INTERSPEECH 2017
August 20–24, 2017, Stockholm, Sweden

Laryngeal Articulation during Trumpet Performance: An Exploratory Study

Luis M.T. Jesus[1], *Bruno Rocha*[2], *Andreia Hall*[3]

[1]School of Health Sciences (ESSUA) and Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro, Portugal
[2]Department of Communication and Art (DeCA), University of Aveiro, Portugal
[3]Department of Mathematics (DMat) and Centre of Research and Development in Mathematics and Applications (CIDMA), University of Aveiro, Portugal

lmtj@ua.pt, brunorocha@ua.pt, andreia.hall@ua.pt

Abstract

Music teacher's reports suggest that the respiratory function and laryngeal control in wind instruments, stimulate muscular tension of the involved anatomical structure. However, the physiology and acoustics of the larynx during trumpet playing has seldom been studied. Therefore, the current paper describes the laryngeal articulation during trumpet performance with biomedical signals and auditory perception. The activation of laryngeal musculature of six professional trumpeters when playing a standard musical passage was analysed using audio, eletroglottography (EGG), oxygen saturation and heart rate signals. Two University trumpet teachers listened to the audio recordings, to evaluate the participants' laryngeal effort (answers on a 100 mm Visual-Analogue-Scale (VAS): 0 "no perceived effort"; 100 "extreme effort"). Correlations between parameters extracted from the EGG data and the perception of the audio stimuli by the teachers were explored. Two hundred and fifty laryngeal articulations, where raising of the larynx and muscular effort were observed, were annotated and analysed. No correlation between the EGG data and the auditory evaluation was observed. However, both teachers perceived the laryngeal effort (VAS mean scores = 61±14). Our findings show that EGG and auditory perception data can provide new insights into laryngeal articulation and breathing control that are key to low muscular tension.

Index Terms: Trumpet Performance, Larynx, Acoustics, Electroglottography, oxygen saturation, heart rate

1. Introduction

Laryngeal articulation is central in the control of airflow during the playing of high-pressure wind instruments [1] but it has not been thoroughly investigated [2], [3].

When a person exhales forcefully with a blocked mouth and nose, changes in the intrathoracic and intra-abdominal pressure take place that dramatically affect venous return, cardiac output, blood pressure, and heart rate. This forced expiratory effort creates an internal pressure in our body that can result in a Valsalva manoeuvre. The Valsalva manoeuvre can be described in four phases [4]. During phase I, heart rate and oxygen saturation decrease because of increased aortic pressure. In phase II the opposite occurs, the heart rate and oxygen saturation increase and the aortic pressure decreases. In stage III the subject begins to breathe normally and the aortic pressure decreases rapidly, because the external compression on the aorta is removed, and the heart rate increases rapidly. Finally, in stage IV, an increase in aortic pressure occurs which in turn decreases the heart rate.

The Valsalva manoeuvre has been observed in wind instrument players [5], especially those that need to exert greater pressure during air exhalation, namely the trumpet and the oboe performers [6]. High expiratory pressure is required in trumpet playing, especially in the upper register. This phenomenon can also occur in beginner trumpet, whose lips do not vibrate when attempting to play notes in the higher register for which they are not yet ready, thus creating a resistance which can, potentially, stimulate the Valsalva manoeuvre.

Laryngeal articulations play a central in trumpet playing and determines airflow characteristics, acting as a control mechanism for muscular tension and air expenditure. However, the behaviour of the larynx during trumpet playing has not been studied with biomedical data. Electroglography (EGG) was therefore used in this study as a non-invasive and simple means to monitor rising and lowering of the larynx [7]. This technique entails placing a pair of electrodes externally on the neck, in the alignment of the thyroid lamina. A high frequency current flows between the electrodes, and the electrical impedance of the contact area of the vocal folds is measured. The EGG waveform allows for the detection of low frequency signals that can be related to the slow movement of the larynx, as can be observed in swallowing [8], [9]. The increase in impedance during laryngeal movement can be explained by the movement of tracheal air to a higher position [10].

Therefore, the current paper describes the laryngeal articulation based on biomedical signals and auditory perception. The activation of laryngeal was analysed using audio, EGG, oxygen saturation and heart rate signals. Two University trumpet teachers listened to the audio recordings, to evaluate the participants' laryngeal.

2. Method

Audio, EGG and pulse oximeter data of all the subjects during performance, were recorded during a first stage of this study. Subsequently, the resulting audio recordings were presented randomly to expert listeners so that they could evaluate laryngeal effort. Finally, a statistical analysis was performed to calculate the correlation of the perception tests results with the EGG data.

2.1 Participants

Six trumpeters (five male and one female) were recruited, with a mean age of 22.3 ± 2.4 years and an average of trumpet playing experience of 11.0 ± 3.2 years.

http://dx.doi.org/10.21437/Interspeech.2017-315

Two expert male listeners aged 39 and 45 years, were involved in the auditory perception experiments. They were orchestral trumpeters and had more than thirty years of experience as performers, having obtained several individual national and international awards. Both had more than fifteen years of teaching experience at various schools and universities. They were also frequently invited to teach courses, give lectures and conduct master classes both nationally and internationally.

2.2 Musical passage

This study is based on the interpretation of a short excerpt: The beginning of Mahler's Fifth Symphony with the Funeral March, performed by one instrument only, the trumpet. This excerpt was chosen because it entails specific aspects of trumpet performance, namely, dynamics, articulation, tuning, register and resistance. Another factor taken into account in the choice of this excerpt was that all the participants were familiar with it, because of its wide use in orchestral work and music education.

2.3 Biomedical signal data collection

Participants were standing at a distance of about 30 cm, in front of an MKH20-P48 omnidirectional condenser microphone (Sennheiser, Germany) connected to a Scarlett 6i6 audio interface (Focusrite, UK) using a Gold Edition XLR Microphone cable (Mogami, USA). An EGG signal was also collected with an EG2-PCX2 electrophotograph (Glottal Enterprises, USA) connected to the second channel of the audio interface, with two 35mm diameter EL-2 electrodes (Glottal Enterprises, USA). The electrodes were placed externally in the alignment of the thyroid lamina and their positioning adjusted so as to obtain the EGG signal with the highest possible peak-to-peak amplitude, when observing the waveform in real time with *Soundcard Scope 1.46*.

The recordings were made with *Adobe Audition 3.0*, at a sampling rate of 48000 Hz, with 16 bits per sample, using the Focusrite USB 2.0 ASIO Driver Audio Driver 1.8. The data was recorded in stereo format .wav (Windows PCM) without compression.

Oxygen saturation and heart rate were also collected with a pulse oximeter Pulsox-300i pulse oximeter (Konica Minolta, Japan) at a sampling rate of 1 Hz. The pulse oximeter was switched on for 1 minute prior to the beginning of the musical performance, in order to establish a baseline of oxygen saturation and heart rate values. The *DS-5 Version 2.00 (090608) 06/2008* program (Konica Minolta Sensing, Japan; Software by Stowood) was used to extract the pulse oximeter data to an ASCII file.

The subjects produced six repetitions of the excerpt with a 30 second pause between each repetition. At the end of the repetitions of the musical piece, the subjects held the vowel [a] for about 2 s, to establish a reference in the audio and EGG signals with regard to amplitude and time.

The recordings took place in an ABS-AUD.45.1 cabin, produced by Absorsor, Portugal, with sound reduction of 45dB, located at University of Aveiro's Speech, Language and Hearing Laboratory (SLHlab).

After storage of the raw recordings, *Adobe Audition 3.0* was used to segment the files, and *Praat version 6.0.05* and *Matlab 8.5.0.197613 (R2015a)* were used for annotation and analysis of audio and EGG signals. The audio, EGG, oxygen saturation and heart rate signals were later synchronised for analysis with *Matlab*.

2.4 Annotation and analysis of the audio and EGG signals

The beginning (LA1) and end (LA2) of the laryngeal articulations (see Figure 1) were annotated with the *Praat version 6.0.05* and stored together with the audio and EGG signals in a binary format.

Figure 1 shows the audio and EGG signals, and the annotation of the laryngeal articulations (LA1 and LA2). The different parameters that were analysed (LA1dur, LA2dur, LA1LA2dur, LA1peak2rms and LA2peak2rms) are also highlighted.

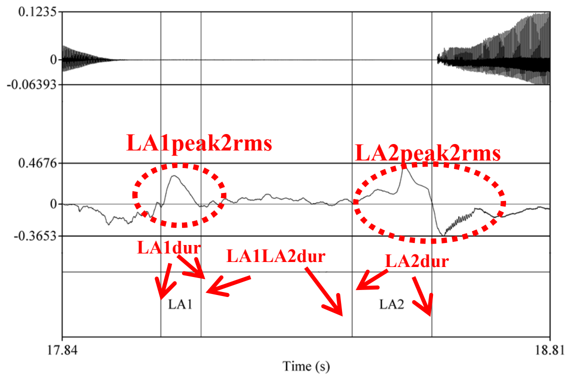

Figure 1: *Audio signal (top), EGG signal (middle) and annotation (bottom) during the first repetition of male participant DO.*

The criteria used to annotate the beginning and the end of the LA1 and LA2 intervals were: Beginning of LA1 –instant at which the EGG signal intercepts the x-axis (amplitude = 0) after a negative deflection resulting from the adduction of the vocal folds; end of LA1 – instant at which the EGG signal intercepts the x-axis again; beginning of LA2 – instant at which the amplitude of the EGG signal initiates a noticeable increase after an interval in which its mean amplitude is very close to zero, resulting from the adduction of the vocal folds; end of LA2 – instant at which the EGG signal intersects the x-axis again before a negative deflection.

The annotations were later exported in ASCII format and read by specific scripts in the *Matlab 8.5.0.197613 (R2015a)* environment where the following parameters were extracted: Duration of laryngeal articulation LA1 (LA1dur); Duration of the laryngeal articulation LA2 (LA2dur); Duration of the interval between the end of the laryngeal articulation LA1 and the beginning of the laryngeal articulation LA2 (LA1LA2dur); Peak-magnitude-to-RMS ratio (crest factor) of the EGG signal, calculated for the intervals LA1 (LA1peak2rms) and LA2 (LA2peak2rms) with *peak2rms* function of *Signal Processing Toolbox Version 7.0 (R2015a)* (MathWorks, Natick). The resulting value was used to analyse how extreme the peak of the waveform is in the range used for its calculation.

The parameter values were finally exported to an *Excel 2013* and *IBM SPSS Statistics version 22* spreadsheets for statistical analysis.

2.5 Auditory perception experiments

The individual repetitions of the six participants (36 items) were segmented with the *Adobe Audition 3.0* and stored in mono files (48000 Hz, 16 bit, uncompressed .wav Windows PCM). The resulting files were then randomised using an online tool (True random integer sequence generator based on atmospheric noise data https://www.random.org/sequences/; RANDOM.ORG; Randomness and Integrity Services Ltd; 2 Waterloo Road, Unit 20; Dublin 4; Ireland) and concatenated on a single mono file (48000 Hz, 16 bit, uncompressed, Windows PCM wav) for later listening.

The listeners were seated comfortably at a table in a quiet room, used closed dynamic transducer HD 380 pro (Sennheiser, Germany) headphones, connected to the internal sound card of a portable computer, running *Windows Media Player 12.0.7601.19148*.

A VAS scale, with a straight 100 mm line, was used for the listeners of the perception test to mark with a cross the degree of effort of the subjects, the left end of the scale representing (minimum) “with no effort” and the right end (maximum) marked with the words “extreme effort”.

A statistical analysis was carried out to determine if there was a relationship between the values of the VAS scale and the LA1dur, LA2dur, La1LA2dur, LA1peak2rms and LA2peak2rms variables. *IBM SPSS Statistics version 22* was used for this purpose. The Pearson correlation coefficient was used to calculate the correlations and a significance level of 5% was considered.

3. Results

3.1 Characterisation of laryngeal articulations

In all the subjects several movements of the larynx were registered; we thus attributed the name of laryngeal articulations (LA) to the phenomenon we observed.

Two laryngeal articulations corresponding to an elevation of the larynx (increased impedance corresponding to an increase in EGG signal amplitude) during inspiration (LA1) and expiration (LA2) are highlighted in Figure 2 with ellipses. These two articulations were observed only in the intervals of silence, in which the subject inspires a great amount of air to continue performing. The inspiration can be heard in the audio signal and its average amplitude is close to zero.

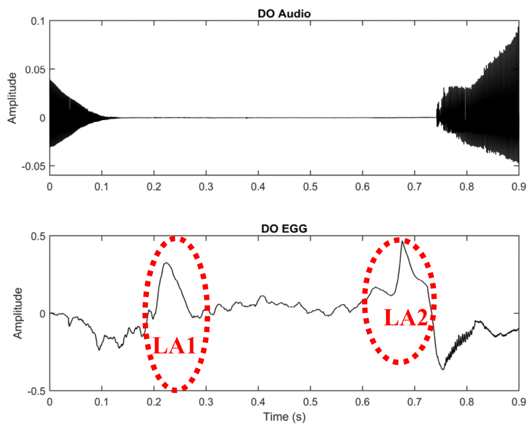

Figure 2: *Audio (above) and EGG (below) signals during the first repetition by participant DO.*

The two laryngeal articulations (LA3 and LA4) shown in Figure 3, corresponding to an elevation of the larynx during the production of 3 short notes were also observed. The participants had to move the tongue to the most anterior region of the vocal tract, which probably resulted in a raised tongue and increased air volume in the oropharynx, blocking the air in the vocal tract. This can happen if the participant only articulates the consonant [t] and not a consonant-vowel syllable (CV), thus contracting the abdominal region.

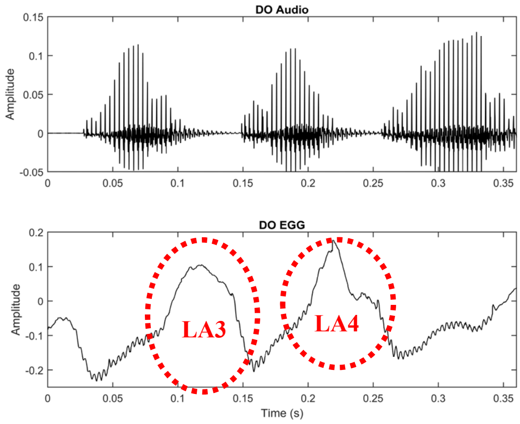

Figure 3: *Audio (above) and EGG (below) signals during the production of 3 short notes by participant DO.*

3.2 Pulse oximetry

Table 1 shows the mean (MEAN), standard error of mean (STD), minimum (MIN) and maximum (MAX) values of oxygen saturation (SpO2) and heart rate (HR), calculated for the interval between the beginning of the first repetition and the end of the last one. The SpO2 values are given in percent (%) and the HR values in beats per minute (bmp).

Table 1: *Pulse oximetry results (one line per participant).*

SpO2 (%)				PR (bpm)			
MIN	MAX	MEAN	STD	MIN	MAX	MEAN	STD
96	101	99	1	82	126	101	9
90	100	97	3	72	116	94	8
95	99	97	1	67	155	88	9
89	100	97	2	98	147	135	11
93	99	97	1	87	146	109	9
96	98	97	1	66	128	104	12

Analysis SpO2 and HR shown in Table 1 revealed that these remained stable during all the subjects' repetitions, disproving the hypothesis that participants used Valsalva manoeuvre during performance. As discussed previously, SpO2 and PR levels have divergent values in the different phases of the Valsalva manoeuvre.

3.3 Parameters extracted from the EGG signal

Table 2 shows the mean (MEAN), standard error of mean (STD), minimum (MIN) and maximum (MAX) values of parameters LA1dur, LA2dur, LA1LA2dur, LA1peak2rms and LA2peak2rms, and number of annotated laryngeal articulations (N). The 250 annotated laryngeal articulations occurred in a short time in all subjects (mean values of LA1dur = 144 ± 53 ms and LA2dur = 160 ± 64 ms), much smaller than the 10 s required for a Valsalva manoeuvre to occur [4]. The value of LA1LA2dur was 201 ± 134 ms, the mean Peak-magnitude-to-RMS ratio (crest factor) of the EGG signal was 1.856 ± 0.292, calculated for the interval LA1 (LA1peak2rms) and of 1.961 ± 0.366 for the interval LA2 (LA2peak2rms).

Table 3: *Parameters used to characterise the laryngeal articulations.*

Participant	LA1dur (ms)				LA2dur (ms)				LA1LA2dur (ms)			STD
	MIN	MAX	MEAN	STD	MIN	MAX	MEAN	STD	MIN	MAX	MEAN	
DM	29	269	130	51	54	248	143	54	47	854	357	195
DO	59	419	138	55	88	265	159	36	6	476	178	101
FM	59	251	151	46	68	308	177	54	43	382	156	80
IR	56	200	132	33	39	332	130	70	11	305	114	75
JR	61	172	113	33	75	361	242	82	116	543	291	91
LC	99	319	192	57	43	266	143	42	86	467	188	76
All	29	419	144	53	39	361	160	64	6	854	201	134

Participant	LA1peak2rms				LA2peak2rms				N
	MIN	MAX	MEAN	STD	MIN	MAX	MEAN	STD	
DM	1.342	2.218	1.771	0.201	1.308	2.582	1.939	0.355	35
DO	1.438	2.471	1.833	0.191	1.434	2.462	1.995	0.210	61
FM	1.331	2.392	1.809	0.262	1.500	3.370	2.023	0.394	37
IR	1.471	3.372	1.784	0.295	1.381	3.900	1.963	0.446	48
JR	1.521	3.258	2.059	0.371	1.351	2.428	1.872	0.268	27
LC	1.443	3.287	1.955	0.346	1.406	3.271	1.931	0.453	42
All	1.331	3.372	1.856	0.292	1.308	3.900	1.961	0.366	250

Table 4: *Auditory perception results.*

	Listener LG				Listener JS			
Participant	MIN (%)	MAX (%)	MEAN (%)	STD (%)	MIN (%)	MAX (%)	MEAN (%)	STD (%)
DM	50	87	62	12	39	89	67	16
DO	32	87	64	17	59	95	81	14
FM	33	80	53	16	21	69	48	16
IR	33	92	62	19	46	100	72	16
JR	28	80	57	17	49	84	69	12
LC	32	63	48	11	31	62	45	10
All	35	82	58	15	41	83	64	14

3.4 Auditory perception experiments

Table 4 presents the mean value (MEAN), standard error of the mean (STD), minimum value (MIN) and maximum value (MAX) of the perception scores for each subject. Table 4 shows high dispersion for all subjects. As a central tendency, both listeners perceived moderate laryngeal effort values. Further analysis using a scatter plot, considering the six replicas of each participant, suggested that the correlation between the two sets of scores is very low. Indeed, the value of Pearson's correlation was 0.079. Due to the dependence between some of these scores the significance of this value could not be accurately assessed. Nevertheless, the p-value was not significant ($p=0.647$). Correlations between each measurement (LA1dur, LA2dur, LA1LA2dur, LA1peak2rms and LA2peak2rms) and the scores assigned by the listeners (total of 250 pairs of observations by each listener) were also calculated. Again, the conditions of applicability of the Pearson correlation test were not completely met, since the observations had dependencies (by groups regarding each player). Therefore, the p-values were not accurate and must be interpreted with care. The data used in this analysis does not allow us to obtain well-founded conclusions about the correlations (the complete set has several dependences that affect the accuracy of the tests and the reduced independent set is too small). Nevertheless the analysis suggests that there are no significant correlations between the measured acoustic variables and scores given by the listeners.

4. Conclusions

In the present study, an average of 42 laryngeal articulations were recorded for each of six players during performance. Analysis of EGG signals' duration, the mean values of the oxygen saturation (97%) and heart rate (105 bpm) for all subjects led us to conclude that there were no Valsalva manoeuvres during performance. Regarding the perception tests, the mean score given by the two raters on a scale of 0 to 100 for all subjects was 61%. Considering the extremes of the VAS scale used, it can be said that the laryngeal effort of the subjects is moderate. Pearson's correlation coefficient between the two listeners was 0.079, indicating small or no correlation. Since the two listeners did not correlate, the comparison of the acoustic measurements and the listeners' scores was performed separately for each listener. However, p-values of the tests led to conclude that overall the acoustic variables and the scores were not correlated.

5. Acknowledgements

This research was partially funded by national funds through the FCT - Foundation for Science and Technology, in the context of the projects UID/CEC/00127/2013 and UID/MAT/04106/2013. The authors would like to thank Maria Teresa Roberto for precious help with writing this manuscript.

6. References

[1] R. R. . Chapman, “A Strategy for Proper Utilization of the Glottis and Implications of Improper Use in Trumpet Performance,” Ph.D. Thesis, University of Miami, 2014.

[2] R. Tubiana and P. C. Amadio, *Medical Problems of the Instrumentalist Musician*. New York: Taylor & Francis, 2000.

[3] M. Schumacher, C. Schmoor, A. Plog, R. Schwarzwald, C. Taschner, M. Echternach, B. Richter, and C. Spahn, “Motor functions in trumpet playing—a real-time MRI analysis,” *Neuroradiology*, vol. 55, no. 9, pp. 1171–1181, Sep. 2013.

[4] R. E. Klabund, *Cardiovascular Physiology Concepts*, 2nd ed. Ohio: Wolters Kluwer, 2012.

[5] L. Schuman, “The Cichowicz School: Comments from Former Studentes of Vincent Cichowicz,” *ITG J.*, vol. 23, pp. 23–25.

[6] K. Steenstrup, *Teaching Brass*, 2nd ed. Aarhus: Royal Academy of Music, 2007.

[7] M. Farooq, J. M. Fontana, and E. Sazonov, “A novel approach for food intake detection using electroglottography,” *Physiol. Meas.*, vol. 35, no. 5, pp. 739–751, May 2014.

[8] J. L. Schultz, A. L. Perlman, and D. J. VanDaele, “Laryngeal movement, oropharyngeal pressure, and submental muscle contraction during swallowing.,” *Arch. Phys. Med. Rehabil.*, vol. 75, no. 2, pp. 183–8, Feb. 1994.

[9] M. S. Klahn and A. L. Perlman, “Temporal and durational patterns associating respiration and swallowing.,” *Dysphagia*, vol. 14, no. 3, pp. 131–8, 1999.

[10] R. Sorin, M. D. McClean, F. Ezerzer, and B. Meissner-Fishbein, “Electroglottographic evaluation of the swallow.,” *Arch. Phys. Med. Rehabil.*, vol. 68, no. 4, pp. 232–5, Apr. 1987.