

AIR SACS AND VOCAL FOLD VIBRATION: IMPLICATIONS FOR EVOLUTION OF SPEECH

Introduction

Air sacs are large cavities that are connected to the vocal tract, and that occur in many primates and other mammals (Frey *et al.*, 2007). In particular, all apes have them, except for humans. There are many possible variations on the exact anatomy of mammalian (and even primate) air sacs, and in the location at which they connect to the vocal tract. They must therefore have evolved separately in different mammalian lineages. However, ape air sacs are all of the same type: lateral ventricular air sacs (Starck & Schneider, 1960; Hayama, 1970; Hewitt *et al.*, 2002) and are therefore likely to be an ancestral trait. The ventricles are small side branches of the vocal tract that connect just above the vocal folds. In chimpanzees, gorillas and orangutans the ventricles are the entrance tubes of large (several liters in some cases) cavities. In humans on the other hand, lateral ventricles only exist as vestigial structures, and only in exceptional (pathological) cases do humans have small air sacs, called laryngocoeles (Stell & Maran, 1975).

It is also interesting to note that, at least in the *Homininae* (gorillas, chimpanzees, bonobos and humans) there appears to be a correlation between the shape of the hyoid bone and the presence or absence of air sacs. The hyoid bone (figure 1) is the only bony part of the vocal tract and the muscles of the larynx and the tongue connect to it. The hyoid bone has a cup shaped extension (called the hyoid bulla) in gorillas, chimpanzees and bonobos, but lacks

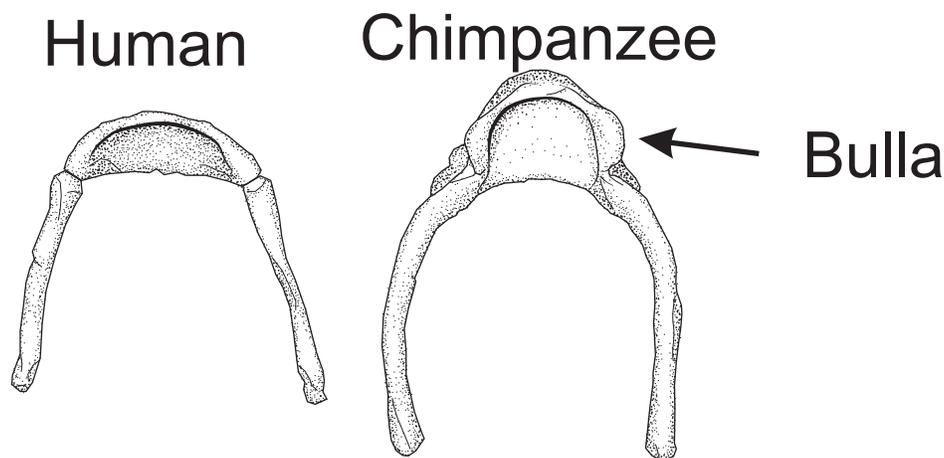


Figure 1: Comparison of a human and chimpanzee (female) hyoid bone. Note the large bulla of the chimpanzee hyoid. This is hollow on the inside and contains the connection to the air sac. In the male hyoid there is no bulla, only a small ridge. Based on replicas made by Bone Clones.

this cup shaped extension in humans (Kohlbrugge, 1896; Brown & Ward, 1988; Aiello & Dean, 2002; Alemseged *et al.*, 2006; Fitch, 2009). The connection between the ventricles and the air sac passes underneath this bulla. Hyoid bones sometimes fossilize, and it has been found that Neanderthals (Arensburg *et al.*, 1989; Arensburg *et al.*, 1990) and *Homo heidelbergensis* (Martínez *et al.*, 2008) have a hyoid bone without a bulla, while *Australopithecus afarensis* (Alemseged *et al.*, 2006) appears to have a bulla. This indicates that air sacs disappeared somewhere between 3.3 million years ago and 500 000 years ago in human evolution.

Primate researchers have written a lot about the function of air sacs (Negus, 1949; Starck & Schneider, 1960; Avril, 1963; Gautier, 1971; Schön, 1971; Marten *et al.*, 1977; Hilloowala & Lass, 1978; Haimoff, 1983; Schön Ybarra, 1986; Harrison, 1995) and also about the relation between human speech and the absence of air sacs in humans (Camper, 1779; Kohlbrugge, 1896; Fitch, 2000; Hewitt *et al.*, 2002; de Boer, 2008b; Fitch, 2009; Hombert, 2010). However, most of this was not based on accurate investigation of the physical effects of air sacs. Recent research that does make use of accurate physical models (Riede *et al.*, 2008; de Boer, 2009) has indicated that air sacs have an important effect on the acoustics of vocalizations and it is therefore interesting to investigate whether disappearance of air sacs could have something to do with complex vocal communication, i. e. speech or singing.

Air sacs can have three effects on vocalization: on the acoustics of the upper vocal tract (which frequencies are amplified and which are attenuated), on the radiation of sound (how efficiently is the sound radiated) and on the source of the sound (vocal fold vibration). Effects on the upper vocal tract and on radiation have been explored elsewhere (de Boer, 2008a; Riede *et al.*, 2008; de Boer, 2009). Interaction on vocal fold vibration with respect to the power of the radiated sound has also been explored elsewhere (Riede *et al.*, 2008). This paper focuses on another aspect of vocal fold vibration: its regularity. It has been proposed that air sacs can cause vocal fold vibration to become irregular (Fitch *et al.*, 2002; Riede *et al.*, 2008). However, the significance for vocal communication of this is unclear. Fitch *et al.* (2002) have proposed that it could be positive, as it might help to highlight the resonance frequencies of the upper vocal tract, while Riede *et al.* (2008) propose that irregular vocal fold vibration might make it more difficult to produce different speech

sounds. There exist computational models of irregularity of voicing of vocal tracts without air sacs (e. g. Hatzikirou *et al.*, 2006; Titze, 2008). There also exists a model that investigates vocal fold-vocal tract interactions in the presence of air sacs, but this has not been used to study irregularity of voicing (Riede *et al.*, 2008). The model presented here investigates the effect of an air sac on the regularity of vocal fold vibration in the presence of an air sac. Irregularity of voicing has been found in many animals (Fitch *et al.*, 2002) most notably in chimpanzees (Riede *et al.*, 2007).

Vocal fold-vocal tract interaction

Most of the study of phonetics is based on the source filter theory (Fant, 1960). This theory describes speech as consisting of two, essentially independent components: a source of acoustic energy and a filter that selectively attenuates and amplifies certain sounds in order to create different signals. The source is usually provided by vibration of the vocal folds, or by turbulent noise produced at narrow constrictions in the vocal tract (other sources, such as grinding of teeth or lip smacks also occur, but are less frequent). The filter is provided by the supralaryngeal vocal tract: the pharyngeal, oral and nasal passages. Its exact shape determines which frequencies are amplified and which are attenuated, and this in turn determines how the produced signal is perceived. The fact that the source and the filter can be controlled separately allows for great flexibility in producing a large range of speech sounds.

Although the independence of source and filter is only an approximation (and this is well-known by phoneticians) it nevertheless works very well to understand human speech. It has been found that the importance of vocal fold-vocal tract interaction increases when a resonance frequency of the vocal tract is close to the frequency with which the vocal folds would vibrate if there would be no interaction. This is usually not the case in human speech, except for high vowels ([i], [u] etc.) in women (who tend to have smaller vocal folds which therefore vibrate at higher frequencies).

Another way in which the interaction between the vocal folds and the vocal tract can be increased is by decreasing the diameter of the epilaryngeal tube (Titze, 2008; Titze *et al.*, 2008). This is the part of the supralaryngeal vocal tract just above the vocal folds, and its diameter can be controlled by actions of the pharyngeal muscles. The reason that a narrower epilaryngeal tube increases interaction is that it results in larger velocities of the

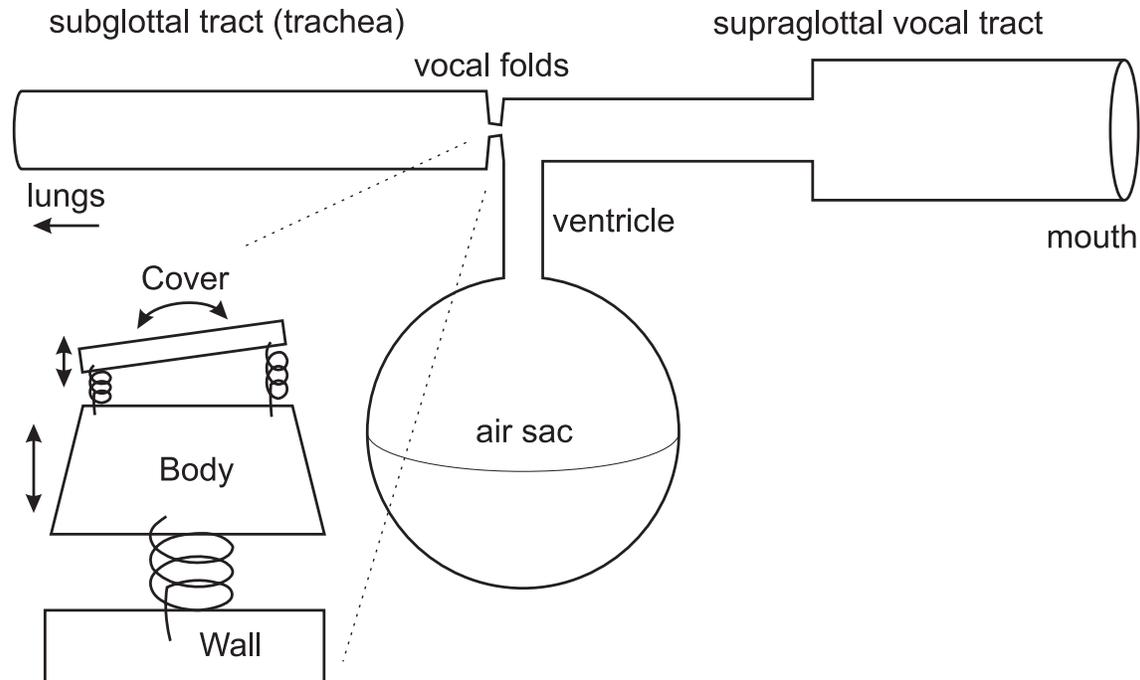


Figure 2: Schematic depiction of the vocal tract model with air sac and the vocal fold model.

air flowing from and to the vibrating vocal folds, thus increasing the forces that act on them.

Interaction between the vocal folds and the vocal tract can have a number of effects. It can change the amount of subglottal pressure required to make the vocal folds vibrate, even rendering it almost impossible to make the vocal folds vibrate at all (Riede *et al.*, 2008; Titze, 2008). Titze (Titze, 2002, 2008) has found that the right combination of subglottal and supraglottal impedances can help to increase output of acoustic power (for tracts without air sacs, but there is no reason why his findings would not be valid for tracts with air sacs). Frequency of vibration is also influenced, such that sudden frequency jumps can occur when the resonance frequencies of the vocal tract cross the vocal fold vibration frequency (Titze, 2008; Titze *et al.*, 2008). These effects can become even more complicated, to the extent that interaction can lead to period doubling or tripling (subharmonics) and even chaotic vibration (Hatzikirou *et al.*, 2006; Titze, 2008).

Air sacs add an extra low-frequency resonance to the acoustic response of the supralaryngeal vocal tract (de Boer, 2008c; Riede *et al.*, 2008; de Boer, 2009). For air sacs of the size as those found in apes, this frequency is in the range of 100–200 Hz. This is very close to the frequency of vocal fold vibration and this would *increase* vocal fold-vocal tract

interaction. On the other hand, air sacs provide an extra output for air flowing through the vocal folds, and as (in lateral ventricular air sacs, as found in apes) the entrance of the air sac is very close to the glottis, this would serve to reduce the acoustic impedance of the epilaryngeal tube and thus to *decrease* air sac-vocal tract interaction. Riede et al. (2008) do find interactions, but it is worthwhile to explore a somewhat larger range of possibilities than they have explored, in order to understand the importance of vocal fold-vocal tract interactions in the acoustic effect of air sacs.

Model and Methods

A computer simulation of vocal tracts with air sacs was used, as using a computer model allows for the largest flexibility in exploring different tract configurations and vocal fold parameters. It is possible to construct real-world physical models of the interaction between vocal tracts with air sacs and vocal fold vibration. This was done (in conjunction with computer simulation) by Riede et al. (2008). However, physical models either use artificial vocal folds made from some flexible material or use vocal folds that have been excised from a dead animal. In both cases, it cannot be expected that the real-world models result in a better approximation to living vocal folds than a computer model would. Of course, the computer model needs to be a simulation of the real vocal folds that has independently been shown to result in realistic behavior.

The vocal fold model that is used here is a reimplementation of Titze's (2002) body-cover model of the vocal folds. Instead of the two coupled masses that are usually modeled in vocal fold models (Dudgeon, 1970; Ishizaka & Flanagan, 1972), it models translational motion of the vocal fold body and translational as well as rotational motion of the vocal fold cover. The model is illustrated in figure 2. It was felt that the body-cover model is a slightly more elegant way of describing the physics of real vocal folds (which also consist of a body and a cover with different elastic properties) than the two-mass model, although they are likely equivalent in their behavior. The model has six degrees of freedom: body translation and translational velocity, cover translation and translational velocity, cover rotation and cover rotational velocity. Parameter settings were as in Titze (2002), unless otherwise stated.

Both the trachea and the vocal tract were modeled using the electrical line analog from Flanagan (1965, section 3.2). The trachea was taken to be 12 cm long, and to have a

diameter of 2 cm (Titze, 2008) while the upper vocal tract had a length of 16 cm and a constant diameter of 2 cm (unless otherwise stated) thus modeling an approximation of the vowel [ə]. The trachea was modeled with 4 circuits, each representing a 3 cm long tube, while the upper vocal tract was modeled as six sections that were all equally long ($2\frac{2}{3}$ cm). Damping due to viscous losses in the air flow and due to thermal conduction was modeled, but flexibility of the walls was not modeled. Flanagan's original specification is for harmonic oscillations at a specified frequency, and its parameters for damping were frequency-dependent. As calculations in the present model were done in the time domain, a fixed value for the damping parameters had to be used. It was decided to calculate these for 100 Hz. This resulted in somewhat stronger damping at high frequencies than for the values specified in Maeda's (1982) time domain model, but subjectively judging vocalizations produced by our model indicate that damping was still less than realistic for higher frequencies (the vocalizations still had a slight "ring" to them). Radiation at the mouth was also incorporated using Flanagan's (1965) model.

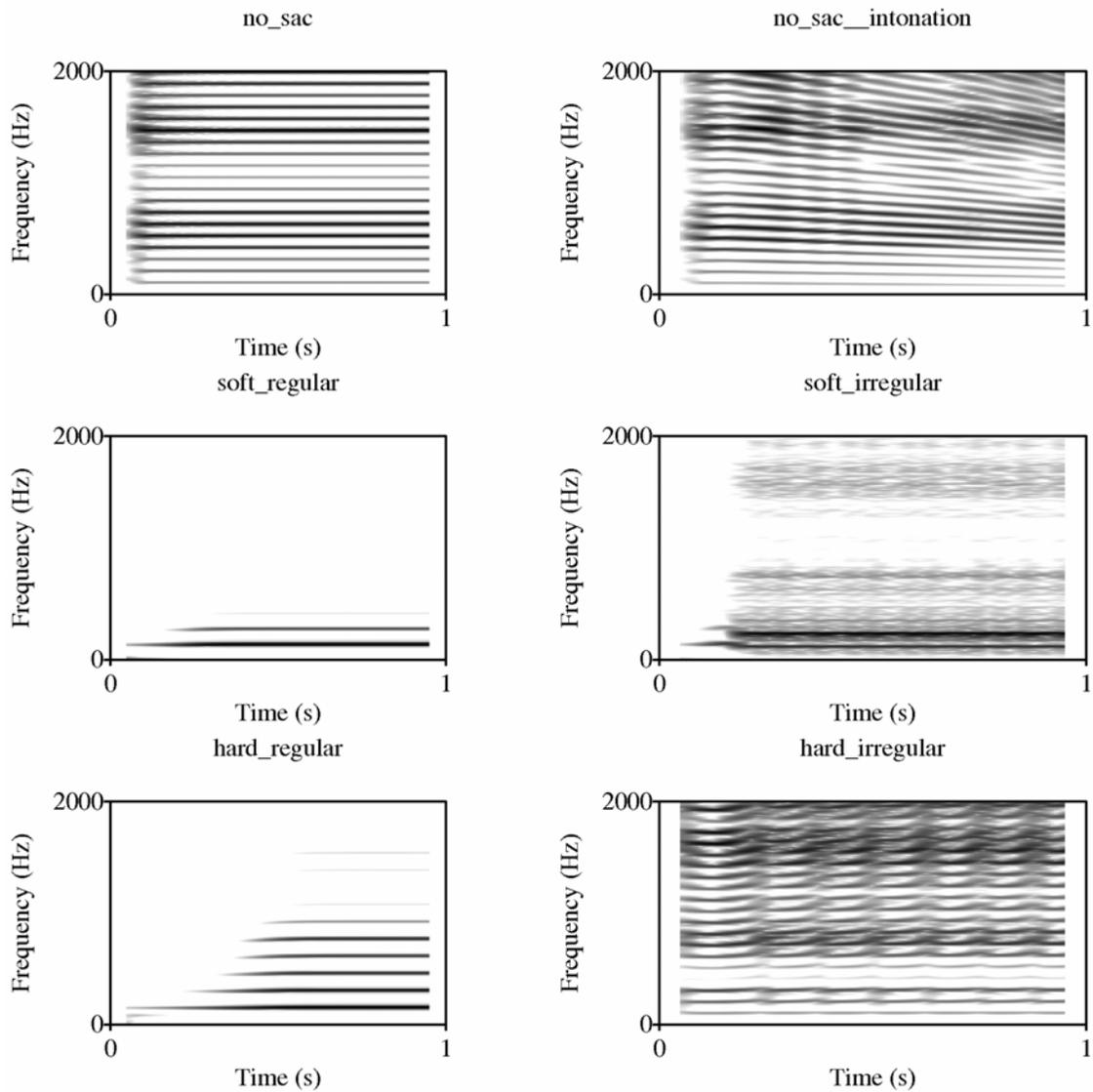


Figure 3: Spectrograms of different vocalizations generated with the model. From top to bottom and from left to right: no air sac without intonation (at default settings), no air sac with intonation (created by manipulating vocal fold elasticity), soft air sac with regular phonation (at 500 Pa lung pressure), soft air sac with irregular phonation (at 800 Pa lung pressure), hard air sac with regular phonation (at 2000 Pa lung pressure) and hard air sac with irregular phonation (at 2000 Pa lung pressure and with $r=2.28$ cm).

The air sac was modeled using de Boer's (2009) model. It has a neck length of 3 cm, a neck diameter of 1 cm, and is spherical with a radius of 5 cm (unless otherwise specified). This model is only valid for low frequencies because at higher frequencies wave propagation in the air sac cavity can no longer be approximated by a one dimensional wave. However, as

the interactions between vocal folds and the upper vocal tract are dominated by the lower frequencies, and as the chaotic nature of the vocal fold vibrations is also already apparent at lower frequencies, a low-frequency approximation was sufficient for the purposes of this paper.

Acoustic output was determined by calculating volume velocity of the flow from the mouth and from the air sac surface. These were added in order to calculate total volume velocity, thus assuming that the distance between the air sac and the mouth is negligible compared to the wavelength of the radiated sound. Again this holds for the frequencies that are relevant for vocal fold-vocal tract interactions.

The complete model had 32 degrees of freedom: 6 due to the vocal fold model, 8 due to the model of the trachea, 5 due to the air sac model, 12 due to the model of the upper vocal tract and one due to the radiation at the lips. It was simulated using the Matlab (version 7.3.0) ode45 function, which is based on a fourth/fifth order Runge-Kutta method using the Dormand-Prince pair (Dormand & Prince, 1980) with default tolerances. In cases where the ode45 function did not give satisfactory results (*out of memory errors* were sometimes generated because the required tolerances could not be met) the ode15s function (Shampine & Reichelt, 1997) was used. This function is more suited for stiff problems. It was used for the simulations with hard-walled air sacs and for the simulations of higher lung pressure in the system without an air sac.

The duration of the simulated signals was two seconds. The first second was discarded in order to get rid of initial fluctuations due to onset of vocalization. At the very beginning of the signal, lung pressure was increased from 0 to the desired value in 0.05 seconds, using a half period of a cosine, in order to prevent an abrupt transition (and the accompanying problems for numerical simulation). Five thousand equally spaced samples were taken from $t = 1-2$ s of the calculated signal, and on the basis of this the correlation dimension (such as used by Kakita & Okamoto, 1995) of the phase portrait of the six degrees of freedom of the vocal folds was calculated. This was done separately for the first 2500 samples and for the last 2500 samples in order to get an estimate of how stable the measure was over time. Before calculating the correlation dimension, the phase portrait was scaled so that the range of all degrees of freedom was between 0 and 1 (this was done for all 5000 samples simultaneously).

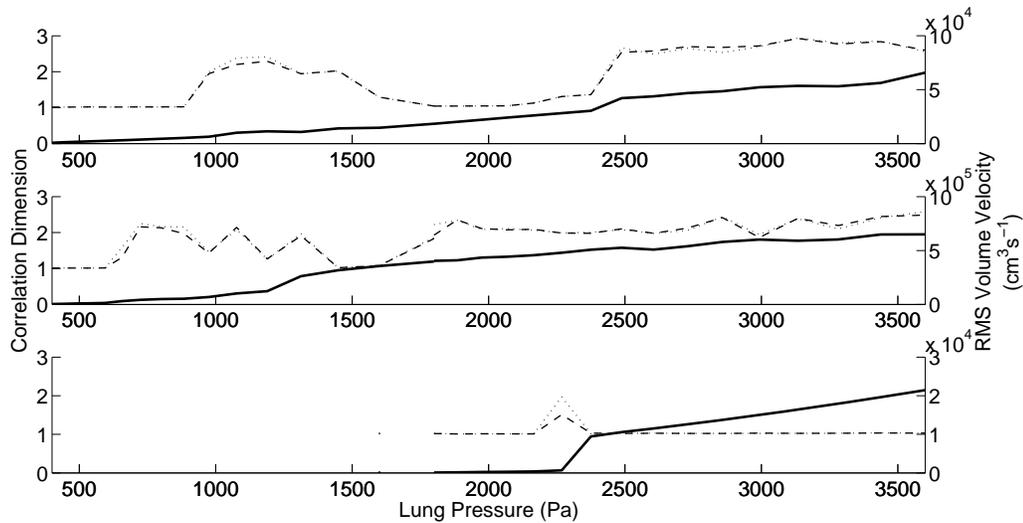


Figure 4: Phonation regularity and intensity versus lung pressure. The upper graph is for tracts without air sac, the middle for a tract with a soft air sac and the lower graph for a tract with a hard air sac. The dotted and dashed lines indicated the correlation dimension of the phase space of the vocal fold vibration. The dotted line is for $t= 1.0\text{--}1.5$ seconds and the dashed line for $t=1.5\text{--}2.0$ seconds of the calculated signal. A value of 1 indicates regular vocalization. The solid line indicates the strength of the sound produced.

Experiments

The influence of different parameters of the system on irregularity (chaoticness) of voicing was investigated. In order to get a first impression of what regular and irregular vocalizations with and without air sacs look like, a number of spectrograms is presented in figure 3. Regular vocalizations are characterized by regularly spaced harmonics of the vocal fold vibration. Irregular (chaotic) vocalizations still have some of this regular structure, but also show energy spread out over the spectrum. It is intriguing to note that in the regular vocalizations with air sacs, most energy is concentrated at the resonance frequency of the air sac, while with irregular vocalizations, relatively more energy appears to be present at the higher frequencies. The longer onset times of the regular vocalizations with air sacs is an artifact of the fact that lung pressures were taken from the lower range at which vocalization occurred.

Riede et al. (2008) have observed that lung pressure is a very important factor in whether vocal fold vibration takes place, and because it can also determine whether vibration is regular or not, a first experiment was run in which lung pressures between 400 and 3600 Pa

were investigated for tracts without and with hard-walled and soft-walled air sacs. The hard-walled air sac had a stiffness of 23 GPa and a quality factor Q of 10 (modeled after the perspex material from which earlier experimental models (de Boer, 2009) were made). The hard-walled models had a wall stiffness of 90 KPa and a quality factor of 1, modeled after real soft tissue (Fant, 1972; Maeda, 1982).

The results of lung pressure on regularity of voicing are presented in figure 4. It can be observed from this figure that regularity of voicing does not just depend on the resonance properties of the vocal tract: for all air sac types, regularity of voicing also depends on lung pressure in a non-linear way. It is not the case that below a certain lung pressure, there is regular voicing and above a certain lung pressure, there is irregular voicing. For all air sac types there are regions where voicing is regular, surrounded by regions where voicing is irregular.

It can also be observed that for tracts with a hard-walled air sac attached, the pressure required for phonation to take place at all is much higher than for tract without air sacs or for tracts with a soft-walled air sac. Whereas for the latter two types, phonation occurred at lung pressures as low as 400 Pa, for the tract with the hard walled air sac, a minimal lung pressure of 1800 Pa was required. This is probably due to the fact that for the dimensions of the air sac selected for this experiment, the resonance frequency of the hard-walled air sac was very close to the preferred frequency of vibration of the vocal folds (about 120 Hz, more about this below).

Finally, it can be observed that acoustic output of the soft-walled air sac is an order of magnitude higher than that of either the tract without the air sac or the tract with the hard-walled air sac. This has to do with the fact that the soft-walled air sac functions as a much more efficient radiator of sound around the frequency at which the vocal folds vibrate.

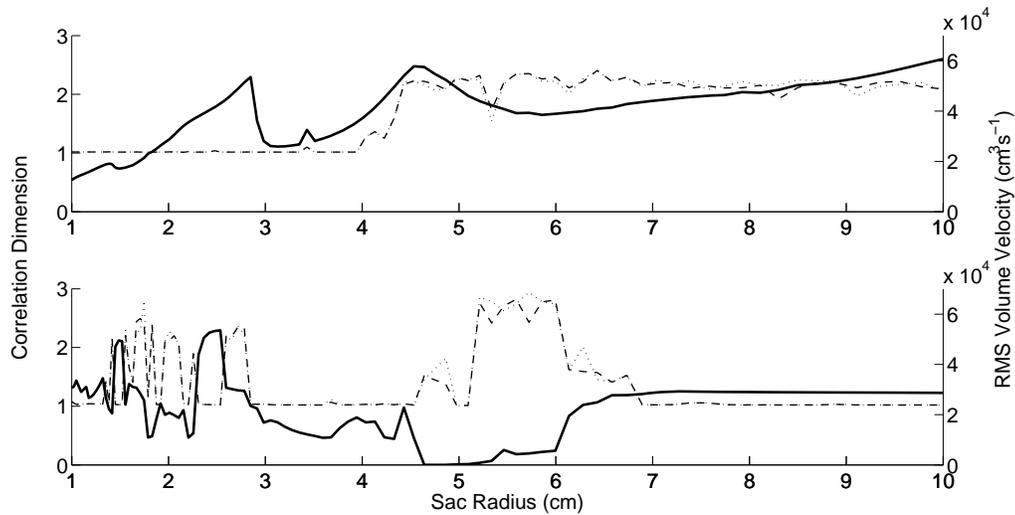


Figure 5: Correlation dimension of the phase space of vocal fold vibration and output intensity for vocal tracts with soft-walled (upper graph) and hard-walled (lower graph) air sacs. Dotted lines show the correlation dimension for $t=1.0\text{--}1.5$ s and dashed lines for $t=1.5\text{--}2.0$ s. A correlation dimension of 1 corresponds to regular phonation. Solid lines show output volume velocity (a measure of intensity).

Lung pressure therefore appears to be an important factor in determining whether voicing is regular, but the acoustic properties of the vocal tract are also important (Kakita & Okamoto, 1995; Hatzikirou *et al.*, 2006; Riede *et al.*, 2008; Titze, 2008; Titze *et al.*, 2008).

It is therefore necessary to investigate what happens when the properties of the vocal tract are changed at constant lung pressure. A range of air sac sizes from 1 cm to 10 cm radius was explored. As 800 Pa has been proposed as a normal lung pressure, it would have been ideal to explore all tracts at this lung pressure. However, tracts with hard-walled air sacs did not achieve phonation at this lung pressure. It was therefore decided to explore hard-walled sacs at 2000 Pa, near the minimal pressure required for phonation to take place. Soft-walled air sacs were explored at 800 Pa lung pressure.

The results of this exploration are presented in figure 5. It can be observed that for soft-walled air sacs a transition from regular to irregular phonation takes place at an air sac radius of about 4 cm. For hard-walled air sacs, the picture is more complicated. There are several regions where voicing is irregular: at a radius between 1.5 and 3 cm and at a radius of 5 to 6 cm. For larger hard-walled air sacs, voicing appears to become regular again. It can also be observed that for air sacs with a radius between 4.5 and 5.5 cm, phonation

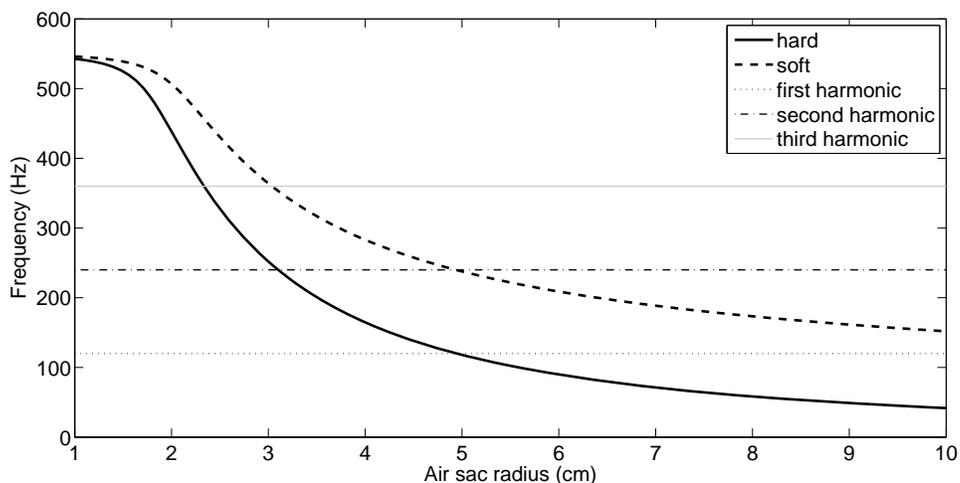


Figure 6: Resonance frequencies of hard-walled and soft-walled air sacs of the dimensions used in figure 5, and the first three harmonics of vocal fold vibration.

becomes very weak. Interestingly at a radius of 5 cm, phonation is regular, but as could be seen in figure 4, also very weak.

These effects are most likely related to the preferred frequency of vibration of the vocal cords being close to the resonance frequencies of the upper vocal tract. Figure 6 shows the frequency of the first resonance of the vocal tract with soft and hard-walled air sacs, as well as the first, second and third harmonics of the unconstrained (rotational) vibration of the vocal fold cover (which for the parameter settings from Titze (2002) is 120 Hz). It can be observed that the dip in output power of the hard-walled air sac occurs almost exactly at the location where the resonance frequency of the vocal tract crosses the preferred frequency of the vocal folds. This concurs with the observations by Titze et al. (Titze, 2008; Titze *et al.*, 2008) on tracts without air sacs. There also appear to be relations between irregular vocalization and the crossing of the third harmonic and the resonance frequency of the tract with hard-walled air sacs, and between onset of irregularity and crossing of the second harmonic and the resonance frequency of the tract with soft-walled air sacs. However, without further theoretical understanding it is difficult to assess whether these relations are accidental or not.

Conclusion

Interaction between vocal fold vibration and the vocal tract was modeled for vocal tracts with air sacs, and it was investigated how the properties of the upper vocal tract influence the regularity of vocal fold vibration. It was found that, for constant vocal tract shape, a first determinant of whether voicing was irregular, was lung pressure. Although it appears that in general, the larger the lung pressure, the more likely it is that vocal fold vibration becomes irregular, this is certainly not a simple relationship: there were cases of low lung pressure with irregular voicing as well as cases of high lung pressure with regular voicing for both soft and hard-walled air sacs, and for vocal tracts without air sacs.

The expectation that irregularity of voicing occurs when (first, second or third) harmonics of the preferred vocal fold vibrations are close to the resonance frequencies of the vocal tract appears to be borne out by the experiments presented here. However, given that the ranges of air sac size where vocalization is uniformly regular or irregular are somewhat broad, there is no very clear correlation between resonance frequency of the air sac and whether vocalization is irregular or not. Also, it appears that when the air sac resonance frequency drops below the first harmonic of vocal fold vibration (such as for hard-walled air sacs with radius larger than about 6 cm) interaction ceases, and vocalization becomes regular again.

Thus the picture of the relation between air sacs and regularity of voicing does not appear to be a simple one. However, for the evolutionarily most interesting case, that of large soft-walled air sacs (such as are found in gorillas, chimpanzees and orangutans) voicing appears to be irregular for large size ranges and large ranges of lung pressure. Although it should be kept in mind that the results presented here have been derived from a simplified model, it is likely that real air sacs would also result in destabilization of vocal fold vibration. Given that interaction increases with larger air flow through the glottis, this is understandable: the vibrating walls of the soft air sac periodically increase and decrease the volume of the air sac and therefore create important air flow, part of which interacts with the vocal folds.

It is possible that irregular vocalization helps to smear out acoustic energy over the spectrum, which as proposed by Fitch et al. (2002) would make it easier to locate the exact position of formant frequencies. On the other hand, the increased interaction between vocal folds and the vocal tract when a soft-walled air sac is present tends to concentrate the acoustic energy at the resonance frequency of the air sac (figure 3). At the same time, the

position of formants can also be clarified using frequency sweeps (intonation) as is usual in human speech (illustrated in figure 3). Therefore, the spread of acoustic energy would not seem to be a very big advantage (if it is an advantage at all). The increased irregularity of voicing would be a disadvantage for producing speech-like sounds, however, as for human speech, timing of onset of vocal fold vibration is very important, for example in the distinction between voiced and voiceless plosives.

The powerful low-frequency vocalizations that can be produced with soft-walled air sacs would help an animal to sound impressive (Fitch & Hauser, 2002) and to project its calls as far as possible in a jungle environment (Marten *et al.*, 1977). It is likely that this would confer an evolutionary advantage.

It has been shown in earlier work (de Boer, 2008b; Riede *et al.*, 2008; de Boer, 2009, 2010) that the effect of air sacs on the resonance patterns of the vocal tract appear to be advantageous for simple primate vocalizations, such as ones used in dominance interactions or used for territorial defense, but disadvantageous for creating subtle acoustic distinctions necessary for speech. From the work presented here, it also appears that the increased interaction between vocal fold vibration and the upper vocal tract is advantageous for producing loud, low-frequency calls, but disadvantageous for producing large sets of precisely timed, subtly distinct sounds. The investigation of the interaction between vocal fold interactions and vocal tracts with air sacs presented here, although preliminary and incomplete, supports the hypothesis that the evolutionary disappearance of air sacs is related to the emergence of complex vocal communication.

Acknowledgement

This work is part of the NWO vidi project “Modeling the evolution of speech” grant number 016.074.324.

References

- Aiello, L. C., & Dean, C. (2002). *Human evolutionary anatomy*. Amsterdam: Elsevier Academic Press.
- Alemseged, Z., Spoor, F., Kimbel, W. H., Bobe, R., Geraads, D., Reed, D., et al. (2006). A juvenile early hominin skeleton from Dikika, Ethiopia. *Nature*, *443*(7109), 296–301.
- Arensburg, B., Schepartz, L. A., Tillier, A. M., Vandermeersch, B., & Rak, Y. (1990). A reappraisal of the anatomical basis for speech in middle palaeolithic hominids. *American Journal of Physical Anthropology*, *83*(2), 137–146.

- Arensburg, B., Tillier, A. M., Vandermeersch, B., Duda, H., Schepartz, L. A., & Rak, Y. (1989). A middle palaeolithic human hyoid bone. *Nature*, 338(6218), 758–760.
- Avril, C. (1963). Kehlkopf und Kehlsack des Schimpansen, Pan troglodytes (Blumenbach 1799). (Mammalia, Primates, Pongidae). *Gegenbaurs morphologisches Jahrbuch*, 105, 75–129.
- Brown, B., & Ward, S. C. (1988). Basicranial and facial topology in *Pongo* and *Sivapithecus*. In J. H. Schwartz (Ed.), *Orang-utan biology* (pp. 247–260). Oxford: Oxford University Press.
- Camper, P. (1779). Account of the organs of speech of the orang outang. By Peter Camper, M. D. Late professor of anatomy, &c. in the University of Groningen, and F. R. S. In a letter to Sir John Pringle, F. R. S. *Philosophical Transactions of the Royal Society of London*, 69, 139–159.
- de Boer, B. (2008a). *The acoustic role of supralaryngeal air sacs*. Paper presented at the Acoustics '08, Paris, France.
- de Boer, B. (2008b). The joy of sacs. In A. D. M. Smith, K. Smith & R. Ferrer i Cancho (Eds.), *The evolution of language* (pp. 415–416). Singapore: World Scientific.
- de Boer, B. (2008c). *Modeling the acoustics of air sacs*. Paper presented at the Acoustic Communication by Animals, Corvallis, OR.
- de Boer, B. (2009). Acoustic analysis of primate air sacs and their effect on vocalization. *Journal of the Acoustical Society of America*, 126(6), 3329–3343.
- de Boer, B. (2010). The perceptual effect of air sacs. In A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith (Eds.), *Proceedings of the 8th international conference on the evolution of language* (pp. 395–396). Hackensack, New Jersey: World Scientific.
- Dormand, J. R., & Prince, P. J. (1980). A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics*, 6, 19–26.
- Dudgeon, D. E. (1970). Two-mass model of the vocal cords. *Journal of the Acoustical Society of America*, 48(1A), 118.
- Fant, G. (1960). *Acoustic theory of speech production*. 'sGravenhage: Mouton.
- Fant, G. (1972). Vocal tract wall effects, losses, and resonance bandwidths. *Speech Transmission Laboratory Quarterly Progress and Status Report*, 13(2–3), 28–52.
- Fitch, W. T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4(7), 258–267.
- Fitch, W. T. (2009). Fossil clues to the evolution of speech. In R. Botha & C. Knight (Eds.), *The cradle of language* (pp. 112–134). Oxford: Oxford University Press.
- Fitch, W. T., & Hauser, M. D. (2002). Unpacking "honesty": Vertebrate vocal production and the evolution of acoustic signals. In A. M. Simmons, R. R. Fay & A. N. Popper (Eds.), *Acoustic communication* (pp. 65–137). New York: Springer.
- Fitch, W. T., Neubauer, J., & Herzog, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, 63(3), 407–418.
- Flanagan, J. L. (1965). *Speech analysis, synthesis and perception*. Berlin: Springer.

- Frey, R., Gebler, A., Fritsch, G., Nygrén, K., & Weissengruber, G. E. (2007). Nordic rattle: The hoarse vocalization and the inflatable laryngeal air sac of reindeer (*rangifer tarandus*). *Journal of Anatomy*, 210(2), 131-159.
- Gautier, J.-P. (1971). Etude morphologique et fonctionnelle des annexes extra-laryngées des cercopithecinae; liaison avec les cris d'espacement. *Biologia Gabonica*, 7, 229-267.
- Haimoff, E. F. (1983). Occurrence of anti-resonance in the song of siamang (*hylobates syndactylus*). *American Journal of Primatology*, 5(3), 249-256.
- Harrison, D. F. N. (1995). *The anatomy and physiology of the mammalian larynx*. Cambridge: Cambridge University Press.
- Hatzikirou, H., Fitch, W. T., & Herzel, H. (2006). Voice instabilities due to source-tract interactions. *Acta Acustica united with Acustica*, 92, 468-475.
- Hayama, S. (1970). The saccus laryngis in primates. *Journal of the Anthropological Society of Nippon*, 78(4), 274-298.
- 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.
- Hilloowala, R. A., & Lass, N. J. (1978). Spectrographic analysis of laryngeal air sac resonance in rhesus monkey. *American Journal of Physical Anthropology*, 49(1), 129-132.
- Hombert, J.-M. (2010). Ecological and sexual explanations for larynx lowering. In A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith (Eds.), *The evolution of language (evolang 8)* (pp. 421-422). Hackensack (NJ): World Scientific.
- Ishizaka, K., & Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell system technical journal*, 51(6), 1233-1268.
- Kakita, Y., & Okamoto, H. (1995). Visualizing the characteristics of vocal fluctuation from the viewpoint of chaos: An attempt toward "qualitative quantification". In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 79-95). San Diego (CA): Singular Publishing Group.
- Kohlbrugge, J. H. F. (1896). Der larynx und die stimmbildung der quadrumana. *Natuurkundig Tijdschrift voor Nederlandsch Indië*, 55, 157-175.
- Maeda, S. (1982). A digital simulation method of the vocal-tract system. *Speech Communication*, 1(3-4), 199-229.
- Marten, K., Quine, D., & Marler, P. (1977). Sound transmission and its significance for animal vocalization: Ii. Tropical forest habitats. *Behavioral Ecology and Sociobiology*, 2(3), 291-302.
- Martínez, I., Arsuaga, J.-L., Quam, R., Carretero, J.-M., Gracia, A., & Rodríguez, L. (2008). Human hyoid bones from the middle Pleistocene site of the sima de los huesos (Sierra de Atapuerca, Spain). *Journal of Human Evolution*, 54, 118-124.
- Negus, V. E. (1949). *The comparative anatomy and physiology of the larynx*. London: William Heinemann Medical Books Ltd.

- Riede, T., Arcadi, A. C., & Owren, M. J. (2007). Nonlinear acoustics in the pant hoots of common chimpanzees (pan troglodytes): Vocalizing at the edge. *Journal of the Acoustical Society of America*, *121*(3), 1758–1767.
- 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.
- Schön, M. A. (1971). The anatomy of the resonating mechanism in howling monkeys. *Folia Primatologica*, *15*(1-2), 117-132.
- Schön Ybarra, M. A. (1986). Loud calls of adult male red howling monkeys (*alouatta seniculus*). *Folia Primatologica*, *47*(4), 204-216.
- Shampine, L. F., & Reichelt, M. W. (1997). The matlab ode suite. *SIAM Journal on Scientific Computing*, *18*, 1–22.
- Starck, D., & Schneider, R. (1960). Respirationsorgane. In H. Hofer, A. H. Schultz & D. Starck (Eds.), *Primatologia* (Vol. III (2), pp. 423–523). Basel: S. Karger.
- Stell, P. M., & Maran, A. G. D. (1975). Laryngocoele. *The Journal of Laryngology & Otology*, *89*, 915–924.
- Titze, I. R. (2002). Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *Journal of the Acoustical Society of America*, *111*(1 Pt 1), 367-376.
- Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *Journal of the Acoustical Society of America*, *123*(5), 2733-2749.
- Titze, I. R., Riede, T., & Popolo, P. (2008). Nonlinear source–filter coupling in phonation: Vocal exercises. *Journal of the Acoustical Society of America*, *123*(4), 1902–1915.