

Vocal-Tract Filtering by Lingual Articulation in a Parrot

Gabriël J.L. Beckers,^{1,3,*} Brian S. Nelson,²
and Roderick A. Suthers^{1,2}

¹Medical Sciences

Indiana University School of Medicine and

²Department of Biology

Jordan Hall

Indiana University

Bloomington, Indiana 47405

³Behavioural Biology

Institute of Biology

Leiden University

Kaiserstraat 63

2311 GP Leiden

The Netherlands

Summary

Human speech and bird vocalization are complex communicative behaviors with notable similarities in development and underlying mechanisms [1]. However, there is an important difference between humans and birds in the way vocal complexity is generally produced. Human speech originates from independent modulatory actions of a sound source, e.g., the vibrating vocal folds, and an acoustic filter, formed by the resonances of the vocal tract (formants) [2]. Modulation in bird vocalization, in contrast, is thought to originate predominantly from the sound source [3], whereas the role of the resonance filter is only subsidiary in emphasizing the complex time-frequency patterns of the source (e.g., [4–7], but see [8]). However, it has been suggested that, analogous to human speech production [9], tongue movements observed in parrot vocalizations modulate formant characteristics independently from the vocal source [10–15]. As yet, direct evidence of such a causal relationship is lacking. In five Monk parakeets, *Myiopsitta monachus*, we replaced the vocal source, the syrinx, with a small speaker that generated a broad-band sound, and we measured the effects of tongue placement on the sound emitted from the beak. The results show that tongue movements cause significant frequency changes in two formants and cause amplitude changes in all four formants present between 0.5 and 10 kHz. We suggest that lingual articulation may thus in part explain the well-known ability of parrots to mimic human speech, and, even more intriguingly, may also underlie a speech-like formant system in natural parrot vocalizations.

Results and Discussion

We obtained feral Monk parakeets that were caught in a government pest control program in Florida, where this South American species has been artificially intro-

duced in the past. We euthanized the animals with an overdose of anesthetic immediately before the start of each experiment. The vocal sound source in parrots is formed by vibrating membranes in the syrinx, situated in the trachea at the tracheobronchial junction [16]. We replaced the syrinx with a small speaker that produced frequency sweeps from 0.5 to 11 kHz and measured the effect of tongue position on vocal-tract resonances. We varied the position of the tongue in the beak systematically in a sagittal plane (front-to-back \times low-to-high placements, with respect to the mandible) by using a micromanipulator. The tongue in parrots is underpinned by the hyoid skeleton, which also supports the larynx [15]. Tongue movements thus cause concurrent movements of the larynx [11, 13].

Figure 1 shows, as an example, the resulting vocal tract resonance patterns in one animal. Four formants are clearly visible (F1–F4, with F1 having the lowest frequency). The center frequencies of F1 and F3 and the amplitudes of all four formants clearly depend on tongue position. The four other animals tested showed similar patterns, and we therefore averaged the amplitudes and center frequencies of each formant at different tongue positions over all five animals. The results are shown in Figure 2.

To our knowledge, these are the first experimental data showing that vocal-tract resonances create strong formant patterns in a parrot. Moreover, our study provides evidence for the idea that such formant patterns are modulated by lingual articulation. This evidence is strong because our methodology enabled us to quantify acoustic filtering of the vocal tract separately from the acoustics of the syringeal sound source.

Formant variation in Monk parakeet vocalizations may be even more extensive than shown in this study. We varied the placement of the tongue (and concurrently that of the larynx because both are attached to the hyoid apparatus) but took care to keep other factors constant. There are, for example, a number of other structures that may play a role in formant modulation in real vocalizations. First, although the position of the larynx is structurally coupled to placement of the tongue, the larynx can also be moved with respect to the lingual apparatus, and net laryngeal movements in live animals will thus be the result of the superimposition of this type of movement on lingual movement [15, 17]. The position of the larynx will likely influence resonance characteristics. Second, beak gape changes often occur during vocalizations, presumably altering resonance characteristics. Third, the length of the trachea is likely an important factor in vocal-tract resonance, and it may be modulated during vocalization. Fourth, the glottal opening may be modulated. All of these factors, and perhaps more, can have their own independent dynamics during vocalization and may thus contribute to a high-dimensional system that determines formant patterns in more elaborate ways than those found in this study.

Conversely, it should be noted that an underlying assumption of our study is that the tested tongue place-

*Correspondence: g.j.l.beckers@umail.leidenuniv.nl

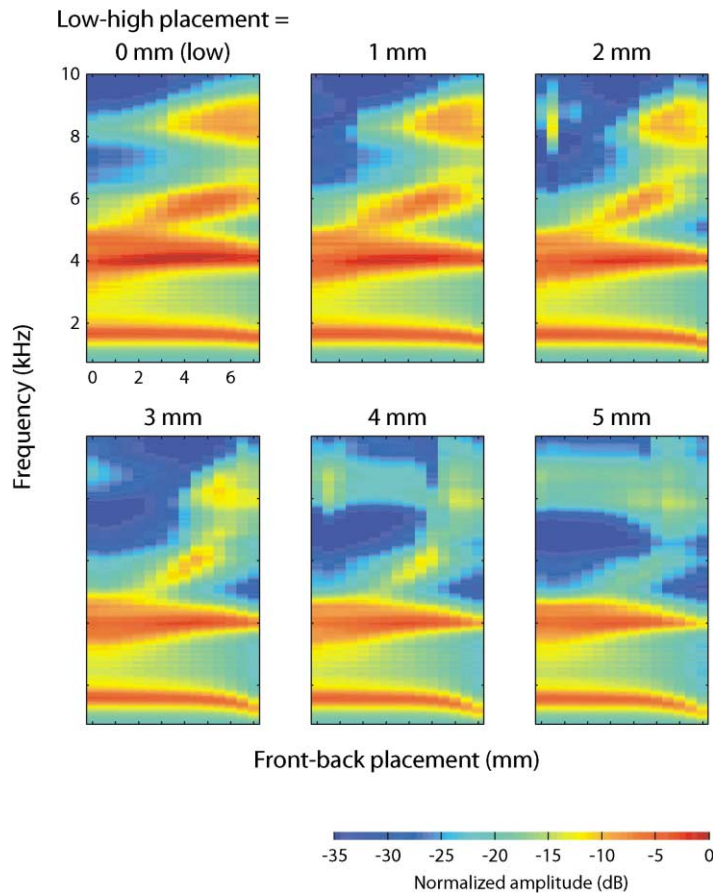


Figure 1. Vocal-Tract Resonance Characteristics as a Function of Tongue Placement in One Individual Parakeet

Each subplot shows the sound energy density as a function of sound frequency and front-to-back placements of the tongue (these placements are expressed as the distance from the most frontal position in the beak). Different subplots correspond to different low-to-high placements (expressed as the distance from the lowest position on the floor of the beak). Red–orange areas indicate relatively high sound levels and correspond to vocal-tract resonances, called formants.

ments are indeed possible in vocalizing parakeets. Tongue movements are often observed in natural Monk parakeet vocalizations (G.J.L.B., personal observations) and in mimic speech of other parrots [10–13], which offers support for this assumption. However, the full range of tongue placements tested in this study, and thus their corresponding formant patterns, need not necessarily be used in vocalizations. Also, there may be differences in the exact vocal-tract geometry and tissue characteristics of vocalizing birds and the dead preparations tested in our study. The degree of glottal opening and the length of the trachea, for example, are unlikely to be exactly the same in both conditions. Such differences may cause deviations in absolute formant frequencies, although we believe they are unlikely to affect the patterns of modulation caused by tongue movements.

In a comparison of our findings in Monk parakeets with human speech production, it is interesting to note that vowels in American English are uniquely determined by F1 and F2 frequencies, the variation in which is caused by variation in the front-to-back and low-to-high placement of the tongue. In Monk parakeets, the same tongue movements also influence formant patterns, but front-to-back tongue movements seem to modulate formant characteristics in a more complex way than do low-to-high movements. Front-to-back movements modulate the center frequencies of F1 and F3, in opposite directions, and the amplitudes of all formants in nonuniform ways. Low-to-high movements have an overall

attenuating effect on the amplitudes of all formants but also influence the center frequencies of F1 and F3.

The frequency axis in Figures 1 and 2 has a linear scale, on which F3 modulation may seem much larger as compared to F1 modulation. However, frequency modulations are not perceived in a linear way in humans, and because basic hearing characteristics are not very different between humans and birds [18, 19], the same may be true for Monk parakeets. In humans, the perception of frequency—pitch—is often expressed on a mel scale, such that frequencies that are spaced at equal distances along this scale are judged to be related by the same pitch ratio [20]. On this scale, F1 variation in Figure 2 has a range of 281 mels, and F3 variation has a range of 392 mels. For a comparison: in American English, F1 and F2 frequencies have ranges of about 500 and 800 mels, respectively [21]. Meaningful differences may be much smaller, however. In female speakers of American English, F1 in /o/ and /a/ vowels, for example, differs about 200 mels, whereas F2 is not different at all.

We believe that a vocal-tract filter with the strong modulatory formant patterns present in Monk parakeets may occur in other parrots as well because the general morphology of the vocal tract is shared among psittaciform birds. In addition to syringeal dynamics [22], this may explain the well-known and remarkable ability of parrots to mimic human speech, in which formant patterns are an important information-coding characteristic. How Monk parakeets mimic speech is an issue not

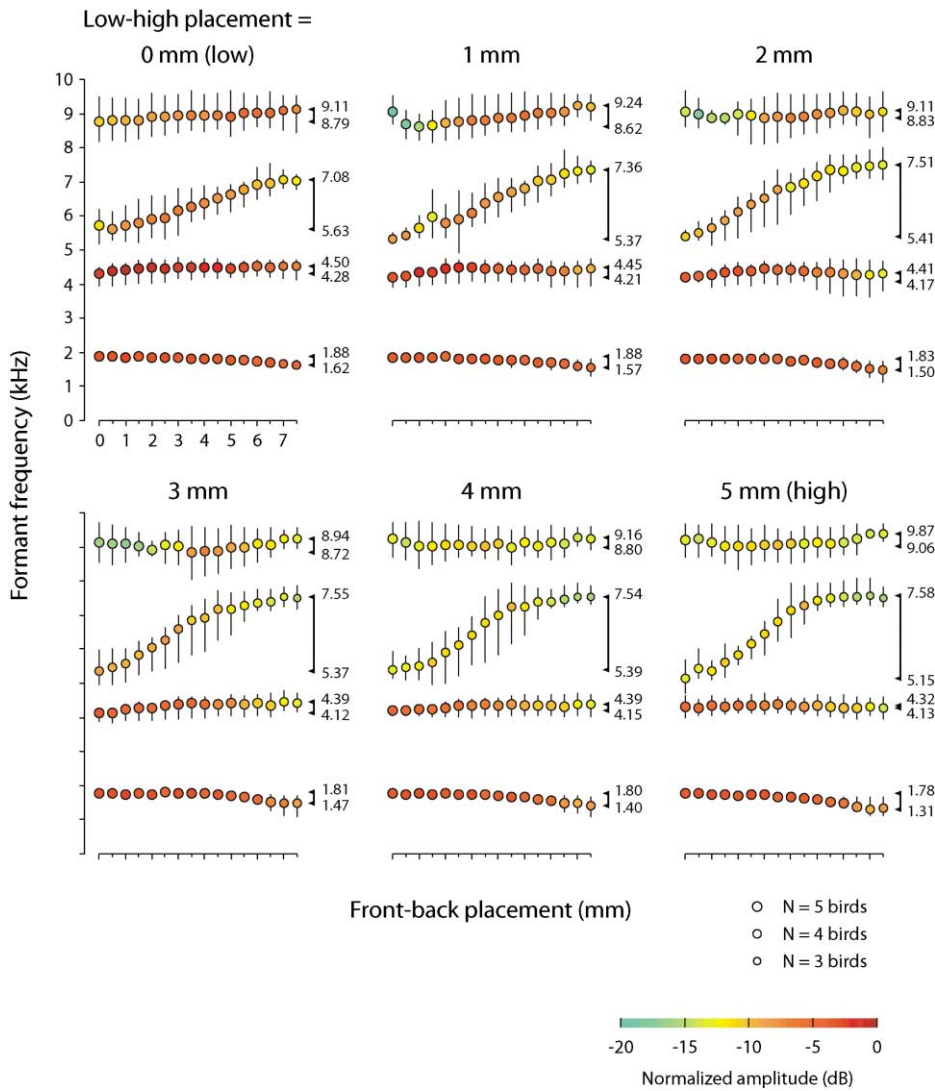


Figure 2. Average Formant Frequencies and Amplitudes as a Function of Tongue Placement in Five Parakeets

Whiskers indicate ranges, and formant amplitude is color-coded in dB relative to the maximum average amplitude. Some tongue placements resulted in less than four significant formants in some animals (see, for instance, frontal and high tongue placements in Figure 1, where F4 is sometimes not present). The number of animals over which the averages are calculated is therefore indicated by the diameter of the circles. Gray vertical bars and numbers at the right of each subplot indicate the range of the average frequency of each formant.

addressed in this study, but it seems clear that matching of F1–F2 absolute frequencies of vowels is unlikely. F1 in speech is far below the lowest formant found in Monk parakeets in this study (0.3–0.9 kHz versus 1.3–1.9 kHz), and it is hard to envision how structures of the relatively small size found in a Monk parakeet vocal tract could produce resonances within the human F1 range. In a series of landmark studies, the acoustic correspondence between human phonemes and their parrot mimic counterparts [12, 14], the correlation between acoustic features and articulation [13], and the morphology of vocal-tract structures [23] have been investigated in Gray parrots, *Psittacus erithacus*. The F1 in mimic vowels of the one individual tested in this species varies between 0.8 and 1.0 kHz and does not match the F1 of tutor vowels outside this range [12]. Nevertheless, mimic vowels were categorized correctly by human listeners,

probably because they rely on F2, which was reproduced correctly, and additional unknown cues. Also, human perception of vowels depends strongly on patterns of formant positions, rather than exact absolute frequencies [24]. The current study shows that the ranges of human F2 and higher formants are within the vocal possibilities of Monk parakeets.

Although the topic of mimic speech in parrots is an interesting one [25], lingual articulation in parrots may also be an important factor in the production of their natural vocalizations. Monk parakeet vocalizations, like many other parrot vocalizations, have a broad-band frequency spectrum [26], and formant patterns such as those found in our study will shape such vocalizations to a significant extent. Figure 3 shows typical examples of natural greeting calls of one of the Monk parakeets used in our study; these examples strongly suggest that

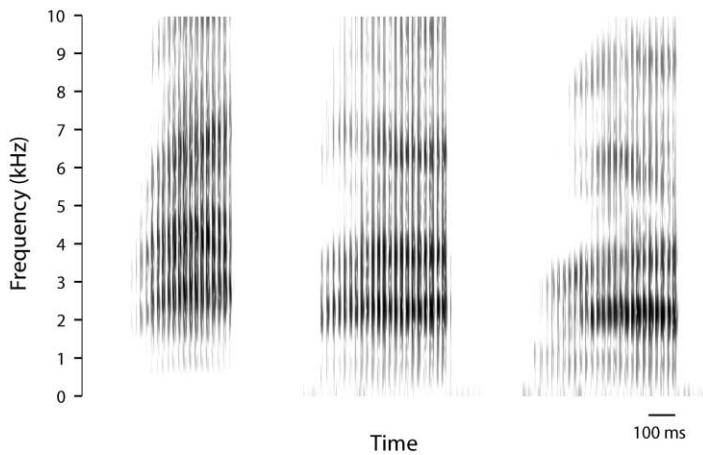


Figure 3. Spectrograms of Three Exemplary Greeting Calls from a Randomly Selected Monk Parakeet Used in this Study

Note the four time-varying darker areas that are suggestive of formants. The other individuals showed similar complex patterns in their greeting calls. The frequency ranges of the highest three of the four areas fall within the frequency ranges of F2, F3, and F4 in Figure 2 very well. The lowest presumed formant, however, is 0.5–0.8 kHz higher than F1 in our measurements. Spectrograms were calculated with a Short-Time Fourier Transform, with a Gaussian 3 ms window and a 30 dB dynamic range.

complex formant patterns are present in the normal vocalizations of this species. Similar patterns can be seen in published spectrograms of a closely related parrot, the Orange-fronted conure, *Aratinga canicularis* [27]. Some caution in this interpretation is needed, however, because vocalizations are the product of the combined action of a sound source and a vocal-tract resonance filter, and we do not know what the contribution of the source sound is in the spectral pattern. The mechanical details of source sound production in parrots are not understood yet. Spectrograms of Monk parakeet vocalizations often show complex patterns of discrete frequency bands, partly harmonically related but partly also arising, apparently, from nonlinear combinations of multiple source frequencies. Without a good understanding of the acoustic properties of the vocal source, the relationship between characteristics of vocalizations and the properties of the vocal-tract filter remains speculative. Nevertheless, it is evident from our results that the tongue movements observed in natural Monk parakeet vocalizations and in mimic speech of other parrots must have a modulatory effect on formant patterns.

The existence of formants has been strongly indicated in a number of other nonhuman vertebrates, such as nonhuman primates, bats, dogs, and birds (for a review, see [28]), although the supporting evidence is not as direct as that in our study. Evidence of the modulation of formant patterns in natural vocalizations of nonhuman animals, however, is rare, although this may be due to a lack of research focused on this specific topic. For the cases in which it has been investigated, the proposed mechanisms of formant modulation are mandible movement and lip protrusion in monkeys [29, 30], as well as the lowering of the larynx in red deer [31]. If formant modulation by tongue movement occurs in parrot natural vocalizations, as our findings suggest, then this has most interesting consequences, for it indicates that the use of lingual articulation to create an extra dimension of vocal complexity may have evolved at least twice, once in humans and once in parrots. Furthermore, if formant modulation by lingual articulation is relevant for vocal communication in parrots, then the lingual neuro-motor program is part of the vocal control system and must be coordinated with syringeal and respiratory vocal motor programs.

The perceptual and functional relevance of complex formant patterns in parrot vocalizations is unknown. To our ears, the formant modulations found in this study are very salient. However, humans are tuned to perceiving formants, which are the basis of information coding in speech. Future research should reveal how salient natural formant patterns are to parrots (see [32–35] for the perception of speech formants in budgerigars, *Melopsittacus undulatus*) and whether or not the complex formant patterns that apparently occur in their natural vocalizations are learned. In addition, playback experiments should provide an insight into the functional relevance of these patterns. More to the point with regard to the current study, further work at the production level should reveal which other vocal-tract structures are articulated by Monk parakeets in the production of their natural vocalizations and perhaps mimic speech; for instance, such work might use cineradiography. With the methodology described here, it will subsequently be possible to investigate the acoustic consequences of identified articulations. We hope that such a combined approach will yield a better understanding of how parrots perceive, learn, and produce their complex vocalizations. Ultimately, this may also contribute to an understanding of the evolutionary origin of human speech [36].

Experimental Procedures

Animals and Preparation

The five adult Monk parakeets used in this study, three females and two males, were anesthetized by inhalation of isoflurane and then euthanized with an overdose of chloroform in the pectoral muscle. The lower end of the trachea and syrinx were made accessible by dissection, ventrally between the crop and clavicles. The trachea was transected at the level of the syrinx, just above the lateral tympaniform membranes (LTMs), which are the sound generators in parrots [16], and left in its relaxed position. A 2 mm length of silastic tube (inner diameter, 1 mm) was then fitted over the port of a small speaker (Knowles EP-7108; 6.3 × 4.3 × 3 mm) and inserted into the end of the trachea so that the speaker was located at the position where the LTMs normally vibrate. The tube was slightly larger than the diameter of the trachea, so that the connection was acoustically sealed. Measuring sound levels after clamping the trachea closed just below the larynx verified the seal at the end of each experiment. The fact that these levels were more than 25 dB lower than in normal recordings showed that very little sound energy leaked through the seal or wall of the exposed part of the trachea and thus that most sound energy radiated from the beak. Each bird

was mounted above the top of a 60 cm steel pole on a small piece of metal mesh wire between two metal bars of a stereotaxic device, which fixed the bird's head. Wedging a thin rectangular steel-wire frame (height 6.9 mm) between the tip of the mandible and halfway along the length of the maxilla and fixing it with tissue adhesive kept the beak gape constant. In natural Monk parakeet vocalizations, beak gape is extremely variable, as verified by video recordings. The beak gape used in our experiments was approximately halfway between the two extremes that occur naturally—almost closed in some variants of the greeting call and wide open in contact calls. The position of the lingual apparatus could be varied with a two-dimensional micromanipulator, which, through a steel rod with a small hook and a suture, was connected to the hyobranchial junction through a slit in the skin of the throat area. Sound-reflecting surfaces were kept to a minimum in this setup. Mounted subjects were positioned at least 60 cm from each wall of a humidified $2 \times 2.4 \times 2.75$ m sound-attenuating chamber (Industrial Acoustic Company) that was lined with at least one layer of acoustic foam (Sonex classic 7.62 cm wedges) to absorb reflections that might otherwise occur from the walls of the chamber.

Acoustic Recordings and Measurements

Linear frequency sweeps (2 s, 0.5–11 kHz) with 10 ms raised cosine onset and offset ramps were generated with an array processor (TDT AP2) and played from the speaker with a digital to analog converter (TDT DD1, 16 bit, 40 kHz) and attenuator (TDT PA4). Sweeps were recorded with a 0.5 inch microphone (Brüel & Kjaer 4189) mounted on the end of a threaded rod (90 cm \times 0.95 cm) and positioned 15 cm from each subject's beak. Recordings were digitized (TDT DD1, 16 bits, 40 kHz, 15 kHz anti-alias filter) and saved to a disk via the array processor. We recorded one sweep for each tongue position and varied tongue position in 0.5 mm front-to-back steps and 1 mm low-to-high steps. The total number of front-to-back and low-to-high steps possible in each animal varied somewhat, but for analyses we used the first 16 front-to-back steps (spanning 7.5 mm) and first six low-to-high steps (spanning 5 mm) in all animals. We thus covered a square matrix of 96 tongue positions in a sagittal plane within the beak of each animal. The frequency response of the speaker system is not flat in the range of interest. Before each session we therefore recorded the speaker output of a frequency sweep separately, from the position where the beak would be during measurements. This recording was used as a reference so that in the analyses we could correct for frequency response deviations of the complete stimulus-generating system. For sound analyses we used the software program Praat (freely available from <http://www.praat.org>), version 4.1.13 for Linux. To quantify vocal-tract resonance patterns, we created a long-time average spectrum of each recorded frequency sweep, with 10 Hz bin widths, and corrected this spectrum for deviations measured in the reference recording. We quantified formant characteristics by using Burg's linear predictive coding algorithm, available in Praat. To this end, we randomized the phase part of the recording's spectrum and applied an inverse Fourier transform to obtain a colored noise sound with an amplitude spectrum identical to that of the recorded frequency sweep. Formant measurements obtained from such sounds were verified by visual inspection of spectrograms.

Acknowledgments

We thank Michael Avery of the United States Department of Agriculture for providing the Monk parakeets and Jack Bradbury, Carel ten Cate, and two anonymous referees for comments on an earlier version of this manuscript. This research was supported by National Institute of Neurological Disorders and Stroke, National Institutes of Health grant R01 NS029467 to R.A.S.

References

1. Doupe, A.J., and Kuhl, P.K. (1999). Birdsong and human speech: Common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567–631.
2. Fant, G. (1960). *Acoustic theory of speech production* (The Hague: Mouton and Co).
3. Greenewalt, C.H. (1968). *Bird song: Acoustics and physiology*. (Washington: Smithsonian Institution Press).
4. Nowicki, S. (1987). Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. *Nature* **325**, 53–55.
5. Westneat, M.W., Long, J.H., Hoese, W., and Nowicki, S. (1993). Kinematics of birdsong: Functional correlation of cranial movements and acoustic features in sparrows. *J. Exp. Biol.* **182**, 147–171.
6. Hoese, W., Podos, J., Boetticher, N.C., and Nowicki, S. (2000). Vocal tract function in birdsong production: experimental manipulation of beak movements. *J. Exp. Biol.* **203**, 1845–1855.
7. Beckers, G.J.L., Suthers, R.A., and ten Cate, C. (2003). Pure-tone birdsong by resonance filtering of harmonic overtones. *Proc. Natl. Acad. Sci. USA* **100**, 7372–7376.
8. Goller, F., Mallinckrodt, M., and Torti, S.D. (2004). Beak gape dynamics during song in the Zebra finch. *J. Neurobiol.* **59**, 289–303.
9. Lieberman, P. (1984). *The biology and evolution of language* (Cambridge, MA: Harvard University Press).
10. Nottebohm, F. (1976). Phonation in the orange-winged amazon parrot, *Amazona amazonica*. *J. Comp. Physiol.* **108**, 157–170.
11. Scanlan, J.P. (1988). Analysis of avian "speech": patterns and production. PhD thesis (London: University College).
12. Patterson, D.K., and Pepperberg, I.M. (1994). A comparative study of human and parrot phonation: Acoustic articulatory correlates of vowels. *J. Acoust. Soc. Am.* **96**, 634–648.
13. Warren, D.K., Patterson, D.K., and Pepperberg, I.M. (1996). Mechanisms of American English vowel production in a Grey parrot (*Psittacus erithacus*). *Auk* **113**, 41–58.
14. Patterson, D.K., and Pepperberg, I.M. (1998). Acoustic and articulatory correlates of stop consonants in a parrot and a human subject. *J. Acoust. Soc. Am.* **103**, 2197–2215.
15. Homberger, D.G. (1999). The avian tongue and larynx: Multiple functions in nutrition and vocalisation. In *Proceedings of the 22nd International Ornithological Congress*. Durban, N.J. Adams and R.H. Slotow, eds. (Johannesburg: BirdLife South Africa), pp. 94–113.
16. Larsen, O.N., and Goller, F. (1999). Role of syringeal vibrations in bird vocalizations. *Proc. R. Soc. Lond. B. Biol. Sci.* **266**, 1609–1615.
17. Homberger, D.G. (1986). The lingual apparatus of the African Grey parrot, *Psittacus erithacus* Linné (Aves: Psittacidae): Description and theoretical mechanical analysis. *Ornithological Monographs*, **39**.
18. Dooling, R.J., Lohr, B., and Dent, M.L. (2000). Hearing in birds and reptiles. In *Comparative Hearing: Birds and Reptiles*. R.J. Dooling, R.R. Fay, and A.N. Popper, eds. (New York: Springer-Verlag), pp. 308–359.
19. Wright, T.F., Cortopassi, K.A., Bradbury, J.W., and Dooling, R.J. (2003). Hearing and vocalizations in the orange-fronted conure (*Aratinga canicularis*). *J. Comp. Psychol.* **117**, 87–95.
20. Stevens, S.S., and Volkman, J. (1940). The relation of pitch to frequency. *Am. J. Psychol.* **53**, 329–353.
21. Stevens, K.N. (1998). *Acoustic phonetics* (Cambridge MA: MIT Press).
22. Lavenex, P.B. (1999). Vocal production mechanisms in the budgerigar (*Melopsittacus undulatus*): The presence and implications of amplitude modulation. *J. Acoust. Soc. Am.* **106**, 491–505.
23. Pepperberg, I.M., Howell, K.S., Banta, P.A., Patterson, D.K., and Meister, M. (1998). Measurement of Grey parrot (*Psittacus erithacus*) trachea via magnetic resonance imaging, dissection, and electron beam computed tomography. *J. Morphol.* **238**, 81–91.
24. Ladefoged, P. (2001). *Vowels and consonants* (Malden, MA: Blackwell Publishers).
25. Thorpe, W.H. (1959). Talking birds and the mode of action of the vocal apparatus of birds. *Proc. Zool. Soc. London* **132**, 441–455.
26. Martella, M.B., and Bucher, E.H. (1990). Vocalizations of the Monk parakeet. *Bird Behaviour* **8**, 101–110.
27. Vehrencamp, S.L., Ritter, A.F., Keever, M., and Bradbury, J.W. (2003). Responses to playback of local vs. distant contact calls

- in the orange-fronted conure, *Aratinga canicularis*. *Ethology* 109, 37–54.
28. Fitch, W.T. (2002). Comparative vocal production and the evolution of speech: reinterpreting the descent of the larynx. In *The Transition to Language*. A. Wray, ed. (Oxford: Oxford University Press), pp 21–45.
 29. Hauser, M.D., Evans, C.S., and Marler, P. (1993). The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations. *Anim. Behav.* 45, 423–433.
 30. Hauser, M.D., and Schön Ybarra, M. (1994). The role of lip configuration in monkey vocalizations—experiments using xylocaine as a nerve block. *Brain Lang.* 46, 232–244.
 31. Fitch, W.T., and Reby, D. (2001). The descended larynx is not uniquely human. *Proc. R. Soc. Lond. B. Biol. Sci.* 268, 1669–1675.
 32. Dooling, R.J., and Brown, S.D. (1990). Speech perception by budgerigars (*Melopsittacus undulatus*): Spoken vowels. *Percept. Psychophys.* 47, 568–574.
 33. Dooling, R.J. (1992). Perception of speech sounds by birds. In Y. Cazals, L. Demany and K. Horner, eds. *Advances in Biosciences: Auditory physiology and perception* (London: Pergamon), pp. 407–413.
 34. 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.
 35. Dent, M.L., Brittan-Powell, E.F., Dooling, R.J., and Pierce, A. (1997). Perception of synthetic /ba/-/wa/ speech continuum by budgerigars (*Melopsittacus undulatus*). *J. Acoust. Soc. Am.* 102, 1891–1897.
 36. Fitch, W.T. (2000). The evolution of speech: a comparative review. *Trends Cogn. Sci.* 4, 258–267.