.29, p < 0.05$ ]. There were also significant differences along both the continua [ $F(7,7) = 4.68, p < 0.05$ ], with discrimination performance better near the center of each continuum and with the discrimination peak in a different location for short compared to long duration syllables (30/50 for CV-S and 40/60 for CV-L).

Humans were tested on both a repeating background discrimination task as well as an identification task. Like the birds, they discriminated among tokens in the CV-S continuum more easily than those of the CV-L continuum [ $F(1,7) = 8.85, p < 0.05$ ]. There were also significant differences along the continua [ $F(7,7) = 6.86, p < 0.05$ ]. When tested on the identification task, humans showed clear phonetic boundaries and boundary shifts similar to those typically reported by speech researchers using other /ba/-/wa/ continua. Humans showed slightly higher boundary locations than the budgerigars but a similar boundary shift with increase in stimulus duration. For the CV-S continuum, the boundary was 49 ms, but for the CV-L continuum, the

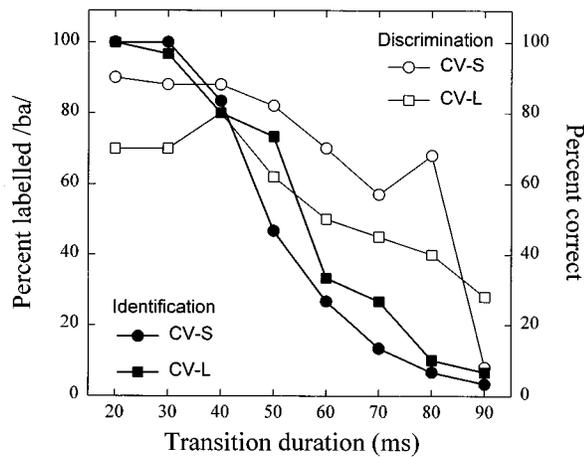


FIG. 4. Discrimination and identification performance for the humans for the two-step stimuli comparisons. Percent correct for the discrimination task (open symbols) show decreases in performance with the increase in transition duration (from /ba/ to /wa/). The boundaries for the identification task (closed symbols) shift from 49 ms for the short to 56 ms for the long stimuli.

boundary was 56 ms. The discrimination and identification results for humans are shown in Fig. 4.

The shift in human boundaries from the identification task were similar to shifts in the peaks in the discrimination functions in the CV-S and CV-L continua for budgerigars. However, the peaks in discrimination for the human subjects were not nearly as clear as for the birds. Instead, human performance showed a change in discrimination that followed Weber's law (percent correct decreased and response latency increased as transition durations increased), an effect noted by others when humans are tested on some speech sound discriminations with a low uncertainty task.

### III. DISCUSSION

It is well known that budgerigars can mimic a variety of human speech sounds. Despite the fact that budgerigar peripheral (Manley, 1990; Manley and Gleich, 1992) and central (Brauth *et al.*, 1987; Brauth, 1988; Striedter, 1994) auditory systems are remarkably different from the mammalian auditory system, budgerigars have nevertheless been shown to discriminate among a variety of speech sounds including /ba-/pa/, /ga-/ka/, /da-/ta/, /ra-/la/ and a variety of natural and synthetic vowels (Dooling *et al.*, 1989, 1995; Dooling and Brown, 1990). The present results add to this database by showing that budgerigars discriminate among /ba-/wa/, phonemes they are capable of producing (Turney *et al.*, 1994), in a way that is consistent with the stimulus-length effect shown in humans.

Studies of speech perception in birds, particularly those that are speech mimics, may have particular relevance for understanding the evolution of human acoustic communication. Several recent studies of parrot phonation have shown that parrots may use their tongue to produce vowels in some ways which are similar to humans (Patterson and Pepperberg, 1994; Warren *et al.*, 1996) and that tracheal resonances may affect certain aspects of vocal production (Brittan-Powell *et al.*, 1997). Moreover, some psittacines also demonstrate a remarkable ability to mimic a variety of nonvocal

human behaviors, including arm, leg, and head movements (Moore, 1992). We know from both anecdotal as well as scientific evidence that budgerigars show socially dependent vocal flexibility throughout life (Farabaugh *et al.*, 1994; Banta and Pepperberg, 1995; Brittan-Powell *et al.*, in press). Taken together, these findings raise the possibility that exposure to talking humans may sensitize these birds to some of the critical acoustic features of human speech. Budgerigars, for instance, show a peak in discrimination of a sinewave /ra-/la/ continuum whereas zebra finches tested under identical conditions do not (Dooling *et al.*, 1995).

A wealth of comparative studies on the perception of speech by animals have shown that uniquely human structures and perceptual mechanisms are not necessary for obtaining humanlike discrimination and classification of English speech sounds (e.g., Kuhl and Miller, 1975; Hienz *et al.*, 1981; Dooling *et al.*, 1989, 1995; Kluender, 1991; Sinnott *et al.*, submitted). Birds provide an important addition to this database both because their auditory system is significantly different from that of humans and other mammals and because they have the ability to regenerate auditory hair cells following acoustic overexposure or treatment with ototoxic drugs (Cotanche, 1987; Corwin and Cotanche, 1988; Ryals and Rubel, 1988). Elsewhere, we have shown that following hair cell damage, both perception and production of species-typical contact calls are only disrupted for a short time before returning to normal (Dooling *et al.*, in press). These vocal signals recover as the auditory periphery becomes repopulated with new hair cells (Dooling *et al.*, in press), so these birds provide a unique model for studying speech perception after repair of the peripheral auditory system.

It is also worthwhile considering whether there is a common mechanism producing the stimulus-length effect in human, nonhuman primates, and budgerigars. At least in budgerigars and nonhuman primates, one nonphonetic explanation for the stimulus-length effect is backward masking, where later information in the syllable masks earlier important cues. This may be true as well of humans. This common acoustic explanation based on backward masking would be strengthened if there was evidence for similar backward masking thresholds in budgerigars and humans for nonspeech stimuli. Fortunately, nonsimultaneous auditory masking data are available for budgerigars, and we know that they show backward (but not simultaneous or forward) masking thresholds that are nearly identical to those described in humans (Dooling and Searcy, 1980). To some extent then, the similar shift in the /ba-/wa/ phonetic labeling boundaries for humans and the shift in the location of the discrimination peak in budgerigars may be due to similar sensitivities to backward masking. Such results would be consistent with the auditory account of the stimulus-length effect in the perception of /ba-/wa/ as proposed by Diehl and Walsh (1989), who found such effects not only with speech continua but also with nonspeech stimuli.

It is important to recognize that there are well-known effects of testing paradigm on speech discrimination and

speech perception findings including the stimulus-length effect in the /ba/-/wa/ continuum. Phonetic boundaries tend to follow Weber's law with weak or nonexistent phonetic boundaries when a low uncertainty discrimination task was used compared to when a high uncertainty identification task was used (Kewley-Port *et al.*, 1988). MacMillan *et al.* (1988) found similar results with vowels, but not consonants. Our human subjects showed clear labeling phonetic boundaries, but no corresponding peaks in their discrimination functions. When tested on an identification task our human subjects showed not only the expected phonetic boundaries for the /ba/-/wa/ continua, but also the expected shift with increasing syllable length. A review of the literature shows that as syllable duration increases from about 100 ms to about 300 ms, the boundary changes from around 30 ms to around 45 ms (Miller and Liberman, 1979; Godfrey and Millay, 1981; Pisoni *et al.*, 1983; Diehl and Walsh, 1989; Sinnott *et al.*, submitted). The budgerigars in this study had peaks in discrimination functions at 40 ms for the 120-ms stimuli and 50 ms for the 320-ms stimuli. This shift falls in the range of the data from humans and monkeys, suggesting common mechanisms.

One popular view of the biology of speech suggests that the auditory system has driven the selection of speech sound categories by providing natural perceptual categories, or broad acoustic targets, for speech sounds that the articulatory system has subsequently evolved to match. In the course of language acquisition, these broad perceptual categories or proclivities may be altered by environmental input during development to become more precisely defined (for a review, see Kuhl, 1989). This leads to, among other things, the many well-known cross cultural differences in the production and perception of speech sound categories (see, for example, Miyawaki *et al.*, 1975). At the other end of the continuum are arguments for phonetic, uniquely human specializations for the perception of speech, such as the findings that even young infants correctly perceive a particular consonant such as /d/ even when the acoustic information for the consonant /d/ varies tremendously depending on speaker, speaking rate, and vowel context, among other things (for a review, see Kuhl, 1989).

The present results contribute to what we know of the biology of speech by providing more evidence that "general" properties of the vertebrate auditory system underlie the perception of speech sound categories—in this case a bird, the budgerigar, shows the vowel length effect. Changing the rate of speaking obviously results in a number of changes in the acoustics of the speech stream, including eliminating and reducing pauses between words and phrases and shortening words. Shortening words brings with it changes in the temporal and spectral cues affecting both vowels and consonants. It seems reasonable to suppose that speakers may vary some features of speech (e.g., vowel duration) over other features when changing speaking rate because the auditory system demands it. Changes in some features may be necessary to maintain optimal processing of other, more critical features. An animal example of such a strategy comes from the horseshoe bat (*Rhinolophus ferrumequinum*). These bats compensate for flight-induced

Doppler shifts in the frequency of their echoes by lowering the frequency of subsequent calls (Schnitzler, 1973). In this way, the frequency of the echo returning to the bat remains in a narrowly tuned frequency range which is disproportionately represented throughout the auditory system from the cochlea to the cortex (Neuweiler, 1980; Pollak, 1980). This specialization allows the horseshoe bat to optimize the processing of the returning echoes.

There are clearly aspects of speech perception that cannot be easily explained by the simple auditory mechanisms as described above, which argue for other hypotheses. For the present speech contrast, the phonetic interpretation argues that human listeners use rate-based articulation cues to adjust their perception of phonemes. A compelling case for this argument is seen in the work of Miller and Liberman (1979) where the addition of a stop consonant on the end of the syllable produced a boundary shift in the opposite direction. It would be interesting to conduct a corresponding experiment in budgerigars and perhaps other birds as one way of testing the generality of auditory accounts of the stimulus-length effect. The point is not to dispute the fact that linguistic experience is important for general phonetic categorizations in humans. Instead, the present data showing that budgerigars exhibit the stimulus-length effect seen in humans, and the comparative approach to speech perception in general, aims to push the auditory account of the stimulus-length effect to its limit. To the extent this can be done, then uniquely common human capabilities are not *required* for perceiving the sounds of speech.

## ACKNOWLEDGMENTS

This work was supported by NIH Grants No. DC-00198 and No. MH-00982 to R.J.D. and No. MH-10993 to E.F.B-P. We thank J. Sinnott, T. Kidd, M. Burr, S. Amagai, K. O'Grady, and K. Nepote for assistance.

- Banta, P. A., and Pepperberg, I. M. (1995). "Learned English vocalizations as a model for studying budgerigar (*Melopsittacus undulatus*) warble song," in *Nervous Systems and Behavior Proceedings of the 4th International Congress of Neuroethology*, edited by M. Burrows, T. Matheson, P. L. Newland, and H. Schuppe (Thieme, New York), p. 335.
- Brauth, S. E., and McHale, C. M. (1988). "Auditory pathways in the budgerigar. II. Intratelencephalic pathways," *Brain Behav. Evol.* **32**, 193–207.
- Brauth, S. E., McHale, C. M., Brasher, C. A., and Dooling, R. J. (1987). "Auditory pathways in the budgerigar. I. Thalamo-telencephalic pathways," *Brain Behav. Evol.* **30**, 174–199.
- Brittan-Powell, E. F., Dooling, R. J., and Farabaugh, S. M. (in press). "Vocal development in budgerigars (*Melopsittacus undulatus*): Contact calls," *J. Comp. Psych.*
- Brittan-Powell, E. F., Dooling, R. J., Larsen, O. N., and Heaton, J. T. (1997). "Mechanisms of vocal production in budgerigars (*Melopsittacus undulatus*)," *J. Acoust. Soc. Am.* **101**, 578–589.
- Burdick, C. K., and Miller, J. D. (1975). "Speech perception by the chinchilla: discrimination of sustained /a/ and /i/," *J. Acoust. Soc. Am.* **58**, 415–427.
- Carr, C. E. (1992). "Evolution of the central auditory system in reptiles and birds," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 511–543.
- Corwin, J. T., and Cotanche, D. A. (1988). "Regeneration of sensory hair cells after acoustic trauma," *Science* **240**, 1772–1774.
- Cotanche, D. A. (1987). "Regeneration of hair cell stereociliary bundles in the chick cochlea following severe acoustic trauma," *Hearing Res.* **30**, 181–196.

- Diehl, R. L., and Kluender, K. R. (1989). "On the objects of speech perception," *Ecol. Psych.* **1**, 121–144.
- Diehl, R. L., and Walsh, M. A. (1989). "An auditory basis for the stimulus-length effect in the perception of stops and glides," *J. Acoust. Soc. Am.* **85**, 2154–2164.
- Dooling, R. J. (1986). "Perception of vocal signals by budgerigars (*Melopsittacus undulatus*)," *Exp. Biol.* **45**, 195–218.
- Dooling, R. J. (1992). "Perception of speech sounds by birds," in *Advances in Biosciences: Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, London), pp. 407–413.
- Dooling, R. J., Best, C. T., and Brown, S. D. (1995). "Discrimination of synthetic full-formant and sinewave /ra-la/ continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*)," *J. Acoust. Soc. Am.* **97**, 1839–1846.
- Dooling, R. J., and Brown, S. D. (1990). "Speech perception by budgerigars (*Melopsittacus undulatus*): Spoken vowels," *Percept. Psychophys.* **47**, 568–574.
- Dooling, R. J., Manabe, K., and Ryals, B. M. (1996). "Effect of masking and hearing loss on vocal production and vocal learning in budgerigars," *Assoc. Res. Otolaryngology Abstr.* 581.
- Dooling, R. J., Okanoya, K., and Brown, S. D. (1989). "Speech perception by budgerigars (*Melopsittacus undulatus*): The voiced-voiceless distinction," *Percept. Psychophys.* **46**, 65–71.
- Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K., and Soli, S. D. (1987). "Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*). II. Vocal signals," *J. Comp. Psych.* **101**, 367–381.
- Dooling, R. J., Ryals, B. M., and Manabe, K. (in press) "Recovery of hearing and vocal behavior after hair cell regeneration," *Proc. Natl. Acad. Sci.*
- Dooling, R. J., and Searcy, M. H. (1980). "Forward and backward auditory masking in the parakeet (*Melopsittacus undulatus*)," *Hearing Res.* **3**, 279–284.
- Dorman, M. F., Studdert-Kennedy, M., and Raphael, L. J. (1977). "Stop-consonant recognition: Release bursts and formant transitions as functionally equivalent, context-dependent cues," *Percept. Psychophys.* **22**, 109–122.
- Farabaugh, S. M., Linzenbold, A., and Dooling, R. J. (1994). "Vocal plasticity in budgerigars (*Melopsittacus undulatus*): Evidence for social factors in the learning of contact calls," *J. Comp. Psych.* **108**, 81–92.
- Gay, T. (1978). "Effect of speaking rate on vowel formant transitions," *J. Acoust. Soc. Am.* **63**, 223–230.
- Godfrey, J. J., and Millay, K. K. (1981). "Discrimination of the tempo of frequency change cue," *J. Acoust. Soc. Am.* **69**, 1446–1448.
- Greenewalt, C. H. (1968). *Bird song: Acoustics and Physiology* (Smithsonian, Washington, DC).
- Hienz, R. D., Sachs, M. B., and Sinnott, J. M. (1981). "Discrimination of steady-state vowels by blackbirds and pigeons," *J. Acoust. Soc. Am.* **70**, 699–706.
- Jamieson, D. G. (1987). "Studies of possible psychoacoustic factors underlying speech perception," in *The Psychophysics of Speech Perception NATO Advanced Research Workshop on Psychophysics of Speech Perception*, edited by M. E. H. Schouten (Kluwer Academic, Hingham, MA), pp. 220–229.
- Kewley-Port, D., Watson, C. S., and Foyle, D. C. (1988). "Auditory temporal acuity in relation to category boundaries; speech and nonspeech stimuli," *J. Acoust. Soc. Am.* **83**, 1133–1145.
- Klatt, D. H., and Stefanski, R. A. (1974). "How does a mynah bird imitate human speech?," *J. Acoust. Soc. Am.* **55**, 822–832.
- Kluender, K. R. (1991). "Effects of first formant onset properties on voicing judgments result from processes not specific to humans," *J. Acoust. Soc. Am.* **90**, 83–96.
- Kuhl, P. K. (1987). "The special-mechanisms debate in speech research: Categorization tests on animals and infants," in *Categorical Perception: The Groundwork of Cognition*, edited by S. Harnad (Cambridge U.P., Cambridge, England), pp. 355–386.
- Kuhl, P. K. (1989). "On babies, birds, modules, and mechanisms: A comparative approach to the acquisition of vocal communication," in *The Comparative Psychology of Audition*, edited by R. J. Dooling and S. H. Hulse (Lawrence Erlbaum, Hillsdale, NJ), pp. 379–419.
- Kuhl, P. K., and Miller, J. D. (1975). "Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants," *Science* **190**, 69–72.
- Kuhl, P. K., and Padden, D. M. (1982). "Enhanced discrimination at the phonetic boundaries for the voicing feature in macaques," *Percept. Psychophys.* **32**, 542–550.
- Macmillan, N. A., Goldberg, R. F., and Braida, L. D. (1988). "Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua," *J. Acoust. Soc. Am.* **84**, 1262–1280.
- Manley, G. A. (1990). *Peripheral Hearing Mechanisms in Reptiles and Birds* (Springer-Verlag, Berlin).
- Manley, G. A., and Gleich, O. (1992). "Evolution and specialization of function in the avian auditory periphery," in *The Evolutionary Biology of Hearing*, edited by D. B. Webster, R. R. Fay, and A. N. Popper (Springer-Verlag, New York), pp. 561–580.
- Miller, J. L., and Liberman, A. M. (1979). "Some effects of later-occurring information on the perception of stop consonant and semivowel," *Percept. Psychophys.* **25**, 457–465.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and Fujimura, O. (1975). "An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English," *Percept. Psychophys.* **18**, 331–340.
- Moore, B. R. (1992). "Avian movement imitation and a new form of mimicry: Tracing the evolution of a complex form of learning," *Behav.* **122**, 231–263.
- Neuweiler, G. (1980). "Auditory processing of echoes: Peripheral processing," in *Animal Sonar Systems*, edited by R. J. Busnel and J. F. Fish (Plenum, New York), pp. 519–548.
- Okanoya, K., and Dooling, R. J. (1987). "Hearing in passerine and psittacine birds: A comparative study of absolute and masked auditory thresholds," *J. Comp. Psych.* **101**, 7–15.
- Okanoya, K., and Dooling, R. J. (1991). "Perception of distance calls by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*): Assessing species-specific advantages," *J. Comp. Psych.* **105**, 60–72.
- Patterson, D. K., and Pepperberg, I. M. (1994). "A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels," *J. Acoust. Soc. Am.* **96**, 635–648.
- Pepperberg, I. M. (1990). "Some cognitive capacities of an African Gray parrot," in *Advances in the Study of Behavior*, edited by P. J. B. Slater, J. S. Rosenblatt, and C. Beer (Academic, New York), pp. 357–409.
- Pisoni, D. B., Carrell, T. D., and Gans, S. J. (1983). "Perception of the duration of rapid spectrum changes in speech and nonspeech signals," *Percept. Psychophys.* **34**, 314–322.
- Ploog, D. W. (1992). "Evolution of vocal communication," in *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, edited by H. Papoušek, U. Jürgens, and M. Papoušek (Cambridge U.P., New York), pp. 6–30.
- Pollak, G. D. (1980). "Organizational and encoding features of single neurons in the inferior colliculus of bats," in *Animal Sonar Systems*, edited by R. J. Busnel and J. F. Fish (Plenum, New York), pp. 549–587.
- Repp, B. H. (1982). "Perceptual integration and differentiation of spectral cues for intervocalic stop consonants," *Percept. Psychophys.* **24**, 471–485.
- Repp, B. H., Liberman, A. M., Eccardt, T., and Pesetsky, D. (1978). "Perceptual integration of acoustic cues for stop, fricative, and affricative manner," *J. Exp. Psychol. Hum. Percept. Perform.* **4**, 621–637.
- Ryals, B. M., and Rubel, E. W. (1988). "Hair cell regeneration after acoustic trauma in adult Corturnix quail," *Science* **240**, 1774–1776.
- Schnitzler, H. U. (1973). "Control of doppler shift compensation in the Greater Horseshoe bat, *Rhinolophus ferrumequinum*," *J. Comp. Physiol.* **82**, 79–92.
- Sinnott, J. M., Brown, C. H., and Borneman, M. A. (submitted). "Effects of syllable duration on stop-glide identification in syllable-initial and syllable-final position by humans and monkeys."
- Striedter, G. S. (1994). "The vocal control pathways in budgerigars differ from those in songbirds," *J. Comp. Neurol.* **343**, 35–56.
- Thorpe, W. H. (1959). "Talking birds and the mode of action of the vocal apparatus of birds," *Proc. Zool. Lond.* **132**, 441–455.
- Turney, S. M., Banta, P. A., and Pepperberg, I. M. (1994). "Comparative acoustical analyses of learned English vocalization of two parrot species," *Anim. Beh. Soc. Abstr.* 140.
- Warren, D. K., Patterson, D. K., and Pepperberg, I. M. (1996). "Mechanisms of American English vowel production in a grey parrot (*Psittacus erithacus*)," *The Auk* **113**, 41–58.