ABDUCTED VOCAL FOLD STATES AND THE EPILARYNX: A NEW TAXONOMY FOR DISTINGUISHING BREATHINESS AND WHISPERINESS

Scott R. Moisik[1], Míša Hejná[2], John H. Esling[3]

[1]Nanyang Technological University, Singapore
[2]Aarhus University, Denmark
[3]University of Victoria, BC, Canada
scott.moisik@ntu.edu.sg; misa.hejna@cc.au.dk; esling@uvic.ca

ABSTRACT

Regarding phonatory settings associated with abducted glottal states, under which fall phenomena such as breath, breathiness, breathy voice, murmur, noisy voice, slack voice, whisper, whisperiness, and whispery voice, researchers assume that the primary (and often the sole) difference between these derives from different degrees of glottal opening. This picture emerges from our overview of foundational documents in phonetics, including introductory as well as advanced textbooks, linguistic dictionaries, seminal papers, and specialised monographs.

The purpose of this paper is to draw attention to the treatment of these abducted glottal states in a theory of laryngeal articulation that demonstrates with articulatory evidence that the assumptions behind these phonetic texts do not adequately capture the articulatory reality, in particular, that they ignore the role of the epilarynx. Presenting a collection of novel research, we show that epilaryngeal function is essential in distinguishing the subtypes of abducted glottal states.

Keywords: abducted glottis, breathiness, whisper, whisperiness, epilaryngeal constriction.

1. INTRODUCTION

Phoneticians working with phonatory settings associated with abducted glottal states have been using a wide range of terms to describe and refer to these, including breath, breathiness, breathy voice, murmur, noisy voice, slack voice, whisper, and whispery voice [1, 4, 7, 10, 13, 17, 19, 20, 29, 34]. This has brought some amount of confusion in the field [10, 17 (49–50)], not only because of the number of terms used but also because a single term is sometimes used to refer to different phenomena (see section 1.2. below). The present paper therefore has the following goals.

Firstly, following Esling & Harris [10], we bring forward a taxonomy in which there are two main phonatory settings associated with abducted glottal states: breathiness and whisperiness (§1.1., but also [10 (350–1)]. Secondly, we set this taxonomy in the context of the existing literature (§1.2.). Finally, we review already existing evidence and present revised diagrams for laryngeal states (§2) and some novel, illustrative articulatory evidence in favour of the taxonomy proposed here (§3).

1.1. Defining breathiness and whisperiness

Breathiness and whisperiness can be thought of as two broader larynx states which subsume breath and breathy voice and whisper and whispery voice, respectively. What breathiness and whisperiness have in common from an articulatory point of view is that they both involve a glottal opening. However, this glottal opening is typically narrower for whisper, although variation exists in this regard [10, 34 (176)]. Crucially, whisperiness involves constriction of the epilarynx [10] – the upper two-thirds of the larynx found immediately above the vocal folds, with the ventricular folds forming the lower epilarynx, the aryepiglottic folds forming the upper epilarynx, and the epiglottis forming its anterior boundary – accompanied by the raising of the larynx, whereas breathiness is characterized by the absence of this supraglottal constriction and is usually accompanied by larynx lowering.

The epilarynx is the central structure constituting the laryngeal articulator, as described in the Laryngeal Articulator Model [9]. The laryngeal articulator, or laryngeal constrictor mechanism, is responsible for glottal and pharyngeal/epiglottal speech sounds, but its activity, ranging from widely unconstricted (as in inspiration) to massively constricted (as in epiglottal stop), accounts for a range of phonatory and vocal register phenomena and covariation between vowel quality and laryngeal state [8, 9, 11, 22, 23, 24].

The distinction between breath and breathy voice then lies in whether the vocal folds vibrate or not. Whilst breath does not include glottal vibration, breathy voice does. The same applies to the distinction between whisper and whispery voice, but with the addition of laryngeal constriction ([10, see also 9, 22]). Note that these categories are phonetic; the conversational register of whispering can be produced with breath and/or whisper.

1.2. Abducted glottal states in previous literature

Some phonetics researchers and teachers working with phonation focus on breathiness and have little to say about whisperiness, if anything at all [4, 13, 7, 29, 17]. At first blush, this may be due to the fact that breathiness would seem to be used for contrastive purposes more frequently than whisperiness, although there is no doubt that both are used widely for a range of paralinguistic purposes (e.g. [2, 5, 12, 14, 19 (135, 140), 18 (200), 21, 27, 30, 32, 33]), which is a good enough reason for the difference to be of relevance to linguists. Nonetheless, both breathiness and whisperiness can be contrastive with modal (and other) phonatory settings, noting also that voiceless pharyngeal fricatives are produced using a whispery laryngeal setting [1 (117), 10 (366–367), 19 (135)]. We therefore conclude that acknowledging and distinguishing breathiness and whisperiness as two different phonatory settings is necessary within phonetics, phonology, typology, sociolinguistics, and language variation and change.

Considering those phonetics researchers and teachers who do distinguish breathiness and whisperiness, they 1. do not necessarily use the same terminology as other researchers and teachers and 2. most often do not depict the articulatory difference accurately. We discuss each point in what follows.

Crystal [7] suggests that the terms breathiness and murmur are interchangeable but seem to be used for breathy voice as defined here. Zemlin [34 (174–176)] puts the terms breathiness, breathy voice, and noisy voice on par, and these are to be distinguished from whispering and whisper (which are used interchangeably). Stevens [29 (87)] and Borden et al. [4 (75)] use the term breathy voicing/voice for breathy voice as defined here. Ball & Rahilly [1 (36)] seem to be using breathy voice and whispery voice as defined here; however, their terminological explanation is not the same solution that we present here. Their terms breathy voice and whisper [1 (117)] are intended to refer to what are labelled breathy voice and whispery voice here. Laver [18 (190, 199, 418–419), 19] and Esling & Harris [10] are the most detailed in distinguishing breath and breathy voice, and whisper and whispery voice, in the usage we adopt here. Laver [18 (200)] further informs us that ‘Ladefoged’s term for whispery voice is murmur’, which differs from the sense of the term ‘murmur’ as outlined in [7].

The vast majority of the literature that acknowledges breathiness and whisperiness as two different settings determines that the articulatory difference between the two is a matter of the degree of glottal opening. Both are described as abducted states of the glottis, but the latter is explained as a more constricted glottis, with the airstream flowing primarily in the opening between the arytenoid cartilages, i.e. through a smaller posterior glottal gap, which is supposed to result in higher intensity friction [1 (35–26), 18 (132, 190, 198), 19 (115, 120–121, 133), 34 (176), additional references in 10 (349)]. Laver [19 (133)] specifically states that, auditorily, ‘the transition from breathiness to whisperiness is part of an auditory continuum, and the placing of the borderline between the two categories is merely an operational decision’, and, articulatorily, the main difference is based on the medial compression involved. Zemlin [34 (176)] adds that, additionally to a difference in the degree of glottal opening, the arytenoids are ‘toed in’ during the production of whisperiness, as opposed to breathiness.

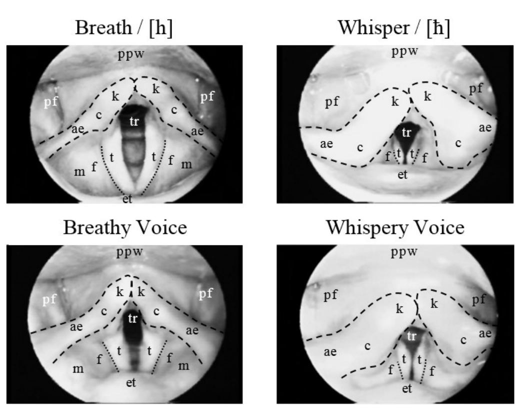
The difference between the two phonatory settings has also been discussed in terms of a degree of laxness or muscular relaxation [18 (418)], although crucially, whisperiness, unlike breathiness, can co-occur with laryngeal tension [19 (146)]. Laver [19 (121)] mentions van den Berg [3 (297)], who presents a rather rare example of a text acknowledging that the air stream may be modified further above the glottis during whisperiness, although the exact words in [3 (297)] state that the friction is generated ‘above the larynx’ rather than above the glottis itself and still *within* the larynx. Finally, Gauffin’s two-dimensional model [20 (151)] portrays the difference between ‘breathy voice’ and ‘whispering’ (as well as ‘murmur’) as based on the degree of glottal adduction but also the degree of laryngealization.

2. EPILARYNGEAL CONSTRICTION

It is clear that breathiness and whisperiness share the state of openness at the glottis but that they can be distinguished by differing states of the epilarynx. It must also be pointed out that breath, without any alterations to the cavities above the glottis, has a fully open, unconstricted laryngeal articulator, where the epilaryngeal tube is open, from glottis to abducted aryepiglottic folds. Thus, the characteristic V shape that is often attributed to the glottis in the state of breath must be treated as a function of the open laryngeal constrictor mechanism, not just of the glottis itself (which will prove crucial in explaining the Y shape associated with whisperiness).

The critical difference between breathiness and whisperiness is the addition of laryngeal constriction, narrowing the tube to generate the auditory effect of whisperiness by causing more turbulent airflow and, therefore, more noise than breath. Fig.1 shows laryngoscopic views of phonetic productions illustrating the difference. This difference is due to the formation of a narrow passage from the glottis through the aryepiglottic folds at the upper edges of the tube. Coronal magnetic resonance imaging in [15] clearly shows the high degree of epilaryngeal narrowing, the vertical extent of which is not entirely apparent in laryngoscopic views. To constrict the epilaryngeal tube, contraction of the external thyroarytenoid, which is thought to be one of the main agonists of such narrowing, recruits fibres that insert into the fovea oblonga on the lateral surface of the arytenoids, and the craniolateral extensions of these muscles even wrap around to interdigitate with the transverse interaytenoid muscle [28]. The bending of the aryepiglottic folds, repositioning the cuneiform cartilages, exerts a significant effect on the reshaping of the glottis. Whereas breathiness involves sufficient interarytenoid activity to maintain corniculate contact with relaxation of the adductors to sufficiently open the glottis to a moderate width (as seen in Fig 1), whisperiness involves activating laryngeal constriction, causing simultaneous narrowing of the glottis, which, when the vocal folds are partially abducted, results in the characteristic Y-shape evident in laryngoscopy.

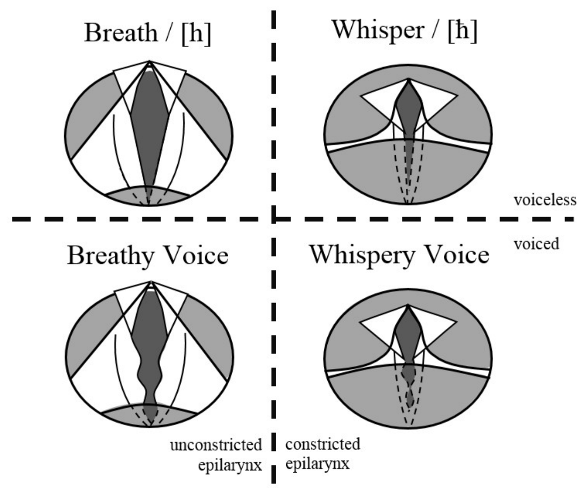
Figure 1: Laryngoscopic view of breathiness vs whisperiness. Abbreviations: ae = aryepiglottic folds, c = cuneiform tubercles, et = epiglottis (tubercle), f = ventricular (false) folds, k = corniculate tubercles, m = inner mucosa of the epilarynx, pf = piriform fossae, ppw = posterior pharyngeal wall, t = vocal (true) folds, tr = trachea. Lines: dashed line = aryepiglottic folds (margin of upper epilarynx), dotted line = medial edge of ventricular folds.



Since ‘laryngeal cartoons’ are popular in pedagogical materials [e.g. 1 (31)], we offer updated versions (Fig. 2), which are a revision enhanced from a flat larynx to a ‘whole larynx’ [24]. As is to be expected, these abstract away much detail, but, with a few evocative lines and supraglottic shading, they emphasize the epilaryngeal component of whisperiness along with concomitant glottal aperture changes. Voicing (the other parameter distinguishing these four states) is depicted with a simple, suggestive wobbling of the vocal folds.

The different varieties of whisper described by Catford [6] can be interpreted as varying functions of the tightness of the epilaryngeal tube and of the consequent vertical compression of this tube due to larynx height adjustments.

Figure 2: Breathiness vs. whisperiness: revised laryngeal cartoons; light gray shading indicates the laryngeal constrictor.



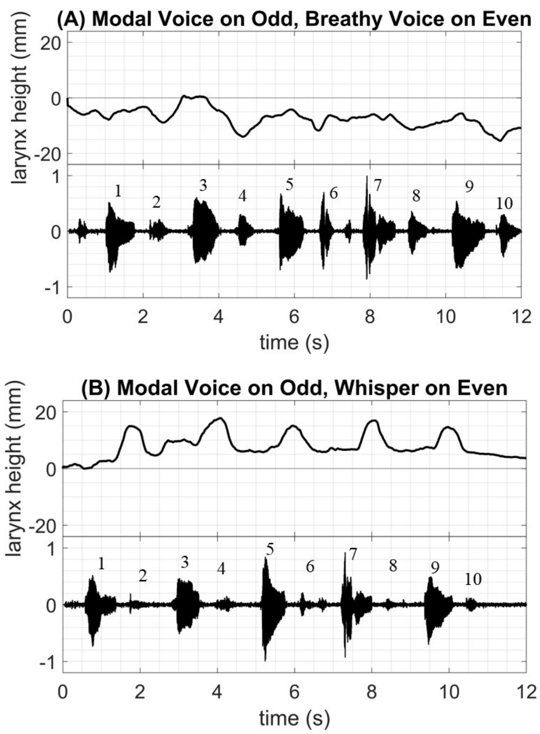
3. LARYNGEAL ULTRASOUND CASE STUDY

Given that whisper is a constricted state, we expect that it will naturally pair with larynx raising, while a state such as breathy voice will pair with larynx lowering, being unconstricted [19 (31), 23]. Here we provide some pilot data obtained using laryngeal ultrasound that suggests this pattern may hold.

The second author counted (in English) from 1 to 10. On odd numbers, modal voice was employed; on even numbers, either (A) breathy voice or (B) whisper (note: not whispery voice) were engaged. Vertical laryngeal displacement was observed by means of laryngeal ultrasound following the methods in [26], but instead using a L12-5L40S 64 (40 mm, 5–12 MHz) linear ultrasound transducer with a SonoSpeech micro ultrasound system and AAA software (Articulate Instruments).

The raw ultrasound data then underwent optical flow analysis, using the `imregdemons()` algorithm in MATLAB 2018b on frame pairs (with reduction to 25% original size) to acquire displacement fields. To obtain a vertical velocity signal, we used the `trimmean()` function, excluding the lowest and highest 25% values (and taken only over the superficial half of the video frame to prevent contamination by vocal fold flutter). Finally, the velocity signal was numerically integrated with the `cumtrapz()` function to obtain the larynx height signal. The results of this procedure appear in Fig. 3.

Figure 3: Larynx height variation during alternation between modal voice and breathy voice (A) or whisper (B) in counting sequences.



As Fig. 3A illustrates, breathy (even numbered) syllables correspond with larynx lowering events. Fig. 3B shows the opposite pattern for whisper, where the larynx raises on whispered (even numbered) syllables. Replication with phonetically naïve participants is desirable, but their behaviour will be linguistically motivated and could vary between phonetically breathy and whispery quality. The differential larynx height patterns can be directly connected with the predispositions of the laryngeal articulator, with unconstricted settings (like breathy voice) favouring a lowered larynx and constricted ones (like whisper) favouring a raised larynx.

4. CONCLUDING REMARKS

In the proposed taxonomy of abducted glottal states [10], the primary articulatory distinction is epilaryngeal. In breathy voice, airflow acts principally on the glottis, while in whispery voice, the airflow is channelled through the narrowed tube formed between the cuneiform cartilages and the base of the epiglottis, generating increased turbulence/ noise, and where the soft tissues adjacent to this channel can be set in motion. Specifically, the change in glottal shape from V-shaped in breathiness to Y-shaped in whisperiness is purely a function of the addition of (epi)laryngeal constrictor adduction. This relationship could be tested using computational biomechanical modelling, as in [25]. Computational fluid dynamic modelling could help to establish more precisely the contributions of glottal shape and epilaryngeal channeling to noise production in whisperiness. Although larynx height differences illustrated here in our case study reflect the expectations of the laryngeal articulator model – with whisperiness engaging epilaryngeal constriction and thereby occurring with larynx raising, which facilitates and is therefore the natural height predisposition for such constriction – further experimental studies are required to provide more conclusive evidence that the relationship between phonation type and larynx height indeed holds.

One parallel element to examine in the presence of voicing is f_o. The question is whether f_o is lower in breathy voice (than in modal voice), as suggested in [31 (3646)] and as indicated by larynx height movements in the case study presented here. Furthermore, it would be worthwhile to examine whether the frequency of glottal vibration alters when whispery voice is engaged. The number of articulatory variables that the laryngeal constrictor brings to the production of voice applies to whispery voice, which could lead voicing to be either higher- or lower-pitched than breathy voice, depending on horizontal and/or vertical compression factors.

A corollary issue is whether the laryngeal constriction that enables whispery voice also induces greater harshness, as a result of the effect of horizontal/vertical compression on the vibratory cycle (the propensity for vocal-ventricular fold coupling to occur). Kohler's [16] note on Lady Bracknell is a possible case of whisper even becoming growl (epilaryngeal vibration). Laver [19] points out the expanded combinatorial possibilities with other phonation types that can be identified with whispery voice. Such combinations are likely to be found in sound systems that exploit the laryngeal articulator, especially in languages that realize tonal registers.

5. REFERENCES

[1] Ball, M.J., Rahilly, J. 2003. *Phonetics: The Science of Speech*. London: Arnold.

[2] Barzeger, R.S., Wogalter, M.S. 1998. Intended carefulness for voiced warning signal words. *Proc. Human Factors and Ergonomic Society Annual Meeting*, 15.

[3] Berg, J. van den. 1968. Mechanism of the larynx and the laryngeal vibrations. In Malmberg, B. (ed), *Manual of Phonetics*. Amsterdam: North-Holland, 278–308.

[4] Borden, G.J., Harris, K.S., Raphael, L.J. 2003. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Philadelphia: Lippincott Williams & Wilkins.

[5] Borsel, J.V., Janssens, J., Bodt, M.D. 2009. Breathiness as a feminine voice characteristic: a perceptual approach. *Journal of Voice* 23, 3, 291–294.

[6] Catford, J.C. 1964. Phonation types: The classification of some laryngeal components of speech production. In: Abercrombie, D., Fry, D.B., MacCarthy, P.A.D., Scott, N.C., Trim, J.L.M. (eds), *In Honour of Daniel Jones*. London: Longmans, Green & Co. Ltd., 26–37.

[7] Crystal, D. 2003. *A Dictionary of Linguistics & Phonetics*. Oxford: Blackwell.

[8] Edmondson, J.A., Padayodi, C.M., Hassan, Z.M., Esling, J.H. 2007. The laryngeal articulator: Source and resonator. *Proc. 16th ICPhS* Saarbrücken, Germany, 2065–2068.

[9] Esling, J.H. 2005. There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics* 50, 13–44.

[10] 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: Erlbaum, 347–383.

[11] Esling, J.H., Moisik, S.R., Benner, A., Crevier-Buchman, L. 2019. *Voice quality: The laryngeal articulator model*. Cambridge: Cambridge University Press.

[12] Farley, S.D., Hughes, S.M., LaFayette, J.N. 2013. People will know we are in love: evidence of differences between vocal samples directed toward lovers and friends. *Journal of Nonverbal Behaviour* 37, 123–138.

[13] Gordon, M., Ladefoged, P. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29, 383–406.

[14] Henton, C.G., Bladon, R.A.W. 1985. Breathiness in normal female speech: Inefficiency versus desirability. *Language & Communication* 5, 3, 221–227.

[15] Honda, K., Kitamura, T., Takemoto, H., Adachi, S., Mokhtari, P., Takano, S., Nota, Y., Hirata, H., Fujimoto, I., Shimada, Y., Masaki, S., Fujita, S., Dang, J. 2010. Visualisation of hypopharyngeal cavities and vocal-tract acoustic modelling. *Computer Methods in Biomechanics and Biomedical Engineering* 13, 4, 443–453.

[16] Kohler, K.J. 2011. Communicative functions integrate segments in prosodies and prosodies in segments. *Phonetica* 68, 1–2, 26–56 DOI:https://doi.org/10.1159/000328433.

[17] Ladefoged, P., Maddieson, I. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

[18] Laver, J. 1994. *Principles of Phonetics*. Cambridge: Cambridge University Press.

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

[20] Lindblom, B. 2009. Laryngeal mechanisms in speech: The contributions of Jan Gauffin. *Logopedics Phoniatrics Vocology* 34, 4, 149–156.

[21] Liu, X., Xu, Y. 2011. What makes a female voice attractive? *Proc. 17th ICPhS* Hong Kong, 1274–1277.

[22] Moisik, S.R. 2013. *The epilarynx in speech*. Doctoral dissertation, University of Victoria.

[23] Moisik, S.R., Czaykowska-Higgins, E., Esling, J.H. 2019. Phonological potentials and the lower vocal tract. *Journal of the International Phonetic Association*, FirstView.

[24] Moisik, S.R., Esling, J.H. 2011. The 'whole larynx' approach to laryngeal features. *Proc. 17th ICPhS* Hong Kong, 1406–1409.

[25] Moisik, S.R., Gick, B. 2017. The quantal larynx: The stable regions of laryngeal biomechanics and implications for speech production. *Journal of Speech Language and Hearing Research* 60, 3, 540–560 DOI:https://doi.org/10.1044/2016_JSLHR-S-16-0019.

[26] Moisik, S.R., Lin, H., Esling, J.H. 2014. A study of laryngeal gestures in Mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association* 44, 1, 21–58.

[27] Poyatos, F. 1991. Paralinguistic qualifiers: our many voices. *Language & Communication* 11, 3, 181–195.

[28] Reidenbach, M.M. 1998. The muscular tissue of the vestibular folds of the larynx. *European Archives of Oto-Rhino-Laryngology* 255, 7, 365–367.

[29] Stevens, K.N. 1998. *Acoustic Phonetics*. Cambridge, MA: The MIT Press.

[30] Teshigawara, M. 2003. *Voices in Japanese animation: a phonetic study of vocal stereotypes of heroes and villains in Japanese culture*. Doctoral dissertation, University of Victoria.

[31] Watson, B.C. 1998. Fundamental frequency during phonetically governed devoicing in normal young and aged speakers. *The Journal of the Acoustical Society of America* 103, 6, 3642–3647 DOI:https://doi.org/10.1121/1.423068.

[32] Winter, B., Grawunder, S. 2012. The phonetic profile of Korean formal and informal speech registers. *Journal of Phonetics* 40, 808–815.

[33] Zhao, Y., Lin, W. 2016. Study of the formant and duration in Chinese whispered vowel speech. *Applied Acoustics* 114, 240–243.

[34] Zemlin, W.R. 1998. *Speech and Hearing Science: Anatomy and Physiology*. Boston: Allyn and Bacon.