News and Views

Multiple discontinuities in nonhuman vocal tracts — A response to Lieberman (2006)

T. Riede [a,*], E. Bronson [b], H. Hatzikirou [c], K. Zuberbühler [d]

[a] *University of Wisconsin, Department of Surgery, 5745 Medical Sciences, Madison, WI, USA*
[b] *Smithsonian's National Zoological Park, 3001 Connecticut Ave., NW, Washington, DC 20008, USA*
[c] *Center for Information Services and High-Performance Computing, Technical University Dresden, 01062 Dresden, Germany*
[d] *School of Psychology, University of St. Andrews, St. Andrews, Scotland, UK*

Male West African Diana monkeys (*Cercopithecus diana*) produce acoustically distinct alarm calls to two of their main predators, the crowned eagle and the leopard. The calls are remarkable in their acoustic structure for at least three reasons. First, they exhibit clear formant frequencies, a defining feature of human speech. Second, Diana monkeys are able to modify the basic structure of these formants by creating formant transitions. In leopard alarm calls, the first formant describes a transition of approximately 150 Hz, while the transition of the second one is about 200 Hz. In striking contrast, the two formants remain relatively constant in eagle alarm calls (Riede and Zuberbühler, 2003a,b). Experiments have shown that these acoustic differences have semantic value to recipients (Zuberbühler, 2003). Third, the first and second formants in Diana monkeys' alarm calls are in close proximity, a feature not normally observed in animal vocalizations. We propose that this formant proximity is the result of a discontinuity (or 'non-uniformity') along the monkey's vocal tract.

In a previous paper (Riede et al., 2005), we sought to understand the production mechanisms of the Diana monkey vocal system using a computer modelling technique, based on anatomical data. Lieberman (2006) has challenged a number of our findings. His concerns relate to the possibility of tongue movements, pharyngeal constrictions, and the position of the larynx. Additionally, the technique of deriving vocal tract area functions from lateral x-rays is criticized. We would like to respond to these criticisms by first summarizing our main point, the notion of non-uniform vocal tracts in nonhumans, before responding in detail to Lieberman's critique.

How do Diana monkeys achieve vocal flexibility?

When an acoustic wave propagates in a cavity, such as a vocal tract, boundary effects determine the acoustic structure of the signal. The acoustic impedance of the cavity (Z) is not constant over the cross section of the cavity but relates to air density (ρ), speed of sound (c), and area of the cross section (A) according to Eq. (1).

$$Z = \frac{\rho^* c}{A} \tag{1}$$

The last factor (i.e., a change in the cross section area) is most relevant in this context. We suggested that the formant patterns observed in Diana monkey alarm calls are the likely result of an area change in the vocal tract caused by an articulatory movement and permanent presence of a constriction (Riede et al., 2005). A uniform vocal tract, in contrast, can only generate equally spaced formant frequencies without any transitions. Using phonetic terminology, a uniform vocal tract takes a "neutral configuration" which produces a "neutral sound" (Laver, 1994), as for example observed in the grunts of chacma baboons (reviewed in Riede et al., 2005). However, some vocalizations, including human speech sounds, dog growls, or rhesus monkey coo calls, deviate from this pattern (Riede and Fitch, 1999; Table 1 of this reply). Diana monkey alarm calls are another good example of such deviation from the "equal formant distance pattern," strongly suggesting that their vocal tract is non-uniform (i.e., constricted somewhere along its length between vocal folds and lips) during vocal production (Riede and Zuberbühler, 2003b; Fig. 1 in Riede et al., 2005).

Additionally, Diana monkeys are able to change the contour of these formants when responding to leopards by creating formant transitions. To account for the transitions, we predicted

* Corresponding author.
    E-mail address: tobiasriede@web.de (T. Riede).

Table 1
Frequency differences between 5 successive formants (F₁, F₂, F₃, F₄, F₅) in vocalizations from non-human mammals and human laughter

Species, vocal type	$F_2 - F_1$	$F_3 - F_2$	$F_4 - F_3$	$F_5 - F_4$	
Female rhesus monkey, coo calls	765	1227	1515	1427	(1)
Female rhesus monkey, grunts	1750	1518	1668	1465	(1)
Male chacma baboon, grunts	730	816	725		(2)
Female chacma baboon, grunts	982	1247	730		(2)
Female chacma baboon, grunts	1010	1259	668		(3)
Male human, open mouth laughing	1057	984	1091	926	(4)
Female human, open mouth laughing	1060	1162	798	833	(4)

Differences were calculated from published formant values. In Chacma baboons the fifth formant is not available. A uniform vocal tract is expected to produce equally spaced resonance frequencies. Formant spacing, i.e., the differences between subsequent formants ($F_2 - F_1$, $F_3 - F_2$, $F_4 - F_3$, $F_5 - F_4$), is expected to be identical in a uniform tube setting. Note that formant patterns deviate from equal spacing, in particular in all female vocalizations.
(1) Table 2 in Rendall et al., 1998; (2) Table 1 in Rendall et al., 2005; (3) Table 1 in Owren et al., 1997; (4) Table 4 in Bachorowski and Owren, 2001.

that Diana monkeys must be able to alter some of the constrictions within their vocal tract through articulatory movements (Riede et al., 2005). We identified three such movements: lowering of the larynx, movements of the mandible, and movements of the lips. When incorporated into our vocal tract model, these factors were sufficient to replicate the formant patterns produced by wild monkeys. We concluded that our model was a good approximation of the actual biological processes taking place in the Diana monkey vocal tract during vocal production.

The role of the tongue and pharyngeal constrictions

Lieberman (2006) first points out that "… the size and position of their tongue preclude the possibility of human like formant formation" and that "… the tongue maneuvers proposed by Riede et al. (2005) are most improbable." We take notice of these points, but continue to observe that the formant patterns produced by Diana monkeys do not fit the expectations for a uniform tube, and that our anatomical investigations suggest the existence of a constriction. Lieberman may be right that tongue movement does not play a major role during call production, but we never suggested that to be the case. Using lateral radiographs, we identified a permanent discontinuity in the vocal tract of anaesthetized Diana monkeys and dead animals (Riede et al., 2005). We found the resting tongue to be involved in producing a 2 to 3 cm long constriction (near the midpoint of the vocal tract) between the frontal oral cavity and the more caudally positioned laryngo-pharyngeal cavity. The intra-laryngeal supra-glottal cavity and the caudal part of the pharynx (both parts considered as the laryngo-pharyngeal

cavity) is wider than the more rostrally located oro-pharynx, the area captured by the back of the tongue and the hard palate (Fig. 4 in Riede et al., 2005). A wide intra-laryngeal supra-glottal cavity preceding the pharyngeal area is expected in species with enlarged laryngeal cavities (Schneider, 1964; Frey and Riede, 2003).

Traditionally, non-human vocal tracts are modelled as uniform or slightly flared tubes. However, to accommodate area changes, a common method has been to model vocal tracts as a series of cylindrical tubes with varying diameters (see for instance Fig. 6.1 in Titze, 1994). Our anatomical findings suggested that Diana monkey vocal tracts should be modelled as a "three-tube approximation," consisting of a wide laryngo-pharyngeal tube, a narrow oro-pharyngeal tube, and a wide oral tube. Adding a fourth very short rostral tube enabled the model to perform fine adjustments, resembling the lip movements observed in wild animals. In short, at no point do we assume a role for tongue movements, although there is no fundamental reason to exclude it.

Of course, Lieberman is right in stating that human speech is directly dependent on a high agility of the human tongue body (Stone et al., 2004). The ability to coordinate a framework of muscles in this region allows humans to produce a wide variety of pharyngeal cross-sectional areas, which generates the formant variability typical for speech (e.g., Story et al., 1996). However, sophisticated tongue movement is also observed in non-humans, typically in the context of food transport. The tongue can be moved upward and pushed against the hard palate. In addition, there is a ventro-dorsal component, which moves the tongue body toward the dorsal pharyngeal wall (cats: Thexton and McGarrick, 1988, 1989; Kobara-Mates et al., 1995; rhesus monkeys: Hiiemae et al., 1995). Whether or not similar movements play a role during vocal behavior is largely unknown.

Humans are special due to the low resting position of their larynx. Despite the more cranial larynx position of most mammals, discontinuities (i.e., constrictions) in the pharyngeal region have been observed. The important point is that none of them rely on movements of the tongue or the larynx. Moreover, there is evidence that the ability to constrict the pharynx is phylogenetically old, tracing as far back as crocodilians. Smith (1992) reports: "… in crocodilians, a process of the hyoid can be elevated to meet the secondary palate and seal the oral cavity (Busbey, 1989)." This constriction is observed during deglutition, the passing of food from the mouth into the esophagus accomplished in part by the superior, middle, and inferior pharyngeal constrictor muscles (pterygo-pharyngeus, cerato-pharyngeus, laryngo-pharyngeus). Interestingly, these muscles are also involved in human speech (Shpritzen et al., 1975; Bell-Berti, 1976; Niimi, 1981).

Dog vocalisations have already been mentioned as an example of animal signals produced by a non-uniform vocal tract (Riede and Fitch, 1999). Dreyfuss (1911) dissected the ventral pharyngeal constrictor of a dog from the thyroid and cricoid cartilage, taking care not to injure the laryngeal nerves and the pharyngeal branches of the glossopharyngeus. After surgery the dog was not able to bark but only to grunt, although

it had no difficulties in swallowing. Grunting then changed into growling, and on the fifth postoperative day the dog barked again normally. Dreyfuss further reported that 'previous resonances' were not restored.

In sum, there is good evidence that animals are able to constrict the pharynx and other regions to achieve acoustic effects similar to those caused by tongue movements. In addition, the head posture and changes in head position during vocalization have an effect on the geometry of the pharyngeal region, which influences the formant frequencies of a vocal utterance (Hellsing, 1989; Muto et al., in press). Our central claim is that constrictions in a nonhuman vocal tract can be caused by other mechanisms apart from tongue movements.

The role of the larynx

Lieberman mentions the implications of larynx position. As pointed out, humans are special compared to the other primates because of the low resting position of the larynx (Flügel and Rohen, 1991). As a consequence, the human pharyngeal region is relatively long, as long as the oral region. In addition, the larynx describes a cranio-caudal movement during speech (Laukkanen et al., 1999; Honda et al., 1999). However, laryngeal descent is not a uniquely human trait; it has also been found in chimpanzees (Nishimura et al., 2003; Nishimura, 2005), Pantherinae (Weissengruber et al., 2002), and red deer and fallow deer (Fitch and Reby, 2001). The crucial issue is not the resting position of the larynx, but its position during vocalization (Nottebohm, 1976). Similar to human speech, temporary lowering of the larynx during vocalization has been reported in rhesus monkeys (Hiiemae et al., 1995), cats (Kobara-Mates et al., 1995), pigs, goats, dogs, squirrel monkeys, red deer, and fallow deer (Fitch, 2000; Fitch and Reby, 2001). Acoustically, larynx-lowering causes equal lowering of all formant frequencies, due to vocal tract elongation (Fant, 1960). However, larynx movement may also lead to changes in pharyngeal vocal tract shape, leading to changes in formant frequencies (demonstrated in humans: Story et al., 1996).

Lieberman (2006) further argues that there is a problem with the length of the Diana monkeys' vocal tract. We think this problem arises because of the assumption that a downscaled human vocal tract is an appropriate model for Diana monkeys. Yet, the human vocal tract shows specific growth patterns in different segments (Vorperian et al., 2005). This almost certainly results in different positions of constrictions relative to vocal tract length, suggesting that downscaling is an inappropriate approach to understand this system.

Assessing the cross-sectional areas of the vocal tract

In our 2005 study, we used anatomical data to model the Diana monkey vocal tract. We were able to measure the midsagittal dorso-ventral distance using lateral radiographs and post-mortem dissection (Riede et al., 2005). We found that the oro-pharyngeal area was narrower than the laryngo-pharynx or the rostral oral cavity. These physiological facts are a prerequisite for the close proximity between the first and second formant in the monkeys' alarm calls. Lateral dimensions in the dead animal were similar from the glottis to the lips.

Lieberman (2006) mentioned that these measurements might not be meaningful because we relied on lateral radiographs, hereby describing the dorso-ventral dimension of the vocal tract only. Although we cannot rule out this possibility completely, our approach is standard. Midsagittal data are used in articulatory models of human speech. They are transformed into an area function, for example by using the midsagittal dorso-ventral distance as the diameter of a circle or one dimension of an elliptic area (e.g., Maeda, 1972; Johannson et al., 1983; Perrier et al., 1992). We used the midsagittal dorso-ventral dimensions as the diameter of a hypothesised circle area to build up a tube-like structure. Based on the dead animal, we have no reason to believe that the lateral dimension of the vocal tract changes significantly over the length of the vocal tract. Thus, the sagittal distances measured in our three anaesthetized Diana monkeys are likely to represent a reasonable estimate of the area function of the monkeys' vocal tract. Nevertheless, we agree with Lieberman that more sophisticated techniques capable of documenting the vocal tract changes in vocalizing animals are desirable for future studies (Riede et al., 2005).

The air sac hypothesis and issues

Diana monkeys possess a single air sac, which extends subcutaneously with an opening (about 0.5 cm in diameter) into the larynx between the epiglottis and thyroid cartilage (Riede and Zuberbühler, 2003b; Riede et al., 2005). Lieberman (2006) suggests that the formant patterns of Diana monkey alarm calls may be a direct result of the air sac. We agree that this is an interesting point and we therefore explored this hypothesis by implementing a side branch to our original model.

We modelled the side branch as a rigid tube, positioned close to the sound source. As a result, both formants decreased with increasing air sac size. Of course, the side branch model explored here is limited and rather artificial. Real air sacs are not rigid tubes but inflatable sacs, which are more likely to emit acoustic energy through their walls, rather than reflecting them back into the laryngeal cavity. This fact alone questions their role in affecting formant patterns as such (Riede et al., 2004; Fletcher et al., 2004). Moreover, one early experimental study showed that puncturing (i.e., disabling) the air sac of a live monkey had no effect on the formant distributions of the call (Gautier, 1971). The actual effects of air sacs on formant frequency distributions may thus be minimal.

It is also important to note that in Diana monkeys the outlet of the air sac is close to the vocal folds. Current theory predicts that vocal fold oscillation is determined not only by the physical characteristics of the vocal folds themselves, but also by

the air stream passing through them (Titze, 1994). Hence, an alternative hypothesis is that, in Diana monkeys, an interaction takes place at the level of the glottis between the air stream coming from the air sac and the glottal oscillation caused by the air stream from the lungs.

Many details, including the acoustic role of the subcutaneous air sac or the motility of the tongue, deserve further investigations. Unlike work with Neanderthals (Boë et al., 2002), however, we are in the fortunate position to address these issues empirically, provided that Diana monkeys manage to evade extinction (http://www.iucnredlist.org/) due to habitat destruction and illegal bushmeat trade in West Africa.

References

Bachorowski, J.A., Owren, M.J., 2001. Not all laughs are alike: voiced but not unvoiced laughter elicits positive affect in listeners. Psychol. Sci. 12, 252—257.

Bell-Berti, F., 1976. An electromyographic study of velopharyngeal function in speech. J. Speech Hear. Res. 19, 225—240.

Boë, J.-L., Heim, J.-L., Honda, K., Maeda, S., 2002. The potential Neandertal vowel space was as large as that of modern humans. J. Phon. 30, 465—484.

Busbey, A.R., 1989. Form and function of the feeding apparatus of *Alligator mississippiensis*. J. Morphol. 202, 99—127.

Dreyfuss, K., 1911. Experimentelle Untersuchungen über den Einfluss des Schlundschnürers auf die Kehlkopfmuskulatur. Verhandlungen Ver. dtsch. Laryng. 18, 492—508.

Fant, G., 1960. Theory of Speech Production. Mouton, The Hague.

Fitch, W.T., 2000. The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. Phonetica 57, 205—218.

Fitch, T.W., Reby, D., 2001. The descended larynx is not uniquely human. Proc. R. Soc. Lond. B 268, 1669—1675.

Fletcher, N., Riede, T., Beckers, G.J.L., Suthers, R.A., 2004. Vocal tract filtering and the 'coo' of doves. J. Acoust. Soc. Am. 116, 3750—3756.

Flügel, C., Rohen, J.W., 1991. The craniofacial proportions and laryngeal position in monkeys and man of different ages, a morphometric study based on CT-scans and radiographs. Mech. Ageing Dev. 61, 65—83.

Frey, R., Riede, T., 2003. Sexual dimorphism of the larynx of the Mongolian gazelle (*Procapra gutturosa* Pallas, 1777) (Mammalia, Artiodactyla, Bovidae). Zool. Anz. 242, 33—62.

Gautier, J.P., 1971. Etude morphologique et fonctionnelle des annexes extralaryngées des cercopithecinae; liason avec les cris d'espacement. Biol. Gabon. 7, 229—267.

Hellsing, E., 1989. Changes in the pharyngeal airway in relation to extension of the head. Eur. J. Orthod. 11, 359—365.

Hiiemae, K.M., Hayenga, S.M., Reese, A., 1995. Patterns of tongue and jaw movement in a cinefluorographic study of feeding in the macaque. Arch. Oral Biol. 40, 229—246.

Honda, K., Hirai, H., Masaki, S., Shimada, Y., 1999. Role of vertical larynx movement and cervical lordosis in f0 control. Lang. Speech 42, 401—411.

Johannson, C., Sundberg, J., Wilbrand, H., Ytterbergh, C., 1983. From Sagittal Distance to Area: A Study of Transverse, Cross-sectional Area in the Pharynx by Means of Computer Tomography. Speech Transmission Lab., Dept. of Speech Communication, Royal Inst. of Technology, Stockholm, pp. 39—49.

Kobara-Mates, M., Logemann, J.A., Larson, C., Kahrilas, P.J., 1995. Physiology of oropharyngeal swallow in the cat: a videofluoroscopic and electromyographic study. Am. J. Physiol. 268, 232—241.

Laukkanen, A.M., Takalo, R., Vilkman, E., Nummenranta, J., Lipponen, T., 1999. Simultaneous videofluorographic and dual-channel electroglottographic registration of the vertical laryngeal position in various phonatory tasks. J. Voice 13, 60—71.

Laver, J., 1994. Principles of Phonetics. Cambridge University Press, Cambridge, New York.

Maeda, S., 1972. On the Conversion of Vocal Tract X-ray Data into Formant Frequencies. Bell Laboratories, Murray Hill, NJ.

Muto, T., Yamazaki, A., Takeda, S., Kawakami, J., Tsuji, Y., Shibata, T., Mizoguchi, I. Relationship between the pharyngeal airway space and craniofacial morphology, taking into account head posture. Int. J. Oral Maxillofac. Surg., in press. doi:10.1016/j.ijom.2005.04.022.

Niimi, S., 1981. Pharyngeal wall movement during speech. Auris Nasus Larynx 8, 27—34.

Nishimura, T., Mikami, N., Suzuki, J., Matsuzawa, T., 2003. Descent of the larynx in chimpanzee infants. Proc. Natl. Acad. Sci. 100, 6930—6933.

Nishimura, T., 2005. Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. Am. J. Phys. Anthropol. 126, 193—204.

Nottebohm, F., 1976. Vocal tract and brain: a search for evolutionary bottlenecks. Ann. N. Y. Acad. Sci. 280, 643—649.

Owren, M.J., Seyfarth, R.M., Cheney, D.L., 1997. The acoustic features of the vowel-like grunt calls in chacma baboons (*Papio cynecephalusursinus*): implications for production processes and function. J. Acoust. Soc. Am. 101, 2951—2963.

Perrier, P., Böe, L.J., Sock, R., 1992. Vocal tract area function from midsagittal dimensions with CT scans and a vocal tract cast: modelling the transition with two sets of coefficients. J. Speech Hear. Res. 35, 53—67.

Rendall, D., Owren, M.J., Rodman, P.S., 1998. The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalization. J. Acoust. Soc. Am. 103, 602—614.

Rendall, D., Kollias, S., Ney, C., Lloyd, P., 2005. Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. J. Acoust. Soc. Am. 117, 944—955.

Riede, T., Fitch, T., 1999. Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). J. Exp. Biol. 202, 2859—2867.

Riede, T., Zuberbühler, K., 2003a. Pulse register phonation in Diana monkey alarm calls. J. Acoust. Soc. Am. 113, 2919—2926.

Riede, T., Zuberbühler, K., 2003b. The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization. J. Acoust. Soc. Am. 114, 1132—1142.

Riede, T., Beckers, G.J.L., Blevins, W., Suthers, R.A., 2004. Inflation of the oesophagus and vocal tract filtering in ring doves. J. Exp. Biol. 207 (23), 4025—4036.

Riede, T., Bronson, E., Hatzikirou, H., Zuberbühler, K., 2005. Vocal production mechanisms in a non-human primate: morphological data and a model. J. Hum. Evol. 48, 85—96.

Schneider, R., 1964. Der Larynx der Säugetiere. In: Handbuch der Zoologie, 8. Band 5, pp. 1—128.

Shpritzen, R.J., McCall, G.N., Skolnick, M.L., Lencione, R.M., 1975. Selective movement of the lateral aspects of the pharyngeal walls during velopharyngeal closure for speech, blowing, and whistling in normals. Cleft Palate J. 12, 51—58.

Smith, K.K., 1992. The evolution of the mammalian pharynx. Zool. J. Linn. Soc. 104, 313—349.

Stone, M., Epstein, M.A., Iskarous, K., 2004. Functional segments in tongue movements. Clin. Linguist. Phon. 18, 507—521.

Story, B.H., Titze, I.R., Hoffman, E.A., 1996. Vocal tract area functions from magnetic resonance imaging. J. Acoust. Soc. Am. 100, 537—554.

Thexton, A.J., McGarrick, J.D., 1988. Tongue movements in the cat during lapping. Arch. Oral Biol. 33, 331—339.

Thexton, A.J., McGarrick, J.D., 1989. Tongue movements in the cat during the intake of solid food. Arch. Oral Biol. 34, 239—248.

Titze, I.R., 1994. Principles of Voice Production. Prentice Hall, Englewood Cliffs.

Vorperian, H.K., Kent, R.D., Lindstrom, M.J., Kalina, C.M., Gentry, L.R., Yandell, B.S., 2005. Development of vocal tract length during early childhood: a magnetic resonance imaging study. J. Acoust. Soc. Am. 117, 338—350.

Weissengruber, G.E., Forstenpointner, G., Peters, G., Kubber-Heiss, A., Fitch, W.T., 2002. Hyoid apparatus and pharynx in the lion (*Panthera leo*), jaguar (*Panthera onca*), tiger (*Panthera tigris*), cheetah (*Acinonyx jubatus*) and domestic cat (*Felis silvestris f. catus*). J. Anat. 201, 195—201.

Zuberbühler, K., 2003. Referential signalling in non-human primates: cognitive precursors and limitations for the evolution of language. Adv. Study Behav. 33, 265—307.