LARYNGEAL ARTICULATORY FUNCTION AND SPEECH ORIGINS

John H. Esling[1], Allison Benner[1], Scott R. Moisik[2]

[1]University of Victoria, [2]Max Planck Institute for Psycholinguistics, Nijmegen
esling@uvic.ca, abenner@uvic.ca, Scott.Moisik@mpi.nl

ABSTRACT

The larynx is the essential articulatory mechanism that primes the vocal tract. Far from being only a glottal source of voicing, the complex laryngeal mechanism entrains the ontogenetic acquisition of speech and, through coarticulatory coupling, guides the production of oral sounds in the infant vocal tract. As such, it is not possible to speculate as to the origins of the speaking modality in humans without considering the fundamental role played by the laryngeal articulatory mechanism. The Laryngeal Articulator Model, which divides the vocal tract into a laryngeal component and an oral component, serves as a basis for describing early infant speech and for positing how speech sounds evolving in various hominids may be related phonetically. To this end, we offer some suggestions for how the evolution and development of vocal tract anatomy fit with our infant speech acquisition data and discuss the implications this has for explaining phonetic learning and for interpreting the biological evolution of the human vocal tract in relation to speech and speech acquisition.

Keywords: laryngeal, larynx, vocal tract anatomy, infant speech, ontogeny

1. INTRODUCTION

The 'laryngeal articulator,' consisting of the glottal mechanism, the supraglottic epilaryngeal tube, the pharyngeal/epiglottal mechanism, and including three levels of folds – the vocal folds, the ventricular folds, and the aryepiglottic folds – is responsible for the generation of multiple source vibrations and for the complex modification of the epilaryngeal and pharyngeal resonating chambers that account for a wide range of contrastive auditory qualities. These qualities are observed in a surprisingly large number of the languages of the world, both linguistically and paralinguistically, and they account for sounds labelled in the IPA as 'pharyngeal' and 'epiglottal,' as various phonation types, as tonal register phonatory contrasts, or as vowel harmony secondary qualities. They reflect an expanding range of what have been known as the 'states of the glottis' (now more properly termed 'states of the larynx') [9, 14, 8, 23]. The laryngeal mechanism constitutes a significantly large and strategic portion of the vocal tract, as depicted in the 'Laryngeal Articulator Model' [10, 11], which has nevertheless been generally overlooked in considering the ontogeny and phylogeny of the phonetic capacity.

It has also been observed that infants, in their first months of life, produce a range of utterances, reflecting both phonatory possibilities and stricture types, that can be directly attributed to the laryngeal articulator mechanism. Systematic observation of infants' early speech production reveals that the control of articulatory detail in the pharynx is mastered during the first year of life [3, 13, 2, 18]. The control and growing understanding of manner of articulation in the pharynx (within the laryngeal mechanism) appears to be a prerequisite for expanding articulatory control into the oral vocal tract. Taking the larynx/pharynx as a starting point for the ontogenetic learning of the speech production capacity is likely to offer productive insights into the phylogenetic development of speech.

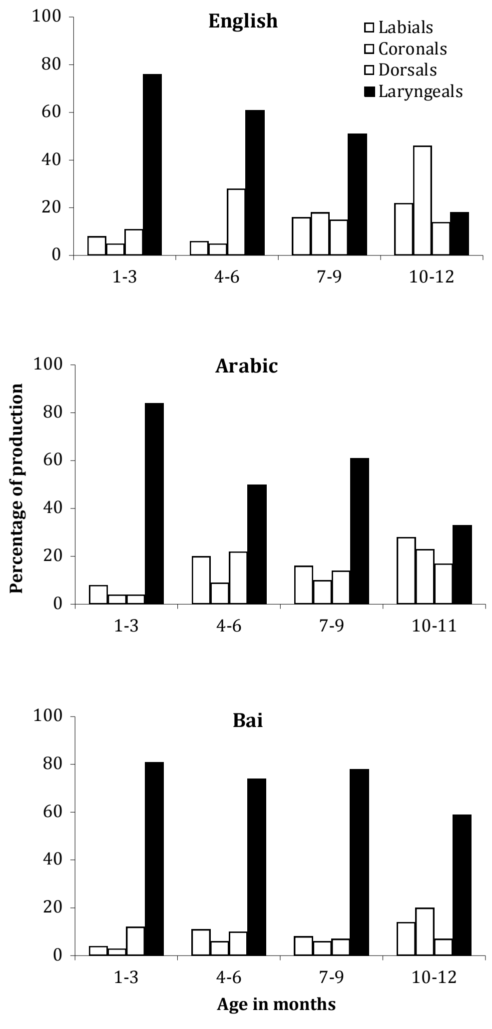
2. INFANT SPEECH ACQUISITION

2.1. Speech begins in the pharynx (with the laryngeal articulator)

Research into the earliest vocalizations by infants in English, French, Arabic, and Bai (Tibeto-Burman) contexts shows that: (1) speech begins in the pharynx, (2) the production of vocalic phonation and of consonantal stricture begins with laryngeally constricted settings, (3) infants actively explore their laryngeal phonetic production capacity through 'dynamic alternations' of paired contrasts, as those contrasts are discovered, and (4) infants often generate oral (lingual, labial, etc.) sounds with a primary laryngeal vocalization which precedes the oral articulation or is maintained as a coarticulation with the oral sound. Evidence from the Infant Speech Acquisition (InSpA) Project [12] illustrates instances of systematic 'phonetic play' that demonstrate how infants acquire basic control over the speech mechanism and the arrays of place and manner of articulation within the larynx during their first year of life.

Anatomically, laryngeal constriction is the first phonetic mechanism available to the infant, since a short (raised) and relatively flat laryngeal vocal tract is predisposed [1]. After (vocalic) crying with constricted (retracted) vowel quality, the 'first sound' that infants can be said to produce as an articulatory (consonantal) stricture is epiglottal stop [ʡ] [13, 12], which they do beginning from the first weeks of life. This stricture is a function of the laryngeal constrictor as the primary airway-protection reflex [16]. Glottal stop [ʔ], requiring more careful control than epiglottal stop, emerges later, early in the second month. Pharyngeal fricatives, approximants and trills appear early. Figure 1 shows the results of an analysis of 4,499 consonantal sounds produced by infants (English: 1,195; Arabic: 1,696; Bai: 1,608). The results clearly illustrate the prevalence of laryngeal sounds (including pharyngeal and glottal sounds) early in infancy and the increase in oral sounds throughout the first year in the production of infants from these three language groups.

Figure 1: Percentage of infants' production in terms of place of articulation according to infants' linguistic background and age group.



Chi-squared and Cramer's V analyses were performed on the consonantal data, split according to the different age groups (1-3, 4-6, 7-9, and 10-12 months) to test the strength of association between language and place of articulation for each of the four age groups. The results indicate that despite the significant association between language and place of articulation for all age groups (for all chi-squared results $p < .01$), the strength of the relationship between these two variables is very weak at 1-3 months (Cramer's V = .104), but considerably stronger at 10-12 months (Cramer's V = .239). These results suggest that as infants approach the end of their first year, their production becomes distinctive from one language group to another, presumably due to the influence of their ambient language. Early in infancy, the prevalence of laryngeal sounds illustrates our hypothesis that speech begins in the pharynx.

Similarly, phonatory configurations where laryngeal constriction dominates (harsh, whispery, and creaky voice) appear before unconstricted (modal, breathy, or falsetto) phonation. In the earliest months, laryngeally constricted production dominates in all languages observed. Analyses of an initial 3,197 utterances (English: 932; Arabic: 1,011; Bai: 1,254), contrasting only auditorily-evaluated constricted vs. unconstricted utterances across age groups, are significant (X^2 (3) = 93.34, $p < .001$), indicating that the incidence of laryngeal constriction in infants' vocalizations varies primarily as an inverse function of age, irrespective of linguistic background [1]. In all language groups, early vocalizations are overwhelmingly constricted, i.e. harsh, creaky, pharyngealized, raised-larynx, etc. As illustrated in Figure 2, the incidence of laryngeal constriction decreased progressively throughout the first year for infants from all three language groups examined, while still forming a major part of their vocal repertoire at the 10-12 month period. In summary, open-airway phonetic realizations occur only rarely until halfway through the first year. It could be said that laryngeally constricted qualities and strictures are reflexively innate, while open (less protective) qualities and strictures are learned.

Figure 2: Constricted and unconstricted voice quality settings produced by English, Arabic, and Bai infants.

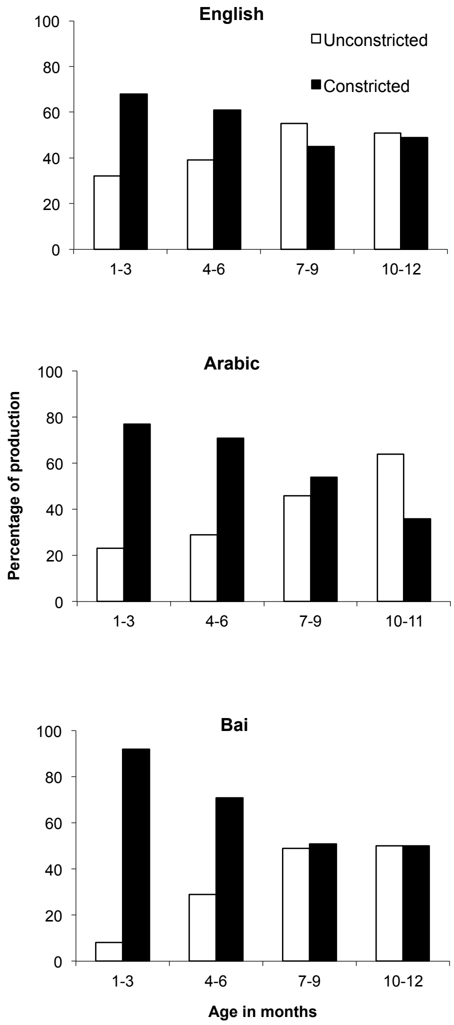
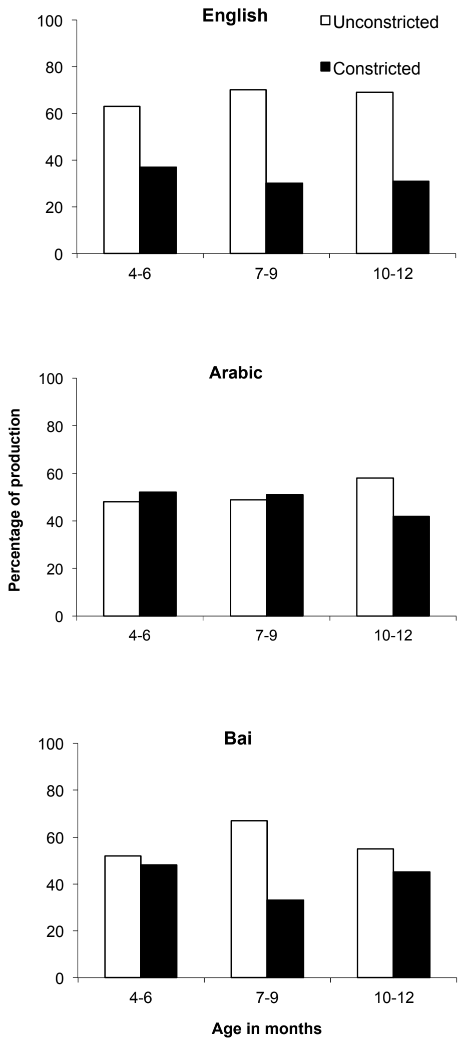


Figure 3: Constricted and unconstricted phonatory settings in the babbling of English, Arabic, and Bai infants.



2.2. Laryngeally constricted vocalization persists

Even during babbling, towards the end of the first year, when oral sounds become preferred, some constricted qualities persist, especially in those languages that contain pharyngeals (Arabic) or constricted registers (Bai) in their phonologies (Figure 3). For example, at the end of the first year, in months 10-12, only 31% of the babbling of English infants includes laryngeal constriction, compared to 42% and 45% for the Arabic and Bai infants, respectively.

Furthermore, as control over the articulators grows and oral strictures begin to be used, sounds that are learned at new oral places of articulation often occur with secondary 'accompaniments' from the original laryngeal articulator: coarticulatory events termed 'pharyngeal priming.' The preference in babbling for oral sounds may relate to the split between brain stem neural control and cortical neural control, where brain stem control can be posited to account for the reflexive emergence of the innate use of the laryngeal articulator and cortical control hypothesized to coincide with the shift from phonetic pre-babbling practice to the primarily oral control exhibited in the babbling stage.

3. EVOLUTIONARY ENTAILMENTS

In Burling's [5] account of the evolution of language and speech, the assertion that human speech sounds have conventional meaning rather than just being iconic from an early stage is supported by our evidence from phonetic ontogeny. What our research adds to the equation is that infants acquire motor control over contrastively useful parcels of speech at a surprisingly early age and in a reflexively rich but visually hidden part of the vocal tract. Any speculation about oscillatory patterns of articulators [22] needs to take into account that these patterns would have developed in the pharynx first, before they progressed to the mouth or the jaw. This provokes speculation about early hominids. If speech sounds develop ontogenetically beginning in the pharynx, as our research has shown, then this invites the possibility that speech sounds could have developed phylogenetically in the pharynx. At the very least, the laryngeal articulator capabilities of early hominid anatomy need to be considered. In reviewing accounts of language evolution such as Burling's, it is important to recognize that the agents of acquisition and change are infants in both cases rather than adults. That is, the speech capacity did not start with an early hominid who had already reached adulthood. Speech representations in every epoch begin with infants, from day one, acquiring phonetic production capabilities in a systematic progression from the larynx/pharynx outwards. At some point in time, infants gained the awareness that their own auto-generated sounds could be used for symbolic meaning. These stimuli would for a time be reflexive, eventually if only occasionally being responded to by an adult (most likely in indirect ways) and reinforced in various directions. In our methodology, it has become clear that adults become intensely aware of the human sound-producing capability when they have infants who are generating the basic elements of phonetic motor production during the first several months of life. The elements become familiar to the adults, but the infant is the driving force; i.e. the sounds are created by each infant, in a logical progression of how sounds can be produced in the pharynx, rather than being 'taught' to the infant. That is, we all learn phonetics 'experimentally' [cf. 6].

The crux of the issue is: if contemporary infants start phonetic acquisition with the laryngeal constrictor mechanism as first articulator, then how far back along the evolutionary path has this been the case? Early hominid infants, once they had the required cognitive criteria for language development that Burling enumerates, could be expected to have generated sounds similar to those pharyngeal sounds that every infant generates today, which have the potential to represent linguistic meaning, and which the infant 'discovers' as having that potential. The mechanism for drawing the phonetic and the semantic processes together would have likely been precisely because of the infant-adult interaction. Burling's observation that 'it is the parent, not the child, who is the imitator' [5:110] is given support by our observations of each infant's autogeneration of laryngeal contrasts (at the purely phonetic level) and remarkable control and early mastery of the innate laryngeal sound-producing instrument.

4. PHYLOGENY, DISCUSSION

4.1. Anatomy and laryngeal articulation

A great deal of attention has been placed on the size of the laryngeal vocal tract in speculation about the phylogenetic substrate necessary for the emergence of speech [15, 20, 21]. The main thrust of this discussion has focused on the proportioning of the oral and pharyngeal cavities (the horizontal/vertical supralaryngeal vocal tract ratio or SVTh/SVTv) in relation to potential phonetic vowel categories and their degree of quantality, *sensu* Stevens [27]. Recently, Boë et al. have asserted the importance of forming oral consonantal stricture [4], and the suggestion of biomechanical limitations on the chimpanzee tongue has been made in favour of this account [28].

The evidence that the larynx is the first domain of phonetic exploration adds yet another degree of complexity to the question of how speech may have evolved. While the human larynx is indeed situated low within the vocal tract, the descent during ontogeny of the laryngeal cartilages relative to the hyoid bone follows a remarkably similar pattern to that observed for chimpanzees [25]. This is thus a phylogenetically old component of the anthropoid vocal tract's developmental sequence; in most other mammals, the hyo-thyroid complex remains bound together and inhibits independent lingual-laryngeal control [19]. The relatively high early position of the larynx relative to the hyoid might have a protective function during infant cry vocalizations (mostly associated with mother-infant separation in non-human primates [24]). This might operate through the action of non-linear source-filter coupling [29], which, when there is substantial epilaryngeal narrowing, serves to increase the acoustic efficiency of the vocal folds. This has the benefit of reducing vocal fold stresses during crying vocalization while still generating a vocalization sufficiently intense to attract the attention of the caregiver (cf. [23]). The naturally constricted larynx also offers other enhancements to the attention-getting function of cry through, for example, the perturbation to phonation (harshness at the vocal fold level) or the accompaniment by extraglottal vibrations associated with the epilarynx, such as those of the ventricular folds, the aryepiglottic folds, or the epiglottis.

As has been shown in our research, this predisposing positioning of the larynx relative to the hyoid bone provides the grounds for the acquisition of the first consonantal stricture (the epiglottal stop) and for the development of manner of articulation (through manipulation of stop, approximant, fricative, and trilling phonetic postures). The pre-constricted posture also has other benefits for phonetic learning. A major challenge in understanding the acquisition of the complex motor control of speech [17] is how the innumerable degrees of freedom of the articulators are mastered. Early hyo-laryngeal approximation and its constraining of infant vocalizations initially to laryngeally constricted sounds serves to reduce considerably the search space of learning the motor control mechanisms behind producing different forms of consonantal strictures. We suspect that these laryngeally enacted processes constitute an early cortical mapping for manner categories upon which oral manners can be developed.

4.2. Unlocking the oral articulators

The other essential component of phonetic behaviour is the development of oral-laryngeal coarticulation, which is critical in the formation of voicing contrasts on obstruents and is essential in the production of tonal and intonational patterns. As the human vocal tract develops [26], the horizontal (i.e. oral) component exhibits a sudden spurt of growth which then nearly halts towards the end of the second year, having attained approximately its pre-adolescence scale. By comparison, there is ongoing growth of the laryngeal vocal tract throughout early childhood, which ultimately gives rise to the characteristic separation between hyoid bone and palate. By comparison, the oral vocal tract of the chimpanzee shows a much faster rate of growth than the laryngeal vocal tract. We might suspect that these continuously changing proportions of the vocal tract would offer some difficulty to the early establishment of place of articulation categories. Whatever ultimately drove the development of a flattened facial profile in humans, we suspect it offers a great advantage for phonetic learning, at least over the chimpanzee vocal tract, by being relatively stable during the post-babbling period (during the second year of life).

It is roughly at the end of the first year, once our larynx has gone through the first crucial 7-8 months of descent in relation to the hyoid bone, that the post-laryngeal phase of phonetic learning begins. By this point we can think of the oral articulators as being ‘unlocked’. The infant now has the challenge of learning to control many more degrees of freedom for phonetic purposes but can draw on control schemes in place for functions such as suckling (control of the lips and the tongue) and swallowing (control of the lips, tongue, soft palate, and larynx) juxtaposed against the cortical setting established for the control of basic phonetic categories of manner of articulation. The vocal behaviour of our primate cousins does not seem to include or at least favour these consonantal properties, being instead characterized primarily by modulation of vowel and phonatory qualities.

5. SUMMARY

The efficacy of vocalization as a social tool is ancient in the primate clade. Humans have taken the remarkable step of exploiting vocalization for the purposes of communication, and, as the predominant modality of human language, it is hard to believe that the need to acquire and use speech did not have some selective effect on our biology. With that stated, it is also the case that those components of ontogeny relating to the position and posturing of the larynx, which we have argued are an essential component of our phonetic learning and capacity, were already in place before language appeared. It strikes us as highly plausible that hominids with which we share much in common, such as Neanderthal [7], had phonetic capacity far in excess of that ascribed to them by some [21]. If the phylogenetic reduction of oral cavity length is really as important as has been suggested [26], we would speculate that the use of laryngeally constricted postures/sounds might have played an even more central role in modulating vowel qualities in Neanderthal phonologies than in those of humans today.

We have ultimately argued that the laryngeal vocal tract is the locus of phonetic exploration and that it would seem that the sequence of phonetic acquisition takes advantage of this initially predisposed constricted posture of the larynx and on its subsequent unlocking. The overall process of phonetic acquisition is thus interacting with an already-in-place sequence of events that unfold during post-natal development and, furthermore, might also have placed some selective pressure on the shape and developmental sequence of the vocal tract itself.

5. REFERENCES

[1] Benner, A. 2009. *Production and Perception of Laryngeal Constriction in the Early Vocalizations of*

Bai and English Infants. PhD dissertation, University of Victoria.
[2] Benner, A., Grenon, I., Esling, J. H. 2007. Infants' phonetic acquisition of voice quality parameters in the first year of life. *Proc. 16th ICPhS* Saarbrücken, 2073–2076.
[3] Bettany, L. 2004. *Range Exploration of Pitch and Phonation in the First Six Months of Life*. MA thesis, University of Victoria.
[4] Boë, L.-J., Badin, P., Ménard, L., Captier, G., Davis, B., MacNeilage, P., Sawallis, T. R., Schwartz, J.-L. 2013. Anatomy and control of the developing human vocal tract: A response to Lieberman. *J. Phonetics* 41(5), 379–392.
[5] Burling, R. 2005. *The Talking Ape: How Language Evolved*. Oxford: OUP.
[6] Catford, J. C. 2001. *A Practical Introduction to Phonetics*, 2nd ed. Oxford: OUP.
[7] Dediu, D., Levinson, S. C. (2013). On the antiquity of language: The reinterpretation of Neandertal linguistic capacities and its consequences. *Frontiers in Psychology* 4. doi:10.3389/fpsyg.2013.00397
[8] Edmondson, J. A., Esling, J. H. 2006. The valves of the throat and their functioning in tone, vocal register, and stress: Laryngoscopic case studies. *Phonology* 23, 157–191.
[9] Esling, J. H. 1996. Pharyngeal consonants and the aryepiglottic sphincter. *JIPA* 26, 65–88.
[10] Esling, J. H. 2005. There are no back vowels: The Laryngeal Articulator Model. *Canadian Journal of Linguistics* 50, 13–44.
[11] Esling, J. H. 2010. Phonetic notation. In: Hardcastle, W. J., Laver, J., Gibbon, F. E. (eds), *The Handbook of Phonetic Sciences,* 2nd ed. Oxford: Wiley-Blackwell, 678–702.
[12] Esling, J. H. 2012. The articulatory function of the larynx and the origins of speech. *Proc. Annual Meeting Berkeley Linguistics Society* 38, University of California, Berkeley. (via *eLanguage* of the LSA)
[13] Esling, J. H., Benner, A., Bettany, L., Zeroual, C. 2004. Le contrôle articulatoire phonétique dans le prébabillage. *Actes des XXVes Journées d'Étude sur la Parole*, 205–208, Fès, Maroc: AFCP.
[14] Esling, J. H., Harris, J. G. 2005. States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In: Hardcastle, W. J., Beck, J. M. (eds), *A Figure of Speech: A Festschrift for John Laver*. Mahwah, NJ: Lawrence Erlbaum, 347–383.
[15] Falk, D. 1975. Comparative anatomy of the larynx in man and the chimpanzee: Implications for language in Neanderthal. *Am J Phys Anthropol.* 43(1), 123–132.
[16] Gauffin, J. 1977. Mechanisms of larynx tube constriction. *Phonetica* 34, 307–309.
[17] Gick, B., Stavness, I. 2013. Modularizing speech. *Frontiers in Cognitive Science* 4, 1–3.
[18] Grenon, I., Benner, A., Esling, J. H. 2007. Language-specific phonetic production patterns in the first year of life. *Proc. 16th ICPhS* Saarbrücken, 1561–1564.
[19] Honda, K. 1995. Laryngeal and extra-laryngeal mechanisms of F0 control. In: Bell-Berti, F., Raphael, L. J. (eds), *Producing Speech: Contemporary Issues: For Katherine Safford Harris*. Woodbury, NY: American Institute of Physics, 215–232.
[20] Lieberman, P., Crelin, E. S. 1971. On the speech of Neanderthal Man. *Linguistic Inquiry* 2(2), 203–222.
[21] Lieberman, P., McCarthy, R. C. 2015. The evolution of speech and language. In: Henke, W., Tattersall, I. (eds), *Handbook of Paleoanthropology*. Berlin: Springer, 1–41. doi: 10.1007/978-3-642-27800-6_79-1
[22] MacNeilage, P. F. 1998. The frame/content theory of evolution of speech production. *Behavioral and Brain Sciences* 21, 499–511.
[23] Moisik, S. R. 2013. *The Epilarynx in Speech*. PhD dissertation, University of Victoria.
[24] Newman, J. D. 1985. The infant cry of primates. In: Lester, B. M., Boukydis, C. F. Z. (eds), *Infant Crying*. New York: Plenum Press, 307–323.
[25] Nishimura, T. 2003. Comparative morphology of the hyo-laryngeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. *Primates* 44(1), 41–49.
[26] Nishimura, T., Mikami, A., Suzuki, J., Matsuzawa, T. 2006. Descent of the hyoid in chimpanzees: Evolution of face flattening and speech. *J. Human Evolution* 51(3), 244–254.
[27] Stevens, K. N. 1989. On the quantal nature of speech. *J. Phonetics* 17, 3–45.
[28] Takemoto, H. 2008. Morphological analyses and 3D modeling of the tongue musculature of the chimpanzee (Pan troglodytes). *American Journal of Primatology* 70(10), 966–975.
[29] Titze, I. R. 2008. Nonlinear source–filter coupling in phonation: Theory. *JASA* 123(5), 2733.