

A COMPARISON OF THE ARTICULATORY PARAMETERS INVOLVED IN THE PRODUCTION OF SOUND OF BONOBOS AND MODERN HUMANS

DIDIER DEMOLIN

Department of Linguistics, Universidade de Sao Paulo, Brasil and Université libre de Bruxelles, Belgium

VÉRONIQUE DELVAUX

FNRS and Université de Mons-Hainaut, Belgium

Most studies of vocalizations with chimpanzees and Bonobos focus on the interpretation of the vocal behaviour of both captive and free-ranging groups to relate sounds produced to their semantic contexts. Spectrographic analyses reveal the acoustic structure of the vocalizations but rarely raise the question of the specific articulatory capacities of Bonobos in relation to the acoustics. This point is essential if one wants to understand the articulatory control that Bonobos have on their vocalizations. It is also important when the vocalizations of Bonobos and the sound produced by modern humans are compared.

1. Introduction

Fitch (2000) in a major paper reviewing the evolution of speech in a comparative perspective shows that two changes were necessary prerequisites for modern human speech abilities: the modification of vocal morphology and the development of vocal imitative ability. Focusing on the first point, one can wonder why do our closest animal relatives, like Bonobos, lack vocal output capabilities comparable with ours. Comparative data from living nonhuman primates will likely shed light on this question and will allow inferring the functions of vocalizations in early hominids. As stated by Fitch (2000: 260), the evolution of speech was closely tied to mechanisms of sound production and the most obvious speech-related difference between humans and other mammals concerns the structure of the human vocal tract. A central point is about the fact that the human vocal tract anatomy differs from other primates. The human larynx rests much lower in the throat than in apes and humans lack laryngeal air sacs. In addition, human speech requires sophisticated nervous control. Therefore as the relation between acoustic and articulatory aspects of human speech are better understood than that of living apes and particularly of Bonobos, it is worth evaluating the vocalizations of the latter from the known principles of the acoustics of human speech. For example taking the perspective of the source-filter theory (Fant 1960) it is possible to discuss the nature of the

source of vocalizations and to make inferences about how the latter is filtered in the vocal tract.

2. Material and method

2.1 Articulatory data

Articulatory data come from dissections, MRI data and 3D scans made on dead specimens from the Zoo of Planckendael in Belgium. Specimens were fixed for 6 months before being dissected. Our data come from 1 male, 1 female and 1 juvenile. 3D scan data come from the male subject and MRI data from the female and juvenile subject. These data allow an accurate description of the size and shape of the Bonobo's vocal tract. From an articulatory viewpoint, the point is to establish the position and the shape of the lips, the jaw, the velum, and the tongue. Two other parameters are considered: (i) the morphology of the larynx and of the vocal folds (ii) the shape and the dimension of the vocal tract.

2.2 Acoustic and video data

2.2.1 Method

Three recording sessions (of approximately 1h30' each) were made with two males, two juveniles and several females, using a digital video camera with an incorporated microphone. These sessions were made at the Zoo of Plankendael in Belgium. One session was made outdoors and two indoors, out of which 45 sequences of vocalizations have been extracted to constitute the corpus for the present study. Each sequence may include several types of vocalizations (e.g. one staccato hooting, then 4 full screams, etc.) produced by different individuals. But we have excluded from the corpus those sequences in which several individuals vocalize together at the same time (mixed vocalizations). In about one third of the sequences (approximately) a Bonobo is filmed by the camera while vocalizing (face or profile), i.e. by recording the sound and the corresponding image.

2.2.2 Repertoire

In order to compare our data with previous work on Bonobo's vocalizations, they were organized within the 16 categories (and subcategories) of De Waal (1988). Our data reveal several occurrences of 11 out of these 16 categories: low hooting, staccato hooting, legato hooting, wieew bark, whistle bark, (food) peep, peep yelp, peep scream, rasp scream, bark scream, full scream. What is missing in our data is: contest hooting (in De Waal 1988, contest hooting was uttered by one single individual); greeting grunt and panting laugh (this is aperiodic sounds of very low intensity hard to extract from ambient noise); pout moans (in De Waal 1988 pout moans were uttered by adolescent males only and we have two adults and two juvenile males but no adolescent in our group); alarm peeps (there has been no opportunity to raise the alarm during our three recording sessions. Some vocalizations in our data do not match any of the categories from

De Waal: (i) two kinds of soft sounds that we would define as: “soft bark” and “soft peep”, (ii) the “piano”, a sequence of non modulated peeps varying in intensity and pitch.

Data were processed with ‘*MovieSignalExplorer*’ (Delvaux & Soquet), that is a customized application that allows (i) to analyze the acoustics of the vocalizations using spectrograms, spectra, energy analyzer and pitch analyzer; (ii) to display every (30 ms long) portion of the acoustic signal together with the corresponding digital picture recorded by the video camera.

3. Results

3.1 Articulatory data

The main features that can be observed from the articulatory measurements are the already known fact that the vocal tract shape of Bonobos is quite different from that of modern humans. The larynx is higher and the tongue is longer and narrower. However there is another important feature to be observed, that is the long flat and narrow palate and the smaller velum.

The larynx of Bonobos shows structures quite similar to that of modern humans with the important difference that the epiglottis seems to be less vertical than in humans as it can be seen in Figure 1. The V shape of the hyoid bone is also different when compared to the U shape of modern humans.



Figure 1. Larynx of a female Bonobo that shows the rather horizontal position of the epiglottis.

3.2 Acoustic data

Acoustic data show that Bonobos seem to modulate the vocal output on the temporal and spectral levels. Modulations of the fundamental frequency accompany variations of intensity. Acoustically it also appears that a second sound source (diplophonia) can frequently be observed in Bonobo's vocalizations. This second source is related to the resonances (formants) of the vocal tract that seem to have a uniform shape most of the time. Noticeable perturbations of the vocal folds vibrations are also frequently observed. Modulations of the harmonics and of the resonances are associated with modulations of the intensity of the sound source. Most of these features can be observed on Figure 2 and 3.

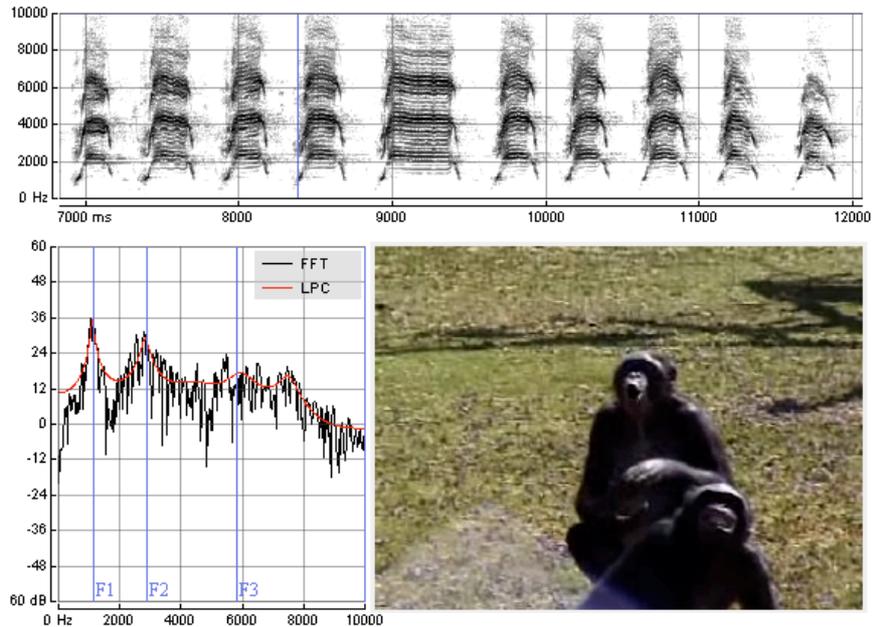


Figure 2. Spectrogram, FFT and LPC spectra of a bark scream produced by the Bonobo staying at the back of the picture. The spectrogram suggests that the formants and the set of harmonics are produced by two different sources. The barks are also modulated in duration.

4. Discussion

From the perspective of the source filter theory and comparing the relation between acoustic and articulatory data there are some striking observations that confirm and complement the observations made in a number of previous studies (e.g. De Waal 1988, Fitch and Hauser 1995, Owren and Rendall 1997 and Bermejo and Omedes 1998).

The source of vocalizations seems to be much more diverse than usually observed in modern humans as noted by Hauser (1996). The vocal fold vibrations that produce a quasi periodic sound and that consist in a fundamental frequency and a set of harmonics is present as a source. However when diplophonia is observed the origin of the second source must also be explained. Schön Ybarra (1995) and Mergell et al. (1999) have suggested that the thin upward extension of the vocal fold that is present in many primate species might serve as independent low-mass oscillators to support the quite high frequencies produced in vocalizations and that these vocal membranes underlie vibration at multiple frequencies, thus giving the observed diplophonia. For the moment, it is difficult to confirm whether this is true or not for the vocalizations of Bonobos. However we would like to draw the attention to the fact that various other

sources can be at play. The sphincteric phonation that implies the vibration of the arytenoids cartilages against the tubercle of the epiglottis can act a phonatory mechanism as well (Traill 1986). Esophageal phonation could also act as source in some vocalizations (Traill 1980). Another possible source is the epiglottis. Indeed as observed in the anatomical data, the epiglottis has a much less vertical position when it is compared to man. Even if it is separated from the root of the tongue in the process of a dissection, this position is striking. The length of the epiglottis that covers an area greater than the larynx itself is also to be noticed. This position could allow the epiglottis to act in a way similar to the opening of a flute to produce the high pitched whistling observed in Bonobo's vocalizations.

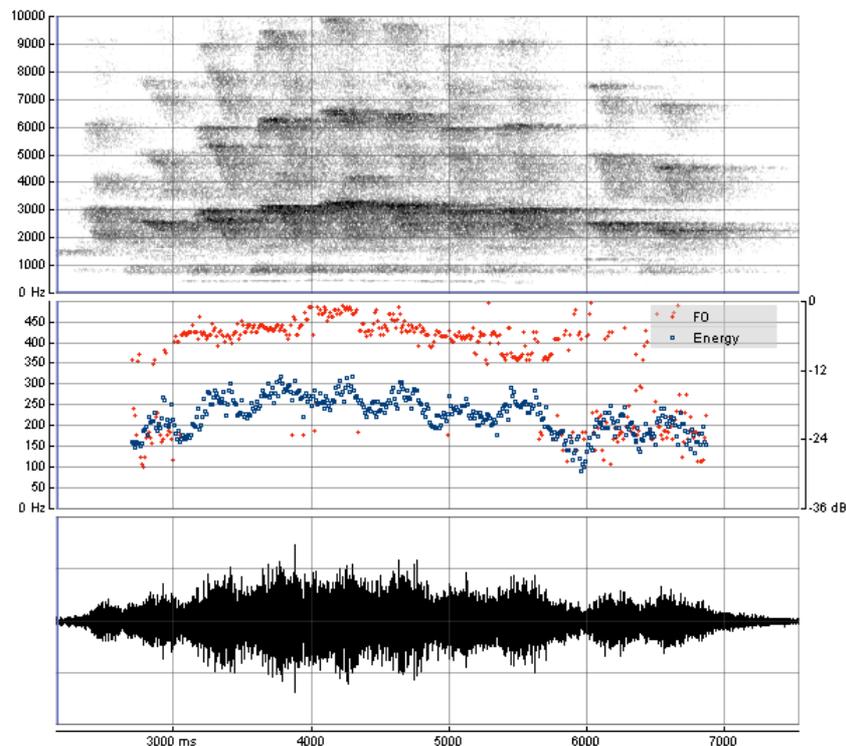


Figure 3. Spectrogram, Fo and energy curves of a piano vocalization. The resonances that can be observed on the spectrogram are accompanied by noise. Modulations of intensity (energy curve) correspond to modulations on the Fo curve.

Finally it has to be noted that if Fo variations are observed they are never independent of intensity. This suggests that there is a close association between Fo, intensity and subglottal pressure in Bonobo's vocalizations. This feature is not true for modern humans contrary to frequent claims (Ohala 1990, Demolin

forthcoming). The fact that F_0 can be controlled independently of subglottal pressure is even a major evolutionary feature of human speech when compared to the vocalizations of primates.

When the filter properties are considered there are several important features to be noticed. The first is that no vowels and consonants are observed during Bonobos vocalizations. When formants are observed in primate vocalizations they almost always suggest that they are produced in a uniform acoustic tube, a feature that is very different of the vast majority of vowels produced by humans. Fitch and Hauser (1995) emphasize the importance of formants in the vocalization of primates and show that the low position of the adult human larynx gives the capacity to produce sounds that have different formant patterns easily. They conclude that the descent of the larynx was a key innovation in the evolution of speech. Fitch (2000: 260) also claims that the change in the larynx position greatly expands the phonetic repertoire of humans because the tongue can move both vertically and horizontally within the vocal tract after this larynx lowering. This is obviously true but it is only part of the explanation. In order for the tongue to move vertically and horizontally, the oral cavity must also be reshaped in these dimensions. This feature and the reshaping of the tongue itself are the consequences of the reduction of prognatism. This has been a major feature in the evolution of hominids and had important consequences on the production of speech because it allows the tongue to make closures and constrictions in the front part of the vocal tract where the vast majority of consonants are produced. Frayer and Nicolay (2000: 231) when comparing the upper palate dimension of modern humans with various hominids and chimpanzees conclude that most specimens older than roughly 250 000 years before the present do not differ from the chimpanzees.

Finally it has to be noted that even if prosodic aspects involving intonation and F_0 contours are not present, there are modulations of duration and intensity like those of Figures 2 and 3 that suggest that these parameters might be controlled in vocalizations.

Acknowledgements

We wish to thank our colleagues Linda Van Elsaker and Kris d'Aout who made the work with the Bonobos of the Zoo in Plankendaal possible and who provided the specimen for dissection. Sergio Hassid and Emile Godefroid of the Hopital Erasme in Brussels made the work dissection possible.

This research was supported by a grant 'Evolution des langues, du langage et de la parole' of the National Bank of Belgium.

References

- Bermejo, M. and A. Omedes (1998). Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan Paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologia*, 70: 329-357.
- Demolin, D. (forthcoming). Phonological universals and the control and regulation of speech production. In *Experimental approaches to Phonology*. P. Beddor, M. Ohala and M.-J. Solé (Eds.). Oxford University Press.
- De Waal, F. (1988). The communicative repertoire of captive Bonobos (*Pan Paniscus*), compared to that of chimpanzees. *Behaviour*, 106: 183-251.
- Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton. The Hague.
- Fitch, W.T. (2000). The evolution of speech a comparative perspective. *Trends in Cognitive Science*, 4 (7): 258-267.
- Fitch, W.T. and M. Hauser (1995). Vocal production in nonhuman primates: acoustics, physiology and functional constraints on 'honest' advertisement. *American Journal of Primatology*, 37: 191-219.
- Freyer, D. W. and C. Nicolay (2000). Fossil evidence for the origin of speech. In N. L. Wallin, B. Merker and S. Brown (eds.) *The origin of Music*. Cambridge: MIT Press.
- Hauser, M. (1996) *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Mergell, P. , W. T. Fitch and H. Herzel (1999). Modelling the role of nonhuman vocal membranes in phonation. *Journal of the Acoustical Society of America*, 105 (3) : 2020-2027.
- Ohala, J.J. (1990). Respiratory activity in speech. In W.J. Hardcastle and A. Marcahl (Eds.) *Speech production and speech modeling*. Kluwer Academic Publishers. 23-53.
- Owren, M. J. and D. Rendall (1997). An affect-conditioning model of nonhuman primate vocal signaling. *Perspectives in Ethology*, 12: 299-346.
- Schön Ybarra, M. (1995). A comparative approach to the nonhuman primate vocal tract: Implications for sound production. In E. Zimmermann, J.D. Newman & U. Jürgens (eds.), *Current topics in primate vocal communication*, New-York. Plenum Press: 185-198.
- Traill, A. (1980). The Laryngeal sphincter as a phonatory mechanism in !xoo Bushman. In *Variation, Culture and Evolution in African Populations: papers in Honor of Dr Hertha de Villiers*. Johannesburg. Witwatersrand University Press: 123-131.
- Traill, A. (1986). A note on pitch control in esophageal speech. *Die Suid-Afrikaanse Tydskrif vir Kommunikasieafwykings*, Vol 27: 95-98.