

01. Authors:

Bart de Boer
Marcus Perlman

02. Word Counts:

Abstract: 59
Main Text: 929
References: 316
Entire Text: 1347

03. Commentary Title:

Physical mechanisms may be as important as brain mechanisms in evolution of speech

04. 04. 06. 07. 08. 09. Contact Details:**Bart de Boer**

Vrije Universiteit Brussel, AI-lab
Pleinlaan 2
1050 Brussels
Belgium
003226293755
bart@ai.vub.ac.be
<http://ai.vub.ac.be/members/bart>

Marcus Perlman

University of California, Merced, Department of Cognitive and Information Sciences
5200 N. Lake Rd.
Merced, CA 95343
United States
mperlman@ucmerced.edu
<http://mperlman.org>

10. Abstract

We present two arguments why physical adaptations for vocalization may be as important as neural adaptations. Firstly, fine control over vocalization is not easy for physical reasons, and modern humans may be exceptional. Secondly, we present an example of a gorilla that shows rudimentary voluntary control over vocalization, indicating that some neural control is already shared with great apes.

11. Main Text

Ackermann, Hage and Ziegler propose a model of the evolution of neural adaptations related to the production of spoken language. Although we are convinced of the importance of such adaptations, and although the authors themselves state that "...the model outlined here addresses only one out of several building blocks..." we would nevertheless like to make two reflections on their article. Our first reflection is on the assumption that independent control

over the vocal folds and the upper vocal tract is somehow a given, and our second reflection is on the ability of apes to control vocalization voluntarily.

Following Fitch (2000a) the authors assume that animal vocalizations have a source and a filter. They also appear to assume that source and filter are independent as they are in modern humans, which is not necessarily the case. In many instances, the behavior of the source is in fact strongly coupled to that of the filter (e.g. in woodwind instruments). Source-filter theory was originally formulated in the context of human speech (Fant, 1960). However, the fact that independence of source and filter is a good approximation for human speech does not mean it is universally valid.

Fletcher (1993) has investigated the theory of vibrating valves and found that the independence of source and filter depends on the precise shape and configuration of the source. In addition, it depends on the ratio of resonance frequencies of the source and the filter. Titze (2008) has adapted the theory to human-like vocal folds, and found that if the frequency at which the vocal folds vibrate is near the resonance frequencies of the vocal tract, strong coupling can occur. Apparently modern human vocal folds and vocal tracts avoid strong coupling, but it is an open question whether this was the case in our evolutionary ancestors.

The little that we do know about ape vocal anatomy appears to argue against independence of source and filter. One instance of this is the large air sacs present in all great apes (Hewitt et al., 2002), which lower the resonance frequency of the upper vocal tract considerably (de Boer, 2008), and would therefore increase coupling (as found in model experiments by Riede et al., 2008). In addition, chimpanzee vocal folds (the only ones about which we have anatomical data) have so-called vocal lips (Kelemen, 1969; Demolin & Delvaux, 2006), and thus a very different shape from human vocal folds. Although we do not know the function of these vocal lips, this difference between two closely related species underscores the point that we should not just assume similar behavior of their vocalization systems.

In systems where source and filter cannot behave independently, the set of signals that can be produced is necessarily more limited. This consequence is demonstrated in a modeling study showing that when source and filter are closely coupled, vocalization may be more chaotic, and thus it may be more difficult to time the onset of vocalization precisely (de Boer, 2012). Given these observations, it may not just be a lack of neural control that makes precise vocalizations difficult for non-human primates. It may also be that the anatomy of their vocal folds and their vocal tracts makes it much harder as well.

Our second point of commentary is to note evidence of at least one case in which a non-human primate appears to have some voluntary control over her larynx in the performance of learned, species atypical vocalizations. Koko, a human-reared, female gorilla (Patterson & Linden, 1981), has been video-recorded performing numerous instances from a repertoire of play behaviors involving voluntary control over her larynx and supralaryngeal vocal tract in coordination

with various manual gestures and action routines (Perlman et al. 2011). This repertoire includes the production of breathy-voiced sounds and glottal stops in situations that are determined by the particular play routine.

Perlman et al. (2011) describe how Koko exhibits vocal control in her play behavior of “talking” into telephones, when she often directs breathy grunt-like vocalizations into the receiver, which she holds to her mouth (voicing was observed in 42 of 68 exhalations over 11 bouts). That she exercises voluntary control over her larynx in these vocalizations is suggested by the contrast of this behavior to her routine of huffing on the lenses of eyeglasses as if to clean them. As in the real human performance of cleaning eyeglasses, Koko produces, in this case, open-mouthed audible huffs that are distinctly and without exception voiceless (observed in 25 exhalations over 12 bouts). Another dimension of vocal control is demonstrated in her voluntary performance of a mock “cough”, which involves a glottal stop, often in coordination with a gesture in which she covers her mouth with an open hand. In several instances, she produces this behavior on command, demonstrating clear voluntary control over the closure of her glottis.

These behaviors appear to be examples of voluntary control over laryngeal motor activity outside of a species-typical audiovisual display, something that the authors say has not been attested yet in great apes. Apparently we should not discount the possibility that apes – and by implication our last common ancestor - have more (rudimentary) abilities to control vocalization voluntarily than is often assumed.

Given that 1) control over vocalization is not just limited by neural factors, but also by purely anatomical and physiological ones, and that 2) a gorilla has been shown to have some rudimentary voluntary control over vocalization, we conclude that in the evolution of speech, anatomical and physiological adaptations to the vocal folds and the vocal tract may have been as important as neural adaptations of their control.

12. Reference List

- De Boer, B. (2008). The acoustic role of supralaryngeal air sacs. *Journal of the Acoustical Society of America*, 123(5 Pt 2), 3732–3733.
- de Boer, B. (2012). Air sacs and vocal fold vibrations: Implications for evolution of speech. *Theoria et Historia Scientiarum*, 9, 13-28.
- Demolin, D., & Delvaux, V. (2006). A comparison of the articulatory parameters involved in the production of sounds of bonobos and modern humans. In A. Cangelosi, A. D. M. Smith & K. Smith (Eds.), *The evolution of language: Proceedings of the 6th international conference (evolang6)* (pp. 67–74). New Jersey: World Scientific.
- Fant, G. (1960). *Acoustic theory of speech production*. 'sGravenhage: Mouton.
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in cognitive sciences*, 4(7), 258–267.
- Fletcher, N. H. (1993). Autonomous vibration of simple pressure-controlled valves in gas flows. *Journal of the Acoustical Society of America*, 93(4 Pt. 1), 2172–2180.

- Hewitt, G. P., MacLarnon, A., & Jones, K. E. (2002). The functions of laryngeal air sacs in primates: A new hypothesis. *Folia Primatologica*, 73, 70–94.
- Kelemen, G. (1969). Anatomy of the larynx and the anatomical basis of vocal performance. In G. H. Bourne (Ed.), *The chimpanzee* (pp. 165-186). Basel: Karger.
- Patterson, F. G. & Linden, E. (1981). *The Education of Koko*. New York: Holt, Rinehart, & Winston.
- Perlman, M., Patterson, F. G., & Cohn, R. H. (September, 2011). The incorporation of learned breathing-related behavior into the multimodal gestures and rituals of an enculturated gorilla. Paper presented at the New College Oxford Conference on Embodied Language, Oxford, United Kingdom.
- Riede, T., Tokuda, I. T., Munger, J. B., & Thomson, S. L. (2008). Mammalian laryngeal air sacs add variability to the vocal tract impedance: Physical and computational modeling. *Journal of the Acoustical Society of America*, 124(1), 634–647.
- Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *Journal of the Acoustical Society of America*, 123(5), 2733-2749.