A STUDY OF MUSCULAR SYNERGIES AT THE GLOTTAL, VENTRICULAR AND ARYEPIGLOTTIC LEVELS

John, H. Esling [1], Chakir Zeroual [2,3], Lise Crevier-Buchman [3,4]

[1] University of Victoria, [2] Faculté de Taza, Morocco, [3] CNRS-UMR7018, Paris III, [4] HEGP, Paris
esling@uvic.ca, chakirzeroual@yahoo.fr, lise.buchman@numericable.fr

ABSTRACT

Normal and high-speed laryngoscopic observations of Moroccan Arabic laryngeal consonants and of phonetically cardinal laryngeals show that: 'glottal stop' cannot be uniquely glottal but is at least glottoventricular; epiglottal (fricative) trills are produced with vibrations of the aryepiglottic folds (AF); epiglottal stop invokes extreme ventricular band (VB) adduction with AF engagement. From these data, we deduce that full VB adduction is a passive effect of partial laryngeal (aryepiglottic) constriction, separate in control from vocal fold (VF) adduction, and that active (total) laryngeal closure results from a progressive contraction of a complex of muscles, the most likely among them (for future study) the thyroarytenoid (TA) muscles.

Keywords: Glottal, ventricular, aryepiglottic, muscular synergy, laryngeal constrictor.

1. INTRODUCTION

This study introduces linguistic physiological data to determine if the laryngeal cavity incorporates only one functional level: glottal (vs. supraglottal), as in the classic literature; or at least three: glottal, ventricular, and aryepiglottic, as proposed in recent literature [2, 3]. We also show that aryepiglottic articulations are not performance styles abnormal to speech [5, 8] but are linguistically common.

The phonatory modes, breath (as in voiceless [h] articulations), voice, and falsetto are produced by adjusting glottal aperture, the tension of the vocal folds (VF), and aerodynamic parameters. The ventricular bands (VB) and the aryepiglottic folds (AF) remain abducted, establishing that the level of the glottis can be controlled (for opening, closing and lengthening) in an independent manner from the immediately supraglottic laryngeal structures.

Sakakibara et al. [9] describe singing voices produced with only VB adduction and only AF adduction. Réthi [7] describes a pathological voice with VB vibration, where the VF remain abducted, but hypotheses explaining VB (and AF) adduction mechanisms remain controversial.

In the laryngeal articulator model, consonants described as pharyngeal or epiglottal share ary-epiglottic constriction [3]. Although adduction of the AF has been attributed to the aryepiglottic muscles (AEm), reinforced by the thyroepiglottic muscles (TEm) [9], we suspect the thyroarytenoid muscles (TAm) are logical agonists, as laryngeal constriction and creaky (laryngealized) voice (VF shortening for low pitch) have similar postures.

According to most physiological analyses, the VB have no intrinsic muscular structure [9]; their adduction is accomplished passively due to glottal-level adductors and especially aryepiglottic-level adductors. Rhéti [7] claims the stylopharyngeal, aryepiglottic, and oblique interarytenoid muscles react concurrently to a single electrical stimulus without implicating the transverse interarytenoid muscles, concluding that these muscles constitute the 'stylopharyngeal muscular system,' predicting that the more the stylopharyngeal muscular system (and therefore the aryepiglottic level) contracts, the more the VB are adducted, and vice versa.

Other physiological analyses suggest that active contraction results from certain muscular fibres located inside the VB. Based on histological and EMG data, it has been suggested [5] that VB adduction can be obtained by contracting the TEm, especially its inferior part, the 'ventricularis' (VENm), whose fibres enter the VB. According to Reidenbach [6], the VB do contain muscular fibres that exert a force towards the interior (lateral-to-medial adduction) and downwards.

To address these controversies from an articulatory perspective, we present endoscopic data on attested segments of Moroccan Arabic (MA) and canonical laryngeal forms produced by a native speaker who is also a trained phonetician.

2. METHOD AND EQUIPMENT

One speaker of Moroccan Arabic, S1 (male, 38), participated in two physiological experiments.

2.1. Experiment 1

An Olympus ENF-P3 nasendoscope, connected to an Olympus CJH 250 camera (25 frames/sec), was inserted through the nostril of S1 and placed above the larynx. S1 pronounced 7 times a list containing:

- [t d ʔ ħ ʕ] in the words [itih idir fuʔad iħil iʕif] and [ʔ] in the nonsense word *[iʔi];
- stop and trills not attested in MA: [ʡ ʜ ʢ] in nonsense words of the structure *[iCi].
- [ħ ʕ ʜ ʢ], held so that airflow interruption was very slight, for comparison with [ʔ ʡ].

2.2. Experiment 2

The Olympus ENF-P3 endoscope was attached to a high-speed camera (500 frames/sec) to observe AF movement. Illumination at this speed is insufficient to visualize the VF but adequate to view the VB and AF levels, so only nonsense words containing [ħ ʕ ʜ ʢ] in the context *[iCi] were analyzed.

3. EXPERIMENTAL RESULTS

3.1. Sounds involving only glottal activity

During the production of modal [u i d], only the VF are adducted and vibrating along their entire length (Fig.1). During [t], the glottis opens along its length, with a maximal aperture much narrower than during breathing (B). During [i t d], the VB remain apart, and the aryepiglottic aperture is wide open (between: 1, arytenoids; 4, AF; 5, epiglottis).

3.2. Sounds involving the ventricular level

3.2.1. Full VB adduction

[ʔ] is produced in [fuʔad] and in *[iʔi] with a total adduction of the VB, preceded and followed by an adduction of the VF that is also certain to last throughout the full stop. Our data show that even a cessation of airflow that is a ‘lax’ version of [ʔ] (stop phonation (SP) [1]) always involves the VB. The aryepiglottic level is more compressed for [ʔ] (postero-anteriorly) than for [i u d t].

3.2.2. Partial VB adduction

Since, based on stroboscopy, only the anterior part of the VF vibrates in creaky voice [1], it is possible that creaky voice may involve some (partial) adduction of the VB in conjunction with postero-anterior aryepiglottic constriction [11]. Given the similarity of creaky voice (for pitch lowering) to [ʔ] (vibratory arrest) as states of the glottis [3, 4], it could be argued that only one element separates the two states. Slight aryepiglottic constriction typifies both the invariable presence of VB compression for [ʔ] and the reduction in VB compression for the production of creaky voice, so we hypothesize that the relaxed VB for creaky voice and compressed VB for harsh voice relate to an optional synergy between the glottal and aryepiglottic levels.

3.3. Sounds involving the aryepiglottic level

3.3.1. Full constriction

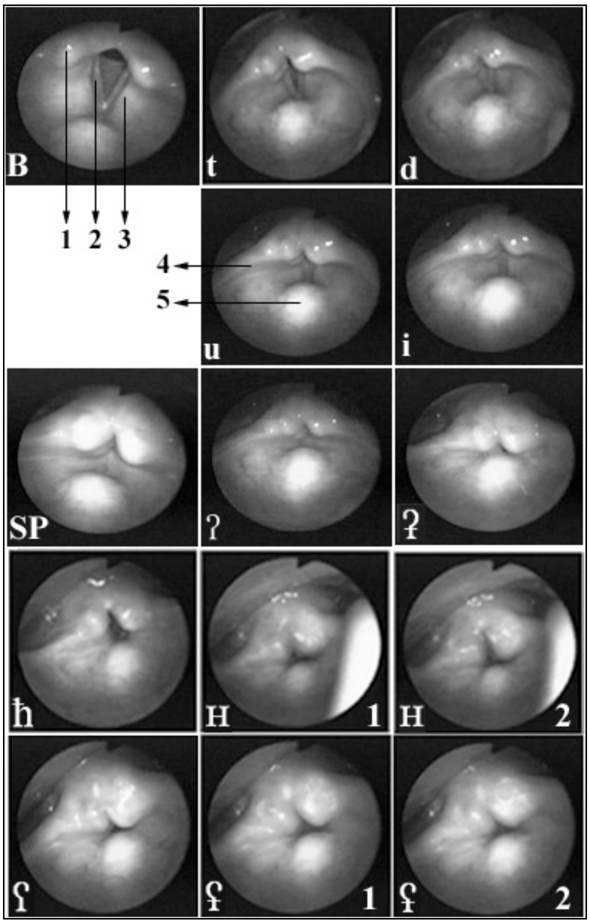
Fig. 1 illustrates that epiglottal stop [ʡ] has full occlusion between the arytenoid apices through the cuneiform cartilages and AF to the surface and sides of the epiglottis. Although tongue retraction and larynx raising are not extreme in this example, sphincteric AF-level closure is complete. This constriction mechanism has also been termed the third valve of the larynx in recent literature [3]. It is assumed that the VF and VB have also reached complete occlusion by the time [ʡ] is achieved.

3.3.2. Partial constriction

The aryepiglottic level during [ħ ʕ] is extremely compressed compared to [t d]. The arytenoids and bent AF are approximated to the epiglottal base, more tightly during [ʕ] than during [ħ]. During [ʕ], the VB are slightly adducted but do not touch. Neither the VF, VB (or AF generally) vibrate during [ħ], and only the VF vibrate during [ʕ].

While during [ħ] the arytenoids remain slightly separated (Fig. 1), the corresponding voiceless trill [ʜ] develops an alternating adhesion to the soft surfaces above the arytenoids. The VF remain abducted for [ħ ʜ]. The posterior part of the glottis appears more closed during [ʜ] compared to [ħ], in all likelihood in order to hold the upper portion of the tube tight enough together to permit the soft tissues to vibrate. Similarly, [ʢ] is also produced with a vibration of the tissues at the superior margins of the constricted supraglottic tube, but the degree of opening is smaller than for [ʜ]. It appears that only the VF and the AF margins vibrate, while the VB do not vibrate. As with MA [ħ ʕ], the glottis abducts for [ʜ], while glottal adduction for [ʢ] adds LCAm compression that presumably synergizes with what we believe to be the (AF-medial) TAm contraction that shuts the sphincter, as seen in the unphonated [SP(=ʔ̞) ʔ ʡ] series continuum in Fig. 1. The voicelessness of [ħ ʜ] (glottally open) clearly argues in favour of independent glottal control from the AF level.

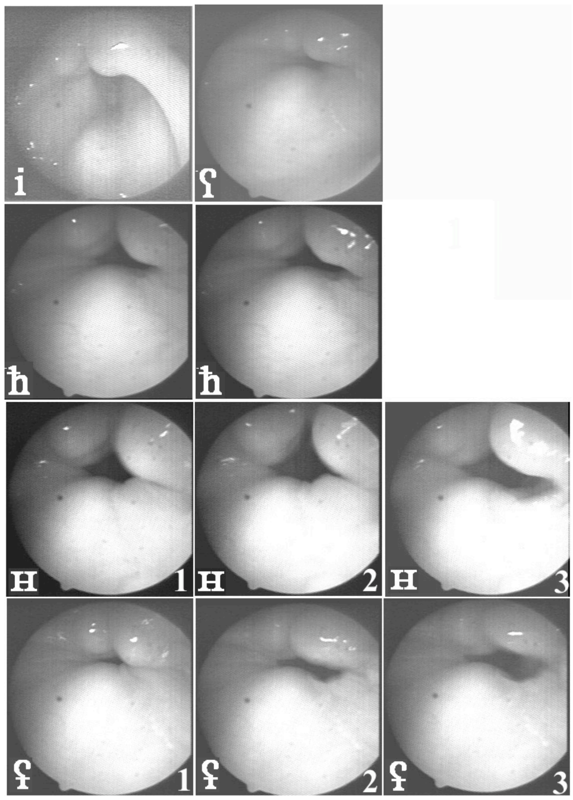
Figure 1: Endoscopic images (25 frames/sec) of the laryngeal cavity (top=posterior) during: B, breathing; SP, 'stop phonation' [1]; [i t d ħ ʕ] in [iCi-]; [u ʔ] in [-uʔa]; and [ʡ ʜ ʢ] in *[iCi-]. ([ʜ ʢ], 1 2): two sequential images. Labels: 1, cuneiform tubercles with arytenoid apices behind, medially; 2, vocal folds; 3, ventricular bands; 4, aryepiglottic folds; 5, base of epiglottis.



High-speed filming reveals that for [ʜ ʢ], not only do the tissues around the arytenoids vibrate, but the AF also vibrate laterally against the base of the epiglottis in an irregular and asymmetrical manner (Fig. 2). The effect is strongest at the right of the image (left AF). The high-speed camera shows that the AF vibrate during [ʢ] without separating interarytenoidally; while an alternating irregular separation of the upper sphincteric structures occurs during [ʜ]. It should also be noted that high-speed films reveal slight vibrations of the upper arytenoid-to-lateral AF margins even during articulations of [ħ] (presumably an aerodynamic effect), but not at all during articulations of [ʕ].

TAm activity is postulated for these pharyngeal strictures because the portions of the AF implicated in laryngeal constriction are the cuneiform (#1) tubercles, halfway along the AF, which are drawn forwards and upwards towards the surface of the epiglottis. It is a misapprehension that it is the arytenoid cartilages that come into contact with the epiglottis during aryepiglottic sphinctering [9]. A close examination of Fig. 1 reveals that it is the cuneiform 'elbows' in the AF that are drawn forwards, whether the sound is voiced or voiceless.

Figure 2: High-speed endoscopic images (500 frames/ sec) of the laryngeal cavity during [ħ ʕ] in [iCi-]; and [ʜ ʢ] in *[iCi-]. ([ʜ ʢ], 1 2 3): three sequential images.



4. DISCUSSION

Modal voicing requires moderate contraction of the lateral cricoarytenoid (LCAm) and lateral thyro-arytenoid (LTAm) muscles to permit the adduction of the ligamental glottis [10] (the LTAm likely dependent on pitch, we suspect). The inter-arytenoid muscle (IAm) aids and maintains closure of the cartilaginous glottis; while the cricothyroid muscles (CTm) prevent the shortening of the VF and excessive augmentation of VF mass during voicing. Only the posterior cricoarytenoid muscles (PCAm) accomplish full VF abduction.

Sawashima et al. stress that during [ʔ], 'the lateral cricoarytenoid muscle appears to show a high degree of activity' [10]. Sakakibara et al. [9] entertain the possibility that the LCAm is responsible for the adduction of the VB, since the electrode they placed under one VB is close to the LCAm. However, the EMG traces they present show almost no electrical activity at the level of this electrode during falsetto, contrary to what would normally be expected for high-pitched voicing [1, 4]. They report that the electrode may have been placed in the TEm (VENm) rather than in the LCAm. The fact that even slight interruptions of airflow are achieved between the VB suggests that their total adduction can occur without the contraction of the LCAm. We speculate that VB adduction is therefore facilitated by adduction at the glottal level (contraction of LCAm and LTAm: passive adduction) but would necessitate other contractions (VENm or other medial/inferior muscle fibres, e.g. portions of the TAm) for VB closure to be complete.

The aryepiglottic space is more compressed during [ʕ] than during SP or [ʔ]. Despite this, the adduction of the VB is stronger during the glottal stops than during [ʕ]. This contradicts Réthi's [7] hypothesis of a positive correlation between the degree of compression at the aryepiglottic level and the adduction of the VB. This observation supports the hypothesis by which the complete adduction of the VB happens in two stages: (1) a passive (partial) adduction, induced by VF adduction but also related to AF constriction, which approximates the VB; and (2) an active (total) adduction resulting from the contraction of muscles lying over the top of the VB, perhaps the VENm, but likely parts of the TAm.

Beyond their role as a glottal-arrest device ([ʔ]), VB compression appears to be stronger in voiced production [ʕ ʢ] than in voiceless production [ħ ʜ], suggesting that glottal adduction is synergistically related to the active compressions at the VB level (and AF level) that eventually close the VB over the VF. This activity must, nevertheless, remain separate from the superior AF-level cuneiform-to-epiglottis tightening that brings about 'laryngeal sphinctering' that causes pharyngeal stricture.

5. CONCLUSION

Our description of SP and [ʔ] suggests that a uniquely 'glottal' explanation is difficult to defend physiologically. The endoscopic data show that in addition to glottal adduction, an engagement of the aryepiglottic level and hence of the VB is required. [ʔ] seems therefore not to be solely 'glottal' but 'glottoventricular.' Massive laryngeal occlusion for epiglottal stop [ʡ] develops significantly greater VF and VB adduction with AF engagement. Our analysis also demonstrates that aryepiglottic trills are produced in linguistic contexts with a vibration of the AF soft-tissue structures at the superior margins of the epiglottic tube, from the cuneiform tubercles laterally against the undersurface of the epiglottis, logically as a function of increased aerodynamic flow.

Glottal adduction, abduction, and lengthening are controlled independently from AF sphinctering. Partial adduction of the VB is facilitated by glottal-level adduction (contraction of LCAm and LTAm) as well as by superior AF-level adduction (AF compression), but their total adduction would necessarily be active, requiring the contraction of an intrinsic muscle, most likely the VENm or other muscles medial and inferior to the cuneiform-thyroid plane, i.e. portions of the AEm. EMG studies based on the articulatory contrasts we have described are invited to resolve these issues.

6. REFERENCES

[1] Catford, J.C. 1968. The articulatory possibilities of man. In: Malmberg, B. (ed), *Manual of Phonetics*. Amsterdam: North Holland, 309–333.

[2] Catford, J.C., Esling, J.H. 2006. Articulatory phonetics. In: Brown, K. (ed), *Encyclopedia of Language and Linguistics* (2nd edn), vol. 9. Oxford: Elsevier, 425–442.

[3] 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.

[4] Laver, J. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

[5] Kimura, M., Sakakibara, K-I., Imagawa, H., Chan, R., Niimi, S., Tayama, N. 2002. Histological investigation of the supraglottal structures in humans for understanding abnormal phonation. *JASA* 112, 2446.

[6] Reidenbach, M. 1998. The muscular tissue of the vestibular folds of the larynx. *Oto. Rh. Lar.* 255, 365–367.

[7] Réthi, A. 1966. Dysphonie spastique, troubles de l'automatisme du langage. In: Moles, A., Vallancien, B. (eds), *Phonétique et phonation*. Paris: Masson, 113–140.

[8] Sakakibara, K-I., Fuks, L., Imagawa, H., Tayama, N. 2004. Growl voice in ethnic and pop styles. *ISMA*, Nara.

[9] Sakakibara, K-I., Kimura, M., Imagawa, H., Niimi, S., Tayama, N. 2004. Physiological study of the supraglottal structure. *ICVPB 2004*, Marseille.

[10] Sawashima, M., Hirose, H. 1980. Laryngeal gestures in speech production. *Ann. Bull. RILP* 14, 29–51.

[11] Zeroual, C., Esling, J.H., Crevier-Buchman, L. 2005. Physiological study of whispered speech in Moroccan Arabic. *Interspeech 2005 Proc.*, Lisbon, 1069–1071.