*The Auk* 113(1):41-58, 1996

# MECHANISMS OF AMERICAN ENGLISH VOWEL PRODUCTION IN A GREY PARROT (*PSITTACUS ERITHACUS*)

DENICE K. WARREN,[1,3] DIANNE K. PATTERSON,[2] AND
IRENE M. PEPPERBERG[1,2]
[1]*Department of Ecology and Evolutionary Biology and*
[2]*Department of Psychology,*
*University of Arizona, Tucson, Arizona 85721, USA*

ABSTRACT.—Parrots are known for their capacity to reproduce the sounds of human speech, but few studies have investigated their mechanisms for producing such vocalizations. Using three methods of noninvasive video imaging (Super VHS video, infrared, X-ray radiography), we examined correlations between several elements of a Grey Parrot's (*Psittacus erithacus*) vocal-tract configuration and the bird's production of two vowels, /i/ ("*eat*") and /a/ ("r*o*ck"). This parrot uses mechanisms that both resemble and differ from those used by: (a) humans to produce English speech; and (b) some avian species to produce conspecific vocalizations. This Grey Parrot, for example, uses its vocal apparatus in some but not all of the ways used by humans to produce vowels. Although our Grey Parrot, like some oscines, appears to use beak opening to modify the characteristics of sound, the specific sound-modification properties of beak opening may or may not be similar to those of nonpsittacids. We describe a first-order model that, although simplistic, may aid our understanding of the mechanisms that enable a parrot to reproduce human speech. *Received 5 December 1994, accepted 19 June 1995.*

ALTHOUGH INSTANCES of allospecific utterances are not uncommon for birds that learn their vocalizations, only a small subset (e.g. corvids, stringillids, cacatuids, psittacids) of two (Passeriformes, Psittaciformes) of the 28 orders of birds are known to reproduce the sounds of human speech. Little is yet understood about the mechanisms of such avian vocal production. Whether mechanisms differ among mimetic species (e.g. Nottebohm 1976, Brackenbury 1982, 1989, Gaunt and Gaunt 1985) is unknown, and only a few studies have compared acoustic and articulatory aspects of avian and human productions (Klatt and Stefanski 1974, Nottebohm 1976, Scanlan 1988, Patterson and Pepperberg 1994).

Specifically, although Homberger (1986) and Nottebohm (1976) published detailed descriptions of the Grey Parrot (*Psittacus erithacus*) lingual apparatus and the Orange-winged Amazon (*Amazona amazonica*) syringeal anatomy, respectively, researchers have not provided correlations between vocal behavior and anatomical data that are adequate for making detailed comparisons between avian and human speech mechanisms. In some cases, the limiting factor was the small lexicon of the avian subject. Scanlan (1988), for example, was able to obtain cineradiographic data for the production of only four instances of an isolated /a/ from a hybrid Amazon parrot (species not provided) and three instances of the name "Coco" from a Grey Parrot.

Our purpose, therefore, was to obtain sufficient data in order to determine which physical structures are used and how they are employed in the production of recognizable psittacine speech, specifically vowels. Our goal was feasible because our experimental subject, a Grey Parrot named Alex, uses English speech to identify, comment upon, refuse, categorize, quantify, or request over 100 different objects, and produces vocal labels for colors, shapes, materials, numbers and categories (Pepperberg 1990a, b, c). Moreover, researchers have analyzed the acoustic characteristics of all the vowels (Patterson and Pepperberg 1994) and many consonants (Patterson and Pepperberg in prep) of English speech that are produced by this parrot in a variety of phonological contexts. Such analyses thus can be correlated with the articulatory data we obtained using three methods of noninvasive video imaging (Super VHS video [SVHS], infrared, and X-ray radiography) in or-

---

[3] Present address: Interpretel, 5210 E. Williams Blvd, Tucson, Arizona 85711, USA.

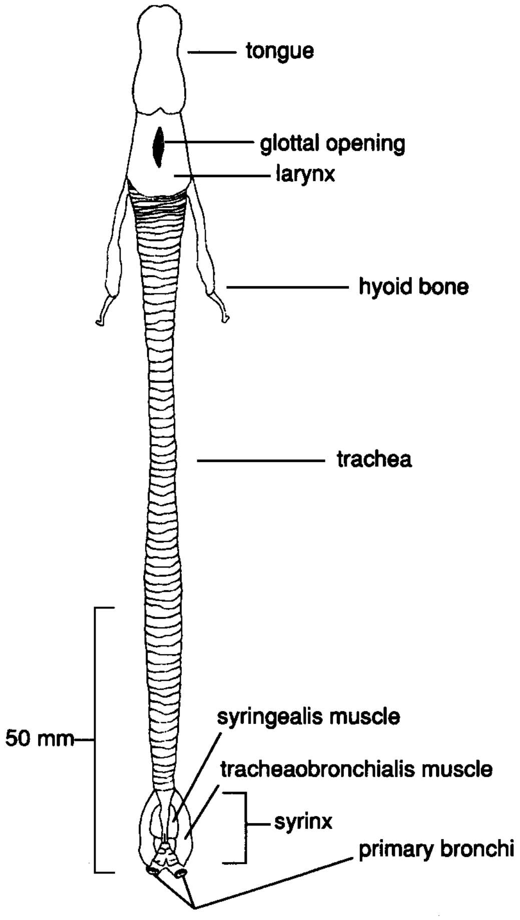

Fig. 1. Dorsal view of Grey Parrot vocal tract.

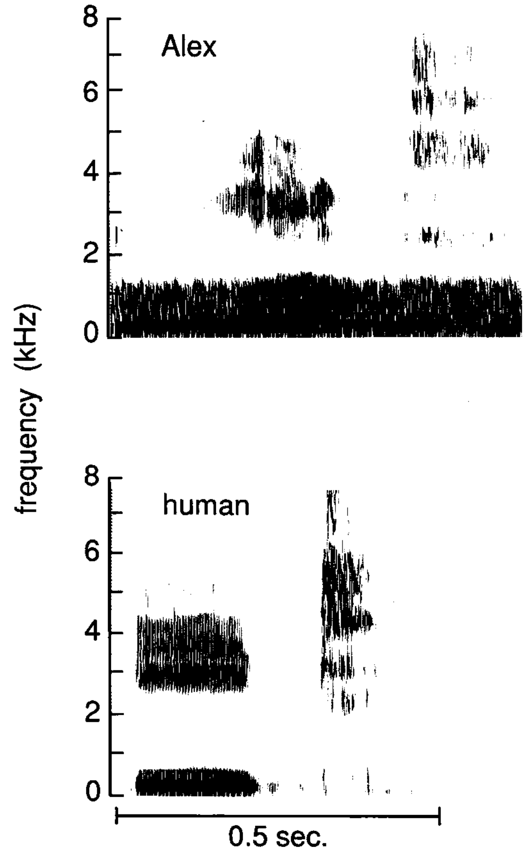

Fig. 2. Spectrograms of the word "eat" (/it/) produced by Grey Parrot (Alex) and human female.

der to develop a preliminary model for the mechanisms of vowel production in this animal.

## ANATOMICAL STRUCTURES USED FOR SPEECH PRODUCTION

### OVERVIEW

Speech production in psittacids, as in humans, is a complex process involving many structures that must be configured differently for each sound. It is generally accepted that, in the avian vocal system, sound is produced in the syrinx (Greenewalt 1968). The extent to which suprasyringeal structures (Fig. 1) such as the trachea, larynx, tongue, and both upper and lower mandibles serve to modify the resonant properties of the vocal tract is still under debate. In contrast, speech in the human vocal system is produced by the vibratory elements of the larynx upon exhalation or, occasionally, inhalation and is known to be modified by the supralaryngeal resonating chambers (Fant 1970, Olive et al. 1993). Structures that are likely involved in this modification are the pharyngeal walls, tongue, tongue root, velum, sinuses, teeth, and lips.

Despite parrots' obvious lack of dentition and lips, and the dramatic differences between the morphology of other human and avian vocal structures, the acoustic characteristics of avian speech may closely resemble that of humans (Klatt and Stefanski 1974, Patterson and Pepperberg 1994). Spectrograms of a word ("eat") produced by a Grey Parrot and a human, for example (Fig. 2 a, b), have striking similarities, although the bird's formants in general are less distinct. Note that these are true formants, not harmonics (Patterson and Pepperberg 1994). Harmonics are integer multiples of the source

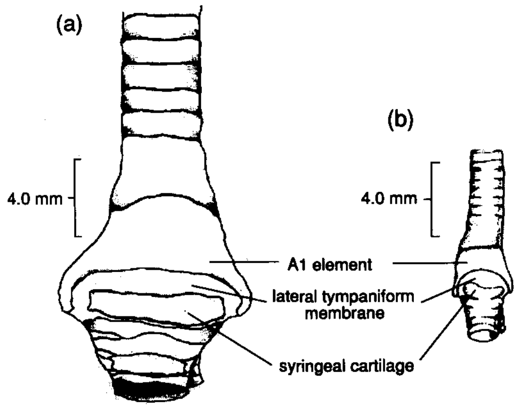

Fig. 3. Cross-sectional view of syrinx of (a) Grey Parrot (drawing by A. Warren) and (b) Budgerigar (*Melopsittacus undulatus*; from King 1989).

vibration (i.e. of the fundamental frequency, $F_0$ of vocal folds of the human larynx or parrot's syringeal membranes); formants (e.g. first and second formant frequencies, $F_1$ and $F_2$) may or may not be linearly related to source frequency, but primarily are a function of the natural resonance of particular configurations of human supralaryngeal or psittacine suprasyringeal tracts. How, then, do such different physical apparatus produce comparable vocalizations? To answer that question, we describe avian vocal systems, with an emphasis on that of the Grey Parrot.

## ROLE OF SYRINX

*Anatomical structure.*—Psittacine syringes, like those of many other birds, are classified as tracheobronchial if they comprise the posterior end of the trachea at its bifurcation and the cranial portions of the bronchi (King 1989) or tracheal if the medial tympaniform membrane is missing or nonfunctional (Gaunt and Gaunt 1985). Unlike many songbirds (e.g. the mimetic Greater Indian Hill Mynah, *Gracula religiosa*), however, several parrot species that have been examined appear to have only one site for sound production within the syrinx (Gaunt and Gaunt 1985, Nottebohm 1976, Patterson and Pepperberg 1994). Moreover, according to Scanlan (1988:140), "The shape and position of the syringeal cartilages in the grey parrot (*Psittacus erithacus*) differ from those of other psittacine species . . . . the dorso-ventrally oriented cranial edges of the syringeal cartilages are straight, not semi-oval in outline . . . . Also, the syringeal cartilages are positioned further craniad than in all other parrot species studied" (Fig. 3 a, b). As a result, he postulated that these modifications may "facilitate control of the intra-syringeal aperture," perhaps precisely controlling contact between the two opposing lateral tympaniform membranes and tightly coupling movements of the cartilages and the membranes. The acoustic effect of these morphological adaptations could be more periodic sound production and greater frequency control in Grey Parrots than in other psittacine species.

*Syrinx as a frequency modulator.*—Several researchers have proposed that all or most frequency modulation can be performed by the syrinx, and that the resonant properties of the rest of the avian vocal tract play little or no part in the modification of sound (Greenewalt 1968). Scanlan (1988), who discussed the roles of several anatomical structures in psittacine speech, noted that one of the syringeal constriction mechanisms in parrots (that involving the lateral tympaniform membranes) functionally resembles that of the human vocal folds in phonation (see also Gaunt and Gaunt 1985); he also, however, discussed the roles of suprasyringeal structures (see below). Interestingly, the degree of the syringeal complexity across avian species is not directly correlated with the complexity of their vocal productions, with some syringeal complexity being necessary but not sufficient for vocal plasticity (Gaunt 1983). For species with relatively simple syringes but complex vocal behavior (like the Grey Parrot), such data imply that other structures must be involved in the modification of syringeal output. Gaunt (1983) and Stein (1968) suggested that vocal plasticity arises in part from neurological adaptations. Although parrots apparently have a complex neurologic system for vocal control (Streidter 1994), we propose that suprasyringeal structures also play a major role.

## ROLE OF SUPRASYRINGEAL STRUCTURES

We consider the possibility that vocal abilities may have arisen through several anatomical substrates, such that the syrinx would work in concert with the elements of the suprasyringeal vocal tract. Many researchers have suggested that suprasyringeal resonating chambers play a considerable role in vocalizations, although most data are for nonpsittacids. Nowicki (1987) argued that the frequency shift observed when oscine whistled songs are produced in a helium atmosphere (e.g. by Song Sparrows [*Melospiza melodia*] and Black-capped Chickadees [*Parus atricapillus*]) demonstrates the effect of suprasyringeal resonance on the output. Correlational evidence also suggests that suprasyringeal structures are capable of modulating frequencies in a number of other species: domestic chickens (*Gallus domesticus;* Myers 1917, Harris et al. 1968, White 1968); Ross' Geese (*Chen rossii*) and Snow Geese (*C. hyperborea hyperborea;* Sutherland and McChesney 1965); Barnacle Geese (*Branta leucopsis;* Hausberger et al. 1991); White-throated Sparrows (*Zonotrichia albicollis*) and Swamp Sparrows (*Melospiza georgiana;* Westneat et al. 1993). In the following sections we describe the major suprasyringeal structures and review the existing correlational evidence for their role in sound modification.

*Trachea.*—The Grey Parrot trachea consists of a series of ossified, complete rings, with minimal intervals between them. These rings can overlap, allowing the trachea to change in length or configuration (Fig. 4). An excised (possibly slightly stretched) Grey Parrot trachea measures approximately 11 cm (Patterson unpubl. data), but the range of length in a talking bird is not known (for critique of measurements taken from formaldehyde-fixed tissue samples, see Hersch 1966).

The nature of the trachea's role in sound modification is debatable. Greenewalt (1968:87), based on research on Song Sparrows, argued that "at least for whistled song, the trachea does not significantly modulate the sounds produced at the source." Other authors, however, have suggested that the trachea may serve as a resonant chamber and modulate frequencies by altering—or being altered in—its length or configuration; such alteration could modify the effective length of the vocal tract (Stein 1968). Several investigators, in fact, have noted apparent changes in tracheal length during vocalization, although most of the data are for nonpsittacids. As early as 1917, Myers noted that artificially shortening the trachea of hens raised the pitch of their calls. In a more exhaustive version of Myers' study, Harris et al. (1968) determined that harmonics of artificially produced sound in the domestic fowl varied according to tracheal length. They hypothesized that the trachea, beak, and oral cavity of the chicken may "tune the sound of vocalization . . . to a resonant frequency, which causes the pitch to be more sharply defined," and concluded that the trachea and primary bronchi combine to form a single resonant tube (p. 112). The elongated tracheal morphology in birds-of-paradise (Paradisaeidae) may "serve to lower the pitch of, and perhaps amplify, their vocalizations" (Clench 1978:428; see also Frith 1994). Warner (1971) believed that one of the functions of the extrinsic syringeal muscles in ducks (*Anas, Aythya*) was to vary the length and, thus, the resonant characteristics of the trachea. Sutherland and McChesney (1965) suggested that tracheal resonance plays a role in modulating the calls of Ross' and Snow geese. Brackenbury (1978), through an acoustic analysis, surmised that sound pulses in the Grasshopper Warbler (*Locusta naeiva*) exhibit features characteristic of a pulsed tracheal resonator. In the one study on psittacids (a Grey Parrot), Scanlan (1988) suggested that the tracheal protraction observed in an *X*-ray film of a parrot's production of the vowel /o/ (in "Coco") also was due to lengthening of the trachea.

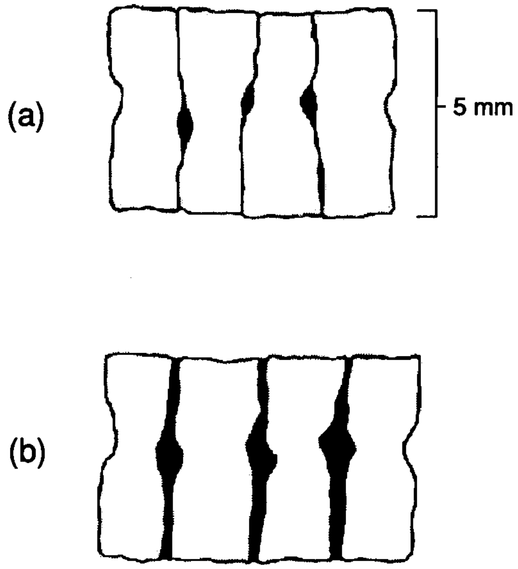

Fig. 4. Dorsal view of Grey Parrot trachea (a) compressed and (b) stretched. Note overlap of tracheal rings in upper panel.

*Larynx.*—Homberger's (1979) detailed study of the larynx of a Grey Parrot concentrated on anatomy and demonstrated that "The extrinsic musculature moves the larynx with respect to the hyoid skeleton" and that "laryngeal action is influenced by the position of the hyoid skel-

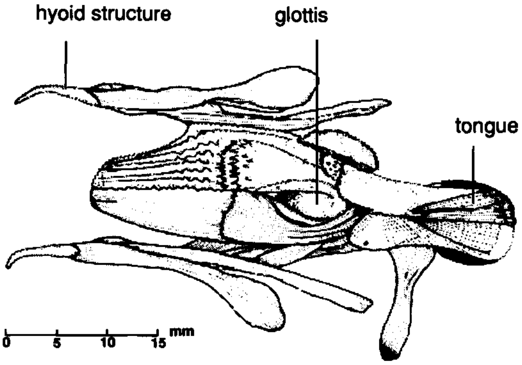

Fig. 5. Dorsal view of Grey Parrot lingual apparatus (from Homberger 1986) with hyoid structure and glottis marked. In *X*-ray in subsequent figures, hyobranchial junction (where hyoid bone of tongue meets larynx) appears as dark spot on ventral side of larynx.

eton with respect to the lower jaw" (p. 988). She found that this psittacine larynx differs anatomically from that of *Corvus* (Bock 1978) and *Gallus* (White 1975), but did not clarify the effect of such differences on vocal production.

The avian larynx, unlike that of humans, does not appear to have any vibratory elements capable of producing sound (McClelland 1989). The avian larynx may be used instead to modify sound produced by the syrinx. In that capacity, the larynx could constrict the glottis to varying degrees, changing the resonant properties of the vocal tract to achieve a target vocalization. White (1968) noted a laryngeal descent during the crowing of chickens, as well as changes in laryngeal configuration. Nottebohm (1976) suggested that the acoustic resonating properties of the nasopharyngeal and buccal cavities and the tracheal tube of the Orange-winged Amazon could be altered by the rostrocaudal movements of the larynx. Scanlan (1988) noted two types of movement in laryngeal configuration in a vocalizing Grey Parrot: preparatory movements that transport the larynx to a vocalizing position; and synchronic movements that occur during vocalization.

*Tongue.*—The skeleton and musculature of the psittacine tongue is unique among birds (Burton 1974). The Grey Parrot tongue is supported by the bony hyoid apparatus, and the glottal opening can be found in the larynx just dorsal to the tongue (Fig. 5; Homberger 1986). Three joints (true diarthroses) in the hyoid apparatus together with six extrinsic and seven intrinsic pairs of lingual muscles can affect motion of the tongue (Homberger 1986). Although likely to have evolved for the purposes of eating, this unique structure allows particularly flexible laryngeal movements within the oro-pharyngeal cavity (Homberger 1986, Scanlan 1988). That is, certain movements of the tongue move the larynx and trachea and, likewise, movement of the trachea may move the tongue. Nottebohm (1976) suggested that parrots may use their tongues to modify the shape and, thus, the resonant properties of the vocal tract. The tongue, its extrinsic musculature, and the hyoid apparatus generally are referred to as the lingual apparatus.

*Upper and lower mandibles.*—The psittacine jaw apparatus is characterized by hinged upper and lower mandibles, with the lower mandible having a wide range of motion. Beak movement during vocalization may serve not only as a visual display, but also as a means to modify the characteristics of the sound, namely amplitude and frequency (Westneat et al. 1993). Increases in beak opening or gape have been positively correlated with higher call frequencies in geese (Hausberger et al. 1991). White-throated Sparrows and Swamp Sparrows may use beak gape to change the effective length of the vocal tract and, thus, track the fundamental frequency produced by the syrinx (Nowicki et al. unpubl. data, Westneat et al. 1993). Data from two experiments, although at first glance in conflict, support this proposal. In one experiment Westneat et al. (1993) found that the extent of beak gape was directly correlated with frequency and not with amplitude. When sparrows' beaks were temporarily immobilized, however, the acoustic frequencies of their song elements remained unchanged, but changes in relative amplitudes occurred (Nowicki et al. unpubl. data). Thus, the function of beak opening is not to modify amplitude independent of frequency (i.e. by projection), but rather to track (and change with) frequency so as to maximize the amplitude for some of the frequencies produced (e.g. to maintain adequate amplitude for sounds of particular frequencies). Similar mechanisms must be considered for parrots.

*Nasal cavity.*—Although little information exists regarding the morphology of psittacine nasal cavities, the nasal cavity also may affect the quality of psittacine vocalization (Nottebohm 1976). Air moves between the pharyngeal and nasal cavities via the choana, which is framed by the paired fleshy choanal folds (Homberger 1980, in press). Thus, the boundary between these two cavities consists of soft tissue and the elasticity of this material is likely to affect how sound energy is transferred from the pharyngeal to the nasal cavities.

## ARTICULATORY PREDICTIONS BASED ON ANATOMICAL AND ACOUSTICAL DATA

*Background.*—Many researchers have designed models of avian vocal mechanisms (e.g. Fletcher 1988). These models are based on analogies either with wind instruments (e.g. Myers 1917) or the human voice (e.g. Nowicki and Marler 1988). In the model based on a wind instrument, the vocal tract is assumed to be a single tube that can be opened at both ends (syrinx and glottis) or only one end (glottis; see Brackenbury 1982); vocal resonances are tightly coupled to the source. In the model based on the human voice, the vocal tract is assumed to have at least two resonating chambers (Olive et al. 1993); vocal resonances need not be coupled to the source.

Most researchers have concentrated on nonpsittacids. Thorpe (1959, 1961) incorporated suggestions from prior studies (e.g. Rüppell 1933) and proposed that application of one or the other model depended upon whether a species did or did not, respectively, have intrinsic syringeal muscles. Greenewalt (1968), however, favored the wind-instrument approach, even for the production of speech sounds by mynahs. Harris et al. (1968), in contrast, proposed a model that considered the trachea to be a cylindrical tube, the pharynx and buccal area an expandable chamber, the syrinx an acoustic oscillator, and the glottis and beak as variable slits. Although their model actually compares the vocal tract to an electrical circuit, these researchers' use of multiple components corresponds to models proposed for humans.

The few researchers who have proposed models for psittacid speech have concentrated on the single-tube approach. Nottebohm (1976: 1633), for example, accounts for the effect of tracheal resonance by assuming that the tube could be either opened or closed at one or both ends, or that "the glottal aperture of the larynx could vary from (a) fully open to (b), nearly closed," such that one end of the tube could assume a variable opening. We believe, however, that a multiple-component model (somewhat along the lines of Harris et al. 1968) is most appropriate. In the following sections we describe the generally accepted human model, explain how psittacine and human structures might be functionally analogous and, taking these analogies into account, propose a model based on, but simpler than, that of Harris et al. (1968).

*Acoustic model based on anatomy.*—A two-chamber model, which includes the oral cavity and pharynx, was one of the earliest-accepted, first-order models for human speech production (Stevens and House 1961); the model excludes the nasal cavity from consideration. More complicated models have been shown to provide little additional advantage over the two-chamber model with respect to prediction of formant frequencies of vowels (Maeda 1991). The simple two-chamber model is consistent with the generally accepted notion that each vowel sound is uniquely defined by its two lowest formant frequencies, $F_1$ and $F_2$. $F_1$ roughly corresponds to tongue height, and $F_2$ to tongue placement with respect to front and back of the oral cavity (Remez et al. 1987). The model proposes that each formant results from resonance in a different chamber along the vocal tract. In a simple two-chamber model, the chambers in the vocal tract are divided by the tongue, one chamber being the pharyngeal cavity and the other the oral cavity. Placement of the tongue changes with different vowel sounds, thus modifying relative and absolute attributes of both chambers and creating unique resonant characteristics for each vowel (Neary 1978). According to various researchers (e.g. Fant 1970, Lieberman 1984), the abrupt area-function discontinuities created by the tongue in the two-chamber model are required for vowel production and cannot be generated in a single-tube system.

Given the acoustic similarities of vowels between humans and our psittacine subject (Patterson and Pepperberg 1994), we hypothesize that a two-chamber model is appropriate for the psittacine speech system. In such a model, reminiscent of that of Harris et al. (1968), sound produced by the syrinx travels up the trachea and most likely encounters a change in impedance where the trachea (which has a small diameter) meets the oro-pharyngeal cavity (which has a large diameter). One potential point for such constriction is the glottis. The change in impedance should cause resonance in the tra-

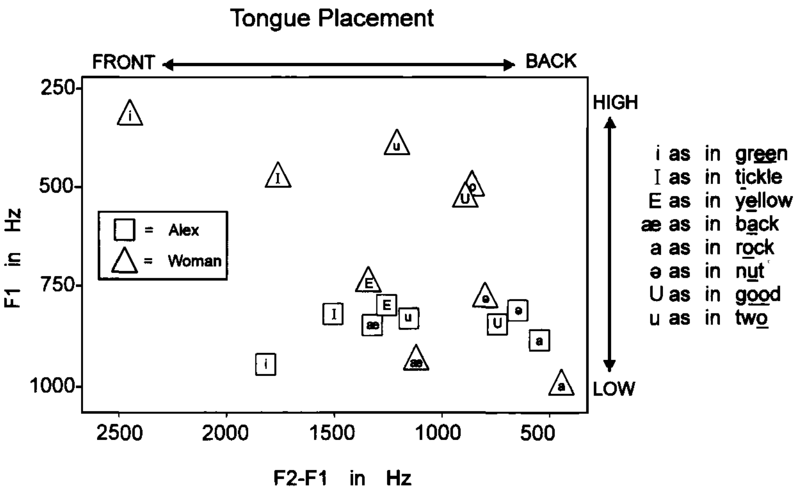

Fig. 6. Predicted tongue placement based on acoustic parameters. International Phonetic Alphabet (IPA; see Pullum and Ladusaw 1986) used throughout.

chea. When the sound exits the oro-pharyngeal cavity and enters what is essentially an anechoic arena, it once again must encounter a change in impedance, causing resonance in an oro-pharyngeal cavity. Hence, there are two major anatomical candidates for resonant chambers in the psittacine vocal tract—the trachea and the oro-pharyngeal cavity.

*Acoustic model based on anatomy and vowel formant frequencies.*—Because the tongue plays an essential role in creating a two-chamber effect in the human vocal tract, tongue placement is highly correlated with a vowel's formant frequencies (Lieberman 1984). For this reason, it is instructive to characterize vowels based on where they fall on a tongue-placement chart (Fig. 6; Borden and Harris 1984). In such charts, $F_1$ (the lowest formant frequency) is plotted against $F_2 - F_1$ (the difference between the two lowest formants). Distribution of vowels along the $y$-axis is correlated with tongue height, whereas distribution along the $x$-axis is correlated with tongue frontedness. Human vowels tend to fall into four categories: high front, high back, low front, and low back. Such a representation makes it possible to predict tongue placement from acoustic information.

Patterson and Pepperberg (1994) similarly characterized the speech produced by a Grey Parrot, Alex (Fig. 6). They determined that Alex's vowels can be classified into "front/back," but not "high/low," categories of tongue placement (i.e. $F_2$, but not $F_1$, varies significantly across vowels). Although this scheme of categorization is specific to human vocal anatomy, the presence of front/back categories in psittacine speech is nevertheless of predictive value.

If we assume a simple two-chamber model of Alex's speech production, the lower frequency formant ($F_1$) should be due to resonance in the longest chamber, and the higher formant ($F_2$) to resonance in the shortest chamber. It follows that, because $F_1$ varies little across Alex's vowels, we would expect $F_1$ to be correlated with a long tube that keeps a relatively constant length. Furthermore, because $F_2$ differs significantly among his vowels, we should find $F_2$ correlated with a shorter tube that quickly and substantially can change in length. Vowels classified as being "fronted" would have shorter $F_2$ resonating chambers and, hence, higher $F_2$ frequencies than "back" vowels.

The obvious candidate for $F_1$ production is the trachea (from syrinx to larynx) because of its length. The trachea could also be a candidate for $F_2$ production because it is capable of changing its length (and thus modulating its resonant frequency) by decreasing the amount of tracheal ring overlap. However, Fletcher (1988) has suggested that this change in length would be slight and likely not result in the type of frequency shift required to produce the variation seen in $F_2$ across vowels.

In contrast, the oro-pharyngeal cavity is ideal for producing the frequency shifts seen in $F_2$. Opening the beak could effectively shorten the resonating chamber and, thus, increase its resonant frequency. Back-front movement of the larynx could also alter the resonant frequency of the oro-pharyngeal cavity by changing the dorsoventral position of the glottal opening. In accordance with this model, the glottis would be a possible place of maximum constriction along the vocal tract and, therefore, the likely divider between the two chambers. Structures such as the tongue also may be responsible for division between the chambers in parrots; such is the case for humans, who accomplish constriction at various points in the vocal tract by using structures such as the tongue, lips, and glottis separately or in combination (Ladefoged 1982).

We find evidence to support this two-chamber model of psittacine speech production by analyzing, via angular and spatial measurements, vocal-tract configurations during vowel production. We used three methods of noninvasive imaging in an attempt not only to test the two-chamber model, but also to describe a number of factors contributing to psittacine vowel production. Our goal was to determine the role, if any, of suprasyringeal structures in vowel production, as well as to correlate anatomical data with data on acoustic output.

## METHODS

*Subject and vowel samples.*—Our subject was a 16-year-old Grey Parrot named Alex, who can produce all the vowels of English speech (Patterson and Pepperberg 1994). For this investigation, we queried him about various items and attributes that he was capable of labeling (Pepperberg 1990a) with the intent of eliciting a wide range of sounds in his acoustic repertoire.

The vowels we analyzed in detail were /i/ (as in *eat*) and /a/ (as in rock), generally in the context of a word but occasionally in isolation. Our notation is that of the International Phonetic Alphabet (IPA), which is the standard system used by phoneticians and linguists to represent human speech (Pullum and Ladusaw 1986). We chose /i/ and /a/ because they are "point vowels" (i.e. vowels that, in humans, are most different from one another in terms of both acoustic characteristics and tongue placement; Borden and Harris 1984). We successfully obtained six samples of each vowel sound (two for each word or isolated instance) with the bird in a lateral position. Words or isolated instances that we elicited containing the vowel /i/ were: "green," "ee," and "eat". For the vowel /a/, words we elicited were: "rock," "want," and "pasta". Alex closely imitates the New York/Boston dialect of his principal trainer, I.M.P., for whom the vowels in "rock," "want," and "pasta" are essentially equivalent. Interobserver reliability between D.K.P. and D.K.W. for identification of Alex's vowels in context was 96% (for details, see Patterson and Pepperberg 1994).

Use of vowels in context rather than in isolation (e.g. Scanlan 1988) generally is preferable for two reasons. First, our ability to identify Alex's targeted vowel is more reliable when the vowel is in the context of a referential term that is specifically being elicited. Vowel sounds from a parrot that does not use referential speech may be subject to instability because of sound play—blends, rearrangements, or substitutions (Todt 1975, Pepperberg et al. 1991, Patterson and Pepperberg 1994). Second, use of words facilitates comparisons with human speech: Most studies on humans rely on vowels in context because human formants, "to be identified with certainty, must often be perceived in relation to the frequencies of some other bit of speech uttered from the same vocal tract" (Borden and Harris 1984:194). Any differences we observed between /a/ and /i/, moreover, would not be significantly affected by use of whole words, as only about 10% of the variation in, for example, Alex's $F_2$ across all his vowels was due to phonological context (Patterson and Pepperberg 1994).

*Imaging techniques.*—We used three noninvasive video-imaging techniques: SVHS video, infrared video, and *X*-ray radiography. The SVHS and infrared data provided qualitative information and a context for the more detailed analysis performed on the radiographic data. For all three imaging methods, we attempted to keep our psittacine subject in a position lateral to the camera.

To observe the external movements associated with speech production, we filmed Alex with a Panasonic SVHS AG-450 camera and Maxell XR-S120 SVHS tape at a rate of 30 frames/s. Although this type of video provided an external context for the internal movements observed in the *X*-ray video, lighting difficulties did not allow clear views of the tongue.

In an attempt to determine tongue position during speech production, we videotaped Alex in the Department of Optical Sciences at the University of Arizona with an infrared camera (ImagIR, Santa Barbara Focalplane, Goleta, California) at a rate of 30 frames/s. The imaging system has a sensitivity better than 0.1°C. In this format, his warm tongue was easily visible when not obstructed by his beak.

In order to see and eventually quantify the movement of internal structures during speech production, we *X*-ray videotaped Alex for 22 min at the University of Arizona Medical Center's Department of Radiology. The center's equipment consists of a Toshiba *X*-ray Machine and an Altronics Medical System HRV 3000 EM High Resolution Multiscan Video Recorder. Merlin Engineering Works downscanned the result-

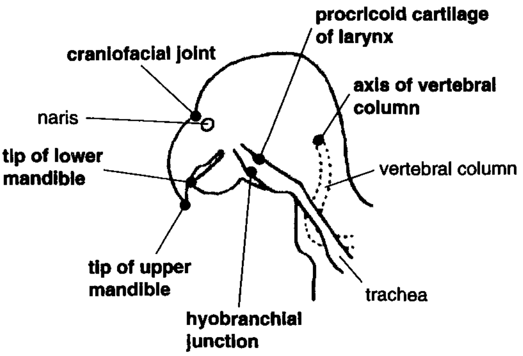

Fig. 7. Anatomical features visible from X-ray of Alex's head and upper vocal tract. Dots are placed on anatomical landmarks used for taking measurements.

ing radiograph from 1,026-line-rate to 512-line-rate format to enable us to view the tape on conventional monitors.

*Digital image processing.*—We digitized video segments at a rate of 30 frames/s using a QuickCapture video-capture card installed in a Macintosh II. Once the data were in a digital format, they could be analyzed on a Macintosh Powerbook 180c computer using the public domain NIH Image program (written by Wayne Rasband, U.S. National Institutes of Health; available from Internet by anonymous ftp from zippy.nimh.nih.gov or on floppy disk from NTIS, 5285 Port Royal Road., Springfield, Virginia 22161, part number PB93-504868).

When assigning acoustic characteristics to a vocal-tract configuration, we identified which sounds were produced (and in which order) in the series of vocal movements. As is standard for studies in humans (e.g. Subtelny et al. 1989), we assumed that a directional change in movement of the vocal structures was correlated with production of the target sound. We then extracted frames containing vowel configurations for further enhancement and analysis.

Digital image processing enabled us to enhance specific anatomical structures. The radiopacity of such structures, which are buried in layers of hard and soft tissue, varies depending on their position in the bird. Consequently, such structures cannot be easily examined on the unprocessed images. We employed different operations, such as sharpening and smoothing, histogram equalization, density slicing, contrast manipulation, image magnification, and image subtraction to enhance each area of interest. For example, sharpening the image generally enhanced the portion of the trachea caudal to the larynx and occupying the jugulum in the neck rostral to the interscapular region, but obscured that portion just caudal to the larynx. Density slicing, however, was useful for visualizing the portion of the trachea obscured by sharpening.

*Analysis.*—Based on the procedures described by Subtelny et al. (1989), we identified visible structures in the X-ray video (e.g. vertebral column, trachea, mandibles) and placed six small dots onto predetermined landmarks on these structures (e.g. axis of vertebral column, hyobranchial junction, procricoid of larynx, craniofacial joint [hinge on upper mandible], and both bill tips; see black dots on Fig. 7). The axis of the vertebral column had the advantage of being the only point dorsal to the trachea that could be identified consistently. The hyobranchial junction is where the hyoid bone of the tongue intersects with the larynx. The hyobranchial junction and the procricoid cartilage, which can be seen on the X-ray, are located respectively on the ventral and dorsal surfaces of the larynx. The visibility of the procricoid and hyobranchial junction and their placement directly opposite one another make these structures important markers. The points on the mandibles were of critical importance for obtaining measurements of head tilt and beak gape.

Having tagged the key structures, we could then place reference lines to enable us to measure the image (Fig. 8). We created a horizontal reference line parallel to the flattest and most cranial portion of the skull. We then created two vertical reference lines perpendicular to the first, with one intersecting the hinge on the upper mandible and the other the vertebral axis. Next, we took a series of angular (Fig. 8a, b) and spatial (Fig. 8c) measurements designed to characterize variation in vocal tract configurations during speech production.

We used several statistical tests to analyze our data. We used a Tukey test (SAS Institute 1989) to determine which measurements differed significantly between the vowels /i/ and /a/. We used GLMs (general linear models; SAS Institute 1989) to learn how much variation in each measurement could be explained by vowel versus word, and to learn if vowel could be used to predict motions in the suprasyringeal tract. The results of the GLMs provide more sophisticated information than do Tukey tests. Finally, we used a correlation matrix to identify measurements that covaried significantly for each vowel. We wanted to determine which, if any, measurements provided distinct information and whether measurements that would be expected to be related were indeed correlated. Moreover, correlations that were not obvious from, for example, inspection of the X-ray data might provide information about additional functional relationships during speech acts.

Although the respiratory system, syrinx, trachea, larynx, glottis, nasal cavity, and mandibles all play some role, direct or indirect, in psittacine speech production, the nature of the imaging techniques we used allowed only for direct measurement of the position of the mandibles, hyobranchial junction, and procricoid of the larynx. The relative position and configuration of the trachea could be determined, but

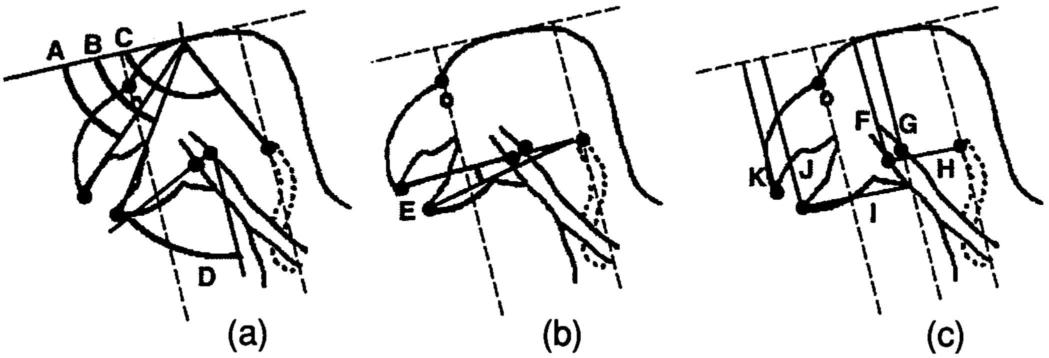

Fig. 8. Depiction of (a and b) angular and (c) spatial measurements used to characterize Alex's vocal tract configurations during his production of /a/ and /i/. A is angle of upper mandible to the horizontal; B is angle of lower mandible to the horizontal; C is angle of head tilting; D is angle of hyobranchial junction to procricoid with respect to the vertical; E is angle of beak gape, which can be a negative value if beak tips overlap; F is vertical position of hyobranchial junction; G is vertical position of procricoid of larynx; H is horizontal position of hyobranchial junction; I is horizontal position of lower mandible; J is vertical position of lower beak tip. (E shown separately from A–D for clarity).

no landmarks caudal to the larynx could be tracked consistently. The position of the tongue can be inferred from the known position of the hyobranchial junction, but the tongue could rarely be viewed during vowel production, even in the infrared video.

## RESULTS

*Articulatory configurations.*—Figure 9 shows the articulatory configurations associated with /a/ and /i/. In the SVHS image (Fig. 9a), note the differences in beak gape and the protracted area below the lower mandible. The infrared images (Fig. 9b) show no evidence of the tongue in a high front position relative to the beak. Were the tongue in such a position, it would be visible as a light shade of grey in infrared. In the *X*-ray images (Fig. 9c), the trachea, the hyobranchial junction (where the hyoid bone of the tongue intersects with the larynx), and the procricoid cartilage of the larynx are highlighted. The vertebral column, beak, and skull also are evident.

Of particular interest are the differences in tracheal configuration during production of /a/ and /i/ (Fig. 9). Although we were unable to extract exact measurements for tracheal length and protraction, the images suggest that, despite the obvious protraction for /a/, there was little change in length. Note that the trachea lies to one side of the medial axis of the bird, so that, depending upon our view, a protracted trachea could be foreshortened or otherwise altered in perspective.

*Statistical results.*—As can be seen from the results of the Tukey tests (Figs. 10a, b), only a subset of measurements differ significantly between vowels. Differences between vowels are significant ($P < 0.05$, df $= 6$) in the angular position of the lower mandible and hyobranchial junction, the overall beak gape, the tilt of the head, protraction of the hyobranchial junction, and the position of the lower beak tip.

GLMs, which are less likely than Tukey tests to attribute significance to a marginal factor (Cohen and Cohen 1983), confirm that vowels are highly correlated to the same six measures shown to differ significantly in the Tukey tests. Although correlation coefficients (discussed below) report the extent to which two variables are positively or negatively correlated, regression coefficients (from GLMs), which are the squares of the correlation coefficients, indicate how much one variable changes as a function of the other (Schroeder et al. 1986). The GLMs, unlike Tukey tests, therefore, indicate the extent to which vowel accounts for the variance in each measurement (see Table 1). Thus, for example, GLMs show that the relationship between vowel and angle of beak gape (E in Fig. 8b; $R^2 = 0.94$) is stronger than between vowel and angle of head tilting (C in Fig. 8a; $R^2 = 0.61$).

GLMs also help to demonstrate the effects of phonetic context of the vowel on the physical measurements. The context of the vowel is the word in which it is embedded; thus, we performed additional GLMs (word and word+

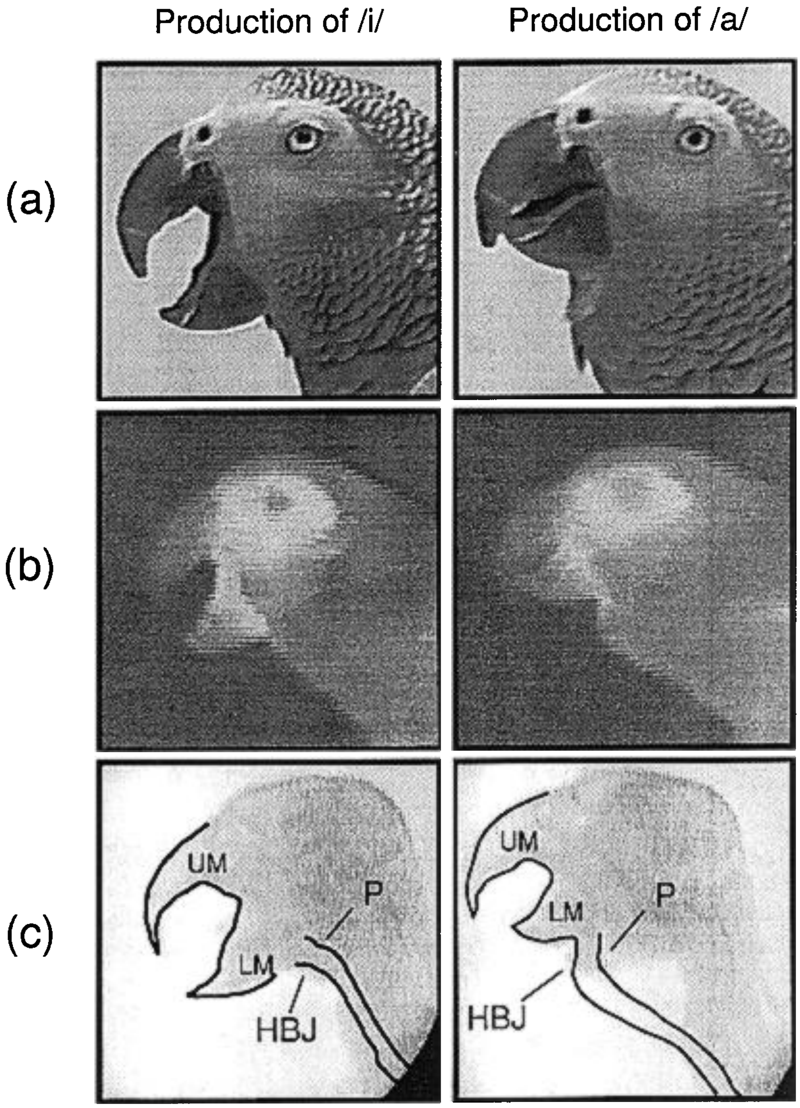

Fig. 9. Alex's articulatory configurations during production of /i/ and /a/. (a) SVHS video; (b) infrared video; (c) *X*-ray video. UM is uppermandable, LM is lower mandable, P is procricoid, and HBJ is hyobranchial junction.

vowel) to determine the extent to which word and word+vowel can account for variance in each measurement. Word and word+vowel provided exactly the same results. Both of these GLMs found that word (or word+vowel) was highly correlated to only five of the six measures discussed above, eliminating angle of head tilt. Note that word accounts more successfully than vowel alone for each of these five remaining measures (Table 1).

We used a correlation matrix to examine the relationships among the various physical measurements, A–K (Fig. 8; see Moore 1992). We found, for example, the expected close correlations ($r > 0.75$) between physically related measurements (Table 2). Other correlations were significant, but not as robust ($r < 0.75$). Moreover, physical measurements could be divided into two groups based on correlational evidence. In group 1, the angle of the lower mandible to the horizontal and the angle of beak gape were closely correlated ($r = 0.88, P < 0.001$),

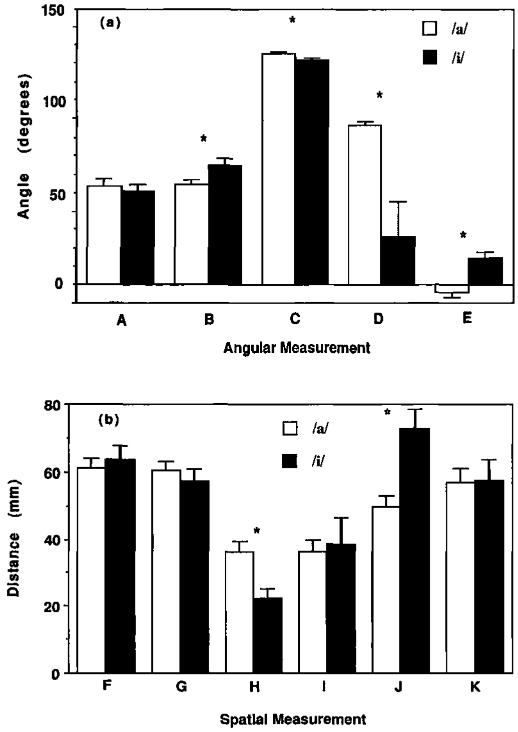

Fig. 10. Means (whiskers indicate SD) of (a) angular and (b) spatial measurements for Alex's production of /a/ and /i/ with statistical significance determined by Tukey tests (*, $P < 0.05$). Capital letters indicate: (A and B) lower mandible opens significantly more during /a/ than /i/; (C) head is tilted back more during /a/ than /i/; (D) procricoid is almost directly above hyobranchial junction in /i/, but is nearly horizontal to junction in /a/; (E) beak is wide open during /i/, but closed during /a/; (F and G) vertical positions of hyobranchial junction and procricoid of larynx do not vary significantly between /i/ and /a/; (H) hyobranchial junction is protracted during /a/ and is positioned medially during /i/; (I) horizontal position of lower mandible does not differ significantly from /i/ to /a/; (J and K) lower beak tip is significantly lower during /a/ than /i/, but upper beak tip does not change its vertical position significantly.

TABLE 1. Results of a GLM to determine extent to which vowel or word (context) accounts for variance in each measurement (letters refer to Fig. 8).

| | $R^2$ | |
|---|---|---|
| Measurement | Vowel | Word (context) |
| B. Angle of the lower mandible to the horizontal | 0.733*** | 0.924** |
| C. Angle of head tilting | 0.607** | — |
| D. Angle of hyobranchial junction to procricoid with respect to the vertical | 0.853*** | 0.947*** |
| E. Angle of beak gape (negative value if beak tips overlap) | 0.941*** | 0.970*** |
| H. Horizontal position of the hyobranchial junction | 0.861*** | 0.950*** |
| J. Vertical position of the lower beak tip | 0.877*** | 0.962*** |

**, $P < 0.005$; ***, $P < 0.001$.

as was the angle of the lower mandible to the horizontal and the vertical position of the lower beak tip (B-E, $r = 0.88$, $P < 0.001$) and, consequently, the angle of beak gape and the vertical position of the lower beak tip (B-J, $r = 0.92$, $P < 0.001$). In group 2, the angle of the hyobranchial junction to the procricoid was closely correlated to the horizontal position of the hyobranchial junction (E-J, $r = 0.93$, $P < 0.001$). The angle of head tilt (C), which decreases as the beak comes up, was significantly negatively correlated with group 1 and positively correlated with group 2. This result shows that the head tilts back as the beak opens.

The division based on correlational evidence suggests that there might be two distinct groups of measurements relevant to distinctions between vowels. However, we also found a consistent pattern of negative correlations between group 1 and group 2 measurements. Thus, for example, the angle of the hyobranchial junction to the procricoid, with respect to the vertical, is negatively correlated to the vertical position of the lower beak tip (D-J, $r = -0.97$, $P < 0.001$).

*Calculations from acoustic and anatomical measures.*—We used previously obtained data on formant values for vowels (Patterson and Pepperberg 1994) and anatomical measures (a dis-

TABLE 2. Pearson product-moment correlations indicating significant associations between measurements. Letters refer to measurements depicted in Figure 8.

| | B | C | D | E | H |
|---|---|---|---|---|---|
| C | −0.456 | | | | |
| D | −0.865 | 0.673 | | | |
| E | 0.877 | −0.750 | −0.922 | | |
| H | −0.892 | 0.604 | 0.928 | −0.956 | |
| J | 0.877 | −0.630 | −0.968 | 0.920 | −0.904 |

TABLE 3. Cavity-length calculations associated with measured formant frequencies.

| Formant (F) value | Calculated cavity length (cm) | Acoustic correlate of F value |
|---|---|---|
| 932 | 9.50 | Value for /i/, vowel with highest mean $F_1$ |
| 805 | 10.99 | Value for /I/, vowel with lowest mean $F_1$[a] |
| 684 | 12.94 | Value for /o/[b] |
| 841 | 10.52 | Mean $F_1$[a] |
| 844 | 10.48 | Mean $F_1$ for Alex's front[c] vowels /i, I, e, E, ae/ |
| 838 | 10.56 | Mean $F_1$ for Alex's back[c] vowels /a, ə, U/ |
| 2,775 | 3.19 | Value for /i/, vowel with highest $F_2$ |
| 1,433 | 6.18 | Value for /a/, vowel with lowest $F_2$[a] |
| 2,028 | 4.36 | Mean $F_2$[a] |
| 2,350 | 3.77 | Mean $F_2$ for Alex's front[c] vowels /i, I, e, E, ae/ |
| 1,506 | 5.88 | Mean $F_2$ for Alex's back[c] vowels /a, ə, U/ |

[a] /o/ is excluded from calculations because of its outlier status. In 29 of 30 cases, /o/ was characterized by a single, low, broad formant (Patterson and Pepperberg 1994).

[b] Value for /o/ presented for completeness.

[c] Designation front and back refer to tongue-placement charts used for humans (see Fig. 6).

sected specimen of *Psittacus erithacus erithacus*; Patterson unpubl. data) to determine whether the trends predicted from the $X$-ray data were consistent with a two-chamber model. Using the formula,

$$X = 34{,}400/4\ F_n, \qquad (1)$$

where $X$ is the length of a uniform tube in cm and $F_n$ is the $n$th formant frequency in hertz (Hz), we calculate the required cavity length for the production of a measured formant frequency. The results are presented in Table 3.

## DISCUSSION

Our results provide considerable information about a Grey Parrot's vocal tract and the mechanisms this parrot uses to produce certain vowels. We discuss our findings with respect to: (1) comparisons between mechanisms for human and Grey Parrot vowel production; (2) possible correlations among the statistically significant measurements we obtained from our $X$-ray data; (3) the extent to which vowels and their context are related to motions in the vocal tract; (4) a two-chamber model for production in a parrot that is comparable to those proposed for humans; and (5) other possible models for vowel production.

*Mechanisms for vocal production in humans and a Grey Parrot.*—Our investigation reveals some of the mechanisms used by a Grey Parrot to produce the human vowels /a/ and /i/. These mechanisms can be distinctly different from those used by humans. Several striking similarities, however, exist between human and avian vowel production.

The differences between human and avian vocal strategies appear to be a consequence of the differing constraints and flexibilities of their vocal structures. Tukey test results suggest that Alex's production of the point vowels /a/ and /i/ is correlated to the functional morphology of his vocal tract. Alex's trachea has considerable flexibility; it is protracted during the production of /a/, but not during the production of /i/. The procricoid cartilage tilts caudally during the tracheal protraction, positioning it horizontal to the hyobranchial junction. A backward tilt of the head also is associated with /a/ and may facilitate the protraction of the trachea. In contrast, the human trachea is stationary and plays little role in sound modification, probably because it is below the laryngeal sound source (Dickson and Maue-Dickson 1982). Although an open mouth is characteristic of a human /a/ (Fromkin and Rodman 1983), a closed beak was associated with Alex's /a/ (Figs. 9a, b). Then too, Alex's /i/ was associated with a very open beak (Figs. 9a, b), in contrast to the relatively closed mouth of humans (Fromkin and Rodman 1983). Although both mandibles are hinged, the Tukey tests and GLMs show that the angle of the lower mandible to the horizontal contributes most to beak gape. The actual horizontal position of the lower mandible, however, does not change significantly between the vowels we studied. Further investigation is needed to determine what roles tracheal configuration and beak gape play in the resonance of other vowels.

Our study also reveals similarities in the speech mechanisms of Alex and humans. Acoustic information on Alex's vowels (Patterson and Pepperberg 1994) suggests a frontal tongue position for /i/ and a back tongue position for /a/, as is seen for humans (Fig. 6). $X$-ray data are consistent with this finding. Alex's trachea is protracted and the hyobranchial junction (on ventral surface of larynx, where larynx and hyoid bone of tongue intersect; see Fig. 5) is in a caudal position during the production of /a/, but the trachea is not protracted and the hyobranchial junction is in a relatively cranial position for /i/. Judging from these data and Scanlan's (1988:155) statement that "Gross movements of the larynx in birds are necessarily associated with movements of the tongue," we suggest that Alex's tongue, like that of humans, is retracted for /a/ and protracted for /i/. Further exploration is needed to determine whether tongue position in the oro-pharyngeal cavity is merely incidental to tracheal position or beak opening, or if the tongue plays an active role in sound modification.

*Correlations between vocal tract measurements for avian /i/ and /a/.*—Many structures appear to work in concert to enable a Grey Parrot to produce human vowels. Front vowels such as /i/ are characterized by an open beak, a nonprotracted trachea, and probably a frontal position of the tongue, with an anterior orientation of the glottal opening. In contrast, back vowels such as /a/ are characterized by a closed beak, a protracted trachea, and probably a back tongue position, with a superior orientation of the glottal opening.

The correlation matrix demonstrates that, as the beak opens, the hyobranchial junction moves forward (i.e. for /i/) and, as the beak closes, the hyobranchial junction retreats toward the spinal cord (i.e. for /a/). Such a correlation was unexpected because the physical relationship of the involved structures was not immediately obvious from the X-ray images. However, anatomical connections, although indirect, may exist between these structures. Whether movement of these two structures is correlated for vowels other than /i/ and /a/—for example, vowels that differ significantly with respect to $F_1$, such as /i/ and /o/, as well as or instead of $F_2$—is still an issue.

Other aspects of how our Grey Parrot produces human speech remain unclear. The actual role of head tilt is unknown; most likely head tilt simply facilitates opening of the lower mandible, much like head tilt enables humans to open their mouths more widely. The function of structures like the nasal cavities, the syrinx, glottis, and tongue, for example, can only be inferred from the imaging processes we have used so far.

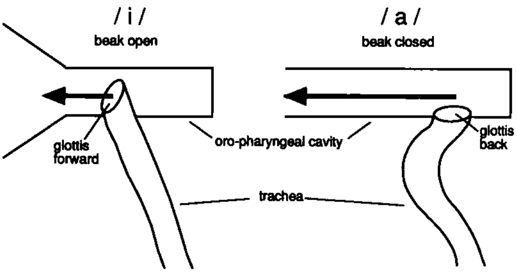

Fig. 11 Schematic drawing of two-chamber model for vowel production in Grey Parrots. Arrows indicate air flow during speech.

*Vowels, context, and motion in the suprasyringeal tract.*—To learn whether vowels or words could be used to predict motions in the suprasyringeal tract, we ran GLMs. We constructed three models (word [i.e. context], vowel, and vowel+word) in order to compare the overlap in variance predicted by context and variance predicted by vowel (i.e. multicolinearity). We needed to learn if the variance could be significantly attributed to vowel rather than to word. Because vowel is included in word, we expected a high degree of multicolinearity, and we actually found that word and word+vowel account for exactly the same amount of variance. Table 1 shows that most of the change in measurements can be attributed to context (i.e. word) but confirm that much of the variance is actually due to vowel. Thus, vowel by itself can explain considerable variance in the physical measurements. Word contains more information because it contains the context as well as the vowel and, in general, additional information increases the amount of variance that can be explained by any model (Cohen and Cohen 1983). Whether we look at word or vowel, we find that motions in the suprasyringeal vocal tract are highly correlated to the particular speech produced. The evidence strongly supports our claim that movement in the suprasyringeal vocal tract is correlated to (and probably causally related to) speech sounds produced by our subject and that the various structures of the vocal tract must move in concert in order to produce the physical configurations that enable a Grey Parrot to emit vowels such as /i/ and /a/.

*Human based model for psittacine speech production.*—Our results suggest that a two-chamber model of speech production (Fig. 11) is a useful starting point for understanding the complex mechanisms enabling speech production in a Grey Parrot. In such a model, the oro-pharyngeal cavity is represented as a short tube that can quickly and substantially alter its configuration, causing relatively large changes in resonant frequency; the trachea is represented as a long tube that keeps a relatively constant length with a relatively constant resonance. The model might allow coupling between chambers in specific instances. The model also provides a means of comparing the mechanisms of Alex's production of vowels with the mechanisms of oscine song production.

Our data are consistent with an association between the oro-pharyngeal cavity and $F_2$, which varies significantly (Tukey test) between /i/ and /a/. We find, for example, that the length of the oro-pharyngeal cavity can be facilely influenced by the degree of beak gape and probably the tongue and glottis. When the beak is open, the tongue forward, and the glottis relatively far to the front, the length of the oro-pharyngeal cavity effectively decreases and its resonant frequency increases, which is what we observe for the vowel /i/ ($\bar{x}$ = 2,775 Hz; Patterson and Pepperberg 1994). When the beak is closed, the tongue back, and the glottis less far to the front, the oro-pharyngeal cavity is effectively enlarged, causing a lowering of its resonant frequency, which is what we observe for /a/ ($\bar{x}$ = 1,433 Hz; Patterson and Pepperberg 1994). The association is also supported by our calculations (Table 3) based on acoustic and anatomical data. To produce the highest and lowest $F_2$'s, Alex's oro-pharyngeal cavity must vary between 3.19 and 6.18 cm in length; to produce the mean $F_2$ across vowels, the cavity would need to be 3.68 cm long. Anatomically, we find that the combined length of the buccal and pharyngeal cavities is 3.69 cm. Whether the length of the oro-pharyngeal cavity can increase by 67% through retraction of the larynx or closing of the beak is unknown, but a 14% decrease from the measured value is within reason. If, however, the amount of vocal tract constriction affects formants to the same extent for parrots as for humans (Gay et al. 1991), Alex might be able to use constriction of the glottis to produce particular resonant frequencies through other, more complex mechanisms than changes in oro-pharyngeal or tracheal length.

Our data are also consistent with an association between the trachea and $F_1$. Although the configuration of the trachea differs considerably between /a/ and /i/, it does not appear to change very much in length. To produce the highest and lowest $F_1$'s, Alex's trachea would have to vary between 9.50 and 10.99 cm; to produce the mean $F_1$ across vowels, the tracheal length should be 10.71 cm. Our observation is supported by data from a dissected Grey Parrot (Patterson unpubl. data), in which the length of a slightly stretched trachea was 11 cm (more than twice as long as structures that comprise the oro-pharyngeal cavity) and could be stretched by only about 10% of its length. Such findings are consistent with an association between the trachea and $F_1$, which does not vary significantly (Tukey test) between /a/ and /i/ ($\bar{x}$ = 872 and 932 Hz, respectively; Patterson and Pepperberg 1994). Possibly the trachea may be significant as a source of low frequencies when protracted; before making any definite claims, however, we must compare X-ray or other, more sophisticated data for tracheal configurations for vowels such as /i/ and /o/, for which $F_1$ does vary significantly (Patterson and Pepperberg 1994).

We also find at least one circumstance which would require coupling of the two chambers. In our calculations, /o/, which in 29 of 30 documented cases (Patterson and Pepperberg 1994) was characterized by only a single formant, requires a tube length of 12.94 cm. Such a length could be produced through a coupling of the tracheal and oro-pharyngeal cavities to form a single tube.

Although our model suggests that beak opening plays an important role in the production of vowels, such an assumption must be evaluated in the light of the role of beak opening in oscine song production (see discussion in Fletcher 1988). Until the results of an immobilization study (Nowicki et al. unpubl. data) were evaluated, beak gape in sparrows was thought to be correlated with frequency rather than amplitude at a given frequency. Conceivably, beak gape in the Grey Parrot, like that in sparrows, tracks the movement of other structure(s) responsible for the $F_1$ and $F_2$ of vowel production.

*Other models.*—Other models for speech production also may be consistent with our data. Models for human speech production, for example, take into account more extensive details of tongue movement (e.g. Stone 1991), nonlinear exchange of energy in the vocal tract to determine what aspects of the speech spectrum should be attributed to the source versus the filter (McGowan 1991), and the limitations that might exist in one- versus three-dimensional models (Kagawa et al. 1992). The general consensus, however, is that, despite the large number of ongoing human studies, models are currently constrained by the quality of empirical knowledge (Crelin 1987). According to Fant (1991:490), "There is an apparent lack of anatomical and physiological data, a lack of insight into dynamic variations of control parameters and lack of flexibility to continuously adapt to configurational variations such as overall tract length, lateral dimensions and essential cavity structures of consonants. We need more insight into voice and noise source interaction, mutually and with respect to the tract system function." Given the paucity of research on psittacine vocal tracts, it is unlikely that any sophisticated models will be developed in the near future. However, although more complex three-dimensional models might provide more detailed correlations between form and function, such models may not be necessary for determining correlations between formant values and physical structures. Different models perform different functions and, even in humans, one-dimensional linear wave-propagation models appear adequate for calculating formant frequencies of a given vocal tract (Maeda 1991).

*General comments.*—Future research into speech mechanisms of psittacids should serve to create a more detailed model of their speech production. Our study is but a first step in determining the mechanisms used by a nonhuman, nonprimate, nonmammal to produce sounds of human speech. We have shown that a Grey Parrot, without benefit of lips and teeth, and with lungs, nasal cavities, trachea, bronchi, larynx, and a tongue that differ considerably from human structures, can produce sounds comparable to those of English speech. The extent to which our findings generalize to other parrots is unknown. The articulatory mechanisms of individual parrots, like those of individual humans (e.g. Johnson et al. 1993), may show considerable speaker-specific differences, and psittacine mechanisms are also likely to differ across species. Our findings nevertheless raise questions about the degree to which human structures and mechanisms are necessary for the production of English speech. Possibly, by comparing and contrasting production mechanisms in "talking" birds and humans, we can more precisely define what is required for speech.

## ACKNOWLEDGMENTS

This work was supported by grants to I.M.P. from the National Science Foundation (BNS 88-20098, 91-96006, and IBN 92-21941) and the NSF Research Experience for Undergraduate program and to D.K.W. by the Howard Hughes Medical Institute Undergraduate Biology Research Program in the Department of Molecular and Cellular Biology and a Barry Goldwater Fellowship. We thank the contributors to the Alex Foundation (particularly The Parrot Club of Manitoba, the Bird Clubs of Virginia, Gateway Parrot Club of St. Louis, the Wasatch Avian Education Society of Salt Lake, the Tennessee Valley Exotic Bird Clubs, and John Paton) for additional support, Art Winfree for introducing D.K.W. to image processing, Audrey Warren for her scientific illustrations, Ying Yong Qi for his patient lessons in image processing and acoustics, Nancy Warren for the purchase of a computer, and the Smithsonian Institution for loan of the *Psittacus erithacus erithacus* specimen (no. 54099). We thank Abbot Gaunt, Sandra Gaunt, and Dominique Homberger for comments on an earlier version of the manuscript. This project would have been impossible without the technical support of the Cardiac Catheterization Laboratory in the Department of Radiology at the University Medical Center, Mike Descour in the Department of Optical Sciences, and Klaus Eichstadt at Merlin Engineering.

## LITERATURE CITED

BOCK, W. J. 1978. Morphology of the larynx of *Corvus brachyrhynchos* (Passeriformes: Corvidae). Wilson Bull. 90:553-565.

BORDEN, G. J., AND K. S. HARRIS. 1984. Speech science primer: Physiology, acoustics, and perception of speech. Williams and Wilkins, Baltimore.

BRACKENBURY, J. H. 1978. A comparison of the origin and temporal arrangement of pulsed sounds in the songs of the Grasshopper and Sedge warblers, *Locusta naeiva* and *Acrocephalus schoenobaenus*. J. Zool. 184:187-206.

BRACKENBURY, J. H. 1982. The structural basis of voice production and its relationship to sound characteristics. Pages 53-73 *in* Acoustic communication in birds. Vol. 1, Production, perception and design features of sound (D. E. Kroodsma, E. H. Miller, and H. Ouellet, Eds.). Academic Press, New York.

BRACKENBURY, J. H. 1989. Functions of the syrinx and the control of sound production. Pages 193-220 *in* Form and function in birds (A. J. King and J. McLelland, Eds.). Academic Press, London.

BURTON, P. J. K. 1974. Jaw and tongue features in Psittaciformes and other orders with special reference to the anatomy of the Tooth-billed Pigeon (*Didunculus stringirostris*). J. Zool. (Lond.) 174:255-276.

CLENCH, M. H. 1978. Tracheal elongation in birds of paradise. Condor 80:423-430.

COHEN, J., AND P. COHEN. 1983. Applied multiple regression/correlation analysis for the behavioral sciences. Erlbaum Associates, Hillsdale, New Jersey.

CRELIN, E. 1987. The human vocal tract: Anatomy, function, development, and evolution. Vantage Press, New York.

DICKSON, D. R., AND W. MAUE-DICKSON. 1982. Anatomical and physiological bases of speech. Little, Brown & Co., Boston, Massachusetts.

FANT, G. 1970. Acoustic theory of speech production, 2nd ed. Mouton & Co., The Netherlands.

FANT, G. 1991. Discussion summaries: Composite models. J. Phonetics 19:489-490.

FLETCHER, N. H. 1988. Bird song—A quantitative acoustic model. J. Theor. Biol. 135:455-481.

FRITH, C. B. 1994. Adaptive significance of tracheal elongation in manucodes (Paradisaeidae). Condor 96:552-555.

FROMKIN, V., AND R. RODMAN. 1983. An introduction to language, 3rd ed. Holt, Reinhart and Williams, New York.

GAUNT, A. S. 1983. An hypothesis concerning the relationship of syringeal structure to vocal abilities. Auk 100:853-862.

GAUNT, A. S., AND S. L. L. GAUNT. 1985. Electromyographic studies of the syrinx in parrots (Aves: Psittacidae). Zoomorphology 105:1-11.

GAY, T., L-J. BOÉ, P. PERRIER, G. FENG, AND E. SWAYNE. 1991. The acoustic sensitivity of vocal tract constrictions: A preliminary report. J. Phonetics 19: 445-452.

GREENEWALT, C. H. 1968. Bird song: Acoustics and physiology. Smithsonian Institution Press, Washington, D.C.

HARRIS, C. L., W. B. GROSS, AND A. ROBESON. 1968. Vocal acoustics of the chicken. Poult. Sci. 47:107-112.

HAUSBERGER, M., J. M. BLACK, AND J. P. RICHARD. 1991. Bill opening and sound spectrum in Barnacle Goose loud calls: Individuals with 'wide mouths' have higher pitched voices. Anim. Behav. 42:319-322.

HERSCH, G. L. 1966. Bird voices and resonant tuning in helium-air mixtures. Ph.D. thesis, Univ. California, Berkeley.

HOMBERGER, D. G. 1979. Functional morphology of the larynx in the parrot *Psittacus erithacus*. Am. Zool. 19:988.

HOMBERGER, D. G. 1980. Funktionell-morphologische Untersuchungen zur Radiation der Ernährungs-und Trinkmethoden der Papageien. Bonn. Zoologische Monographien, No. 13.

HOMBERGER, D. G. 1986. The lingual apparatus of the African Grey Parrot, *Psittacus erithacus Linne* (Aves: Psittacidae). Description and theoretical mechanical analysis. Ornithol. Monogr. 39.

HOMBERGER, D. G. In press. Psittacidae. *In* Fauna of Australia, vol. 2 (G. B. Ross, Ed.). Australian Government Publishing Service, Canberra, Australia.

JOHNSON, K., P. LADEFOGED, AND M. LINDAU. 1993. Individual differences in vowel production. J. Acoust. Soc. Am. 94:701-714.

KAGAWA, Y., R. SHIMOYAMA, T. YAMABUCHI, T. MURAI, AND K. TAKARADA. 1992. Boundary element models of the vocal tract and radiation field and their response characteristics. J. Sound Vib. 157: 385-403.

KING, A. S. 1989. Functional anatomy of the syrinx. Pages 105-192 *in* Form and function in birds (A. S. King and J. McClelland, Eds.). Academic Press, London.

KLATT, D. H., AND R. A. STEFANSKI. 1974. How does a mynah bird imitate human speech? J. Acoust. Soc. Am. 55:822-832.

LADEFOGED, P. 1982. A course in phonetics. Harcourt, Brace, Jovanovich, San Diego, California.

LIEBERMAN, P. 1984. The biology and evolution of language. Harvard Univ. Press, Cambridge, Massachusetts.

MAEDA, S. 1991. Discussion summaries: Toward better models of speech production. J. Phonetics 19: 493-495.

MCCLELLAND, J. 1989. Larynx and trachea. Pages 69-103 *in* Form and function in birds, vol 4 (A. S. King and J. McClelland, Eds.). Academic Press, London.

MCGOWAN, R. S. 1991. Nonlinearities for one-dimensional propagation in the vocal tract. J. Phonetics 19:425-432.

MOORE, C. A. 1992. The correspondence of vocal tract resonance with volumes obtained from magnetic resonance images. J. Speech Hearing Res. 35:1009-1024.

MYERS, J. A. 1917. Studies on the syrinx of *Gallus domesticus*. J. Morphol. 29:165-214.

NEARY, T. 1978. Phonetic features for vowels. Indiana Univ. Linguistics Club, Bloomington.

NOTTEBOHM, F. 1976. Phonation in the Orange-winged Amazon Parrot, *Amazona amazonica*. J. Comp. Physiol. 108:157-170.

NOWICKI, S. 1987. Vocal tract resonances in oscine bird sound production: Evidence from bird song in a helium atmosphere. Nature 325:53-55.

NOWICKI, S., AND P. MARLER. 1988. How do birds sing? Music Percept. 5:391-426.

OLIVE, J. P., A. GREENWOOD, AND J. COLEMAN. 1993. Acoustics of American English speech: A dynamic approach. Springer-Verlag, New York.

PATTERSON, D. K., AND I. M. PEPPERBERG. 1994. A comparative study of human and parrot phonation: Acoustic and articulatory correlates of vowels. J. Acoust. Soc. Am. 96:634-648.

PEPPERBERG, I. M. 1990a. Cognition in an African Grey Parrot (*Psittacus erithacus*): Further evidence for comprehension of categories and labels. J. Comp. Psychol. 104:41-52.

PEPPERBERG, I. M. 1990b. Referential mapping: A technique for attaching functional significance to the innovative utterances of an African Grey Parrot. Appl. Psycholinguist. 11:23-44.

PEPPERBERG, I. M. 1990c. Some cognitive capacities of an African Grey Parrot (*Psittacus erithacus*). Pages 357-409 *in* Advances in the study of behavior (P. J. B. Slater, J. S. Rosenblatt, and C. Beer, Eds.). Academic Press, New York.

PEPPERBERG, I. M., K. J. BRESE, AND B. J. HARRIS. 1991. Solitary sound play during acquisition of English vocalizations by an African Grey Parrot (*Psittacus erithacus*): Possible parallels with children's monologue speech. Appl. Psycholinguist. 12:151-178.

PULLUM, G. K., AND W. A. LADUSAW. 1986. Phonetic symbol guide. Univ. Chicago Press, Chicago.

REMEZ, R., P. RUBIN, L. NYGAARD, AND W. HOWELL. 1987. Perceptual normalization of vowels produced by sinusoidal voices. J. Exp. Psychol. Human Percept. Perform. 13:40-61.

RÜPPELL, W. 1933. Physiologie und Akustik der Vögelstimme. J. Ornithol. 81:433-542.

SAS INSTITUTE. 1989. SAS/STAT user's guide, version 6, 4th ed. SAS Institute, Inc., Cary, North Carolina.

SCANLAN, J. 1988. Analysis of avian "speech": Patterns and production. Ph.D. dissertation, University College, London.

SCHROEDER, L. D., D. L. SJOQUIST, AND P. E. STEPHAN. 1986. Understanding regression analysis: An introductory guide. Series on Quantitative Applications in the Social Sciences, 07-057. Sage Publications, Newbury Park, California.

STEIN, R. 1968. Modulation in bird sounds. Auk 85: 229-243.

STEVENS, K. N., AND A. S. HOUSE. 1961. An acoustical theory of vowel production and some of its implications. J. Speech Hear. Res. 4:303-320.

STONE, M. 1991. Toward a model of three-dimensional tongue movement. J. Phonetics 19:309-320.

STREIDTER, G. 1994. The vocal control pathways in budgerigars differ from those in songbirds. J. Comp. Neurol. 343:35-56.

SUBTELNY, J., W. L. R. WHITEHEAD, AND J. D. SUBTELNY. 1989. Cephalometric and cineradiographic study of deviant resonance in hearing impaired speakers. J. Speech Hear. Disord. 54:249-263.

SUTHERLAND, C. A., AND D. S. MCCHESNEY. 1965. Sound production in two species of geese. Living Bird 4:99-106.

THORPE, W. H. 1959. Talking birds and the mode of action of the vocal apparatus of birds. Proc. Zool. Soc. Lond. 132:441-455.

THORPE, W. H. 1961. Bird-song. Cambridge Univ. Press, Cambridge.

TODT, D. 1975. Spontaneous recombinations of vocal patterns in parrots. Naturwissenschaften 62:399-400.

WARNER, R. W. 1971. The structural basis of the organ of voice in the genera *Anas* and *Aythya* (Aves). J. Zool. (Lond.) 164:197-207.

WESTNEAT, M., J. H. L. LONG, W. HOESE, AND S. NOWICKI. 1993. Kinematics of birdsong: Functional correlation of cranial movements and acoustic features in sparrows. J. Exp. Biol. 182:147-171.

WHITE, S. S. 1968. Movements of the larynx during crowing in the domestic cock. J. Anat. 103:390-392.

WHITE, S. S. 1975. The larynx. Pages 1891-1897 *in* Sissons and Grossman's the anatomy of domestic animals, 5th ed., vol. 2 (R. Getty, Ed.). Saunders, Philadelphia.