Original Article

Design of Wearable Electrolarynx with Automatic Control

M. Madhushankara[1], Somashekara Bhat[2], Keerthana Prasad[1]

[1]Manipal School of Information Sciences, Manipal Academy of Higher Education,
[2]Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India

Abstract

Objective: The current work aims to design and develop an automatically controlled wearable electrolarynx, a voice substitution device for laryngeal carcinoma survivals. **Methods:** The physical activity of mouth opening is sensed, amplified, and made to act as an enable signal to trigger the wearable electrolarynx. The resulting speech is recorded and compared for its voice reaction durations with that of manual electrolarynx and normal speaking methods. Perception evaluations of 5 subjects from 10 speech-language therapists are obtained. **Results:** The wearable electrolarynx turn-on in 13 µs once the mouth movement for speech is sensed. The voice initiation time and termination durations are 215.68 m and 231.41 ms, respectively. Results indicate that there is no significant difference ($P < 0.05$) between the voice reaction durations of wearable electrolarynx and normal speaking methods. The subjective evaluation results show that there is a significant improvement ($P < 0.05$) in intelligibility and noise reduction when compared to a commercially available electrolarynx with an average intra-class correlation coefficient of 0.68 from analysis of variance two factors without replication. **Conclusions:** The assessment of the wearable and automatically controlled electrolarynx provides hands-free speech and easy control over the device.

Keywords: *Automatic control, carcinoma, electrolarynx, wearable*

Submitted: 17-Aug-2021 **Revised:** 02-Nov-2021 **Accepted:** 20-Dec-2021 **Published:** 10-Nov-2022

Introduction

A larynx is a flexible organ between the pharynx and the trachea, an essential component in the speech production system. A laryngectomy is a surgical procedure to remove cancer cells in the larynx. The complete removal of the larynx leads to the disconnection of the trachea from the pharynx, preventing any natural speech after the surgery.[1] To produce the speech, the laryngectomees need alaryngeal methods such as electrolarynx. This device can also help patients with oral intubation for speech production.[2,3]

Speech rehabilitation is a vital part to establish verbal communications in speech-deprived patients. There are mainly three different speech restoration methods.[4] Speech production by the sudden release of injected air into the oral cavity is known as esophageal speech.[5] The tracheoesophageal prosthesis is another method to produce voice with the help of a one-way valve.[6] When the patient wishes to speak, he or she manually occludes the valve so that air is directed towards the esophagus. The prosthesis allows the air to vibrate the pharynx muscles. The pulsation of the upper vocal cord is then modulated to a speech by the proper movement of articulators. The electrolarynx is a handheld electromechanical device to vibrate the pharynx muscles externally. Speech is produced by the appropriate movement of articulators.[7]

In the present situation, this speech aid suffers from manual control. This condition could be awkward for the user where he/she cannot start speaking instantly when needed. Automatic turn-on/off of the electrolarynx is necessary for the user to speak spontaneously. Many researchers have attempted to solve this problem either with the intra-oral electrolarynx or with neck muscle activities to control the electrolarynx.

An intra-oral electrolarynx in the form of a synthetic tooth inside the oral cavity is one of the methods to avoid carrying it with a hand.[8,9] Movement of the finger is used to turn-on/turn-off and also for pitch variation.[8,10-12] The lateral movement of the thumb is related to

---

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

How to cite this article: Madhushankara M, Bhat S, Prasad K. Design of wearable electrolarynx with automatic control. J Med Sign Sens 2022;12:317-25.

Address for correspondence:
Dr. M. Madhushankara,
Manipal School of Information Sciences, Manipal Academy of Higher Education,
Manipal - 576 104, Karnataka, India.
E-mail: madhushankar.m@manipal.edu

Access this article online
Website: www.jmssjournal.net
DOI: 10.4103/jmss.jmss_147_21
Quick Response Code:

the fundamental frequency variation, and the up-down movement is associated with device onset and offset. These finger-controlled switches need the user's promptness when initiating verbal communications.

The activity of the anterior part of the neck and chin muscles as the secondary source of speech information is used to regulate the electrolarynx.[13-17] Using sufficiently amplified electromyographic sensor signals captured from positions of the neck and face surface, the electronic circuitry controls the activation and deactivation of the electrolarynx. However, carrying these electrodes along with the electrolarynx makes a patient's life more troublesome. The device's turn-on and turn-off capability of the neck and face muscles of patients who have undergone laryngeal surgery is carried out to ascertain its acceptance.[18,19] A wireless connection between the electromyography signals and transducers was established to further increase the ease of use.[20,21] However, the reliability of the device for activation, deactivation, and control over the pitch needs to be examined and there was no further work reported by the authors.

A real-time approach to control the voicing is performed using video cameras and artificial neural technology.[22-25] The artificial neural network (ANN) system including a camera facing the mouth for controlling the device has resulted in better turn-on control and less turn-off control than the existing electrolarynx. However, it comes with an extra cost of a video camera along with the carrier headset. The complex ANN is trained with a conjugate gradient back-propagation algorithm. These systems of automatic control of the device suffer either from complex algorithms and the challenge of capturing the myographic activity or come with an extra budget in the form of video cameras.

In the present work, a sensor interface to automatically control the electrolarynx when the mouth opens for speech is presented. The noninvasive and nonobtrusive method controls the device when required. The movement of the mandible during speech is detected using a flex sensor. The variation in the resistance of the sensor is converted into the variation in voltage. The device triggers when the voltage level is more than a threshold value; which defines the mouth opening for speech.

Methodology

Figure 1 depicts the methodology used in the present work. The process of detecting the intention to speak begins when the electrolarynx is powered on. On the successful detection of the mouth opening for speech, an excitation pattern is generated. The excitation source is provided to the neck surface through a power amplifier and a transducer. A power-down mode is activated when there is no mouth movement to save the battery power. The electrolarynx resumes normal mode of operation if there are any mouth movement activities without manual intervention.

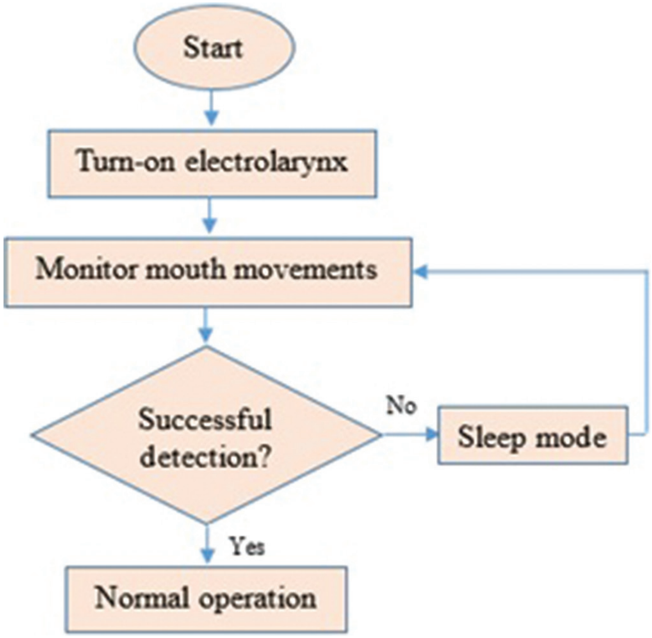


Figure 1: Flowchart

We have followed a simple approach to control the electrolarynx by using mouth movements for speech. The electrolarynx is either turn-on or turn-off depending on the open and closing actions of the mouth. An anatomical structure called mandible bone in the human face is responsible for the oral cavity movements as shown in Figure 2. The mean value of the angle of the mandible is found to be 123.5° with a range of 106°–141°.[27] During the mouth opening, the mandible moves downward and translates forward. During mouth closing, this pattern is inverted.[28] A strain is produced on the mandibular bone during both speech and mastication. For speech, the amount of the strain is lesser in magnitude than that of chewing action.[29] The strain variation is converted into variation in voltage by the appropriate circuit configuration and thereby used to control the electrolarynx.

Several stretchable, soft electronic sensors which can be placed on the body are readily available and can be categorized into ultrasonic, piezoelectric, and resistive sensors. Based on the current application and nature of the mandible, a resistive flex sensor is employed in detecting the movement of the mouth.

A low-cost, commercially available Flex Sensor from SpectraSymbol is employed to capture the movement of the mandible thereby the movement of the lower jawbone, which helps to identify the mouth opening for speech.[30,31] It is coated with resistive carbon components on a thin elastic substrate. The sensor is located just underneath the ear, where the mandible motion can be detected due to the changes in skin curvature.[26] The resistance is directly proportional to its degree of bending. Figure 3 shows the placement of a resistive flex sensor with an active length of 5.54 cm attached just below the ear.

The block diagram of the prototype electrolarynx along with the automatic control is shown in Figure 4. Figure 5 shows the sensor interface for automatic control. The sensor interface consists of the flex sensor as a part of Whetstone's bridge network along with the instrumentation amplifier to amplify the signal voltage.[32] The flex sensor is connected at R_3 of the Wheatstone bridge circuit. The two ends of the bridge network are connected to an instrumentation amplifier configured using operational amplifiers U_1, U_2, and U_3. The resistance of the flex sensor varies when bent from its horizontal position. It has a resistance of 31.1k Ω resistance in the horizontal position. When bent at an angle of 90°, the resistance is 111.2k Ω. The sensitivity of the flex sensor is 8.75k Ω/10° change in angle.

From the Wheatstone bridge network, the equations for V_{01} and V_{02} can be obtained as,

$$V_{01} = \frac{R_2}{R_1 + R_2} \times V_B \tag{1}$$

$$V_{02} = \frac{R_4}{R_3 + R_4} \times V_B \tag{2}$$

Since the operational amplifier U_1 and U_2 are in voltage follower mode the voltage V_{03} and V_{04} are the same as V_{01} and V_{02} respectively. The amplifier U_3 acts as the difference amplifier and amplifies the difference of voltage, V_{03}, and V_{04}. The voltage V_{06} is given by Eq. 4.[32]

$$V_{06} = \frac{R_8}{R_8 + R_6} \times V_{04} \tag{3}$$

If, $V_{04} = 0$, then output at V_{OUT} due to V_{03} is,

$$V_{OUT} = -\frac{R_7}{R_5} \times V_{03} \tag{4}$$

If, $V_{03} = 0$, then output at V_{OUT} due to V_{04} is,

$$V_{OUT} = \frac{R_8}{R_6 + R_8} \times V_{04} \times \frac{R_5 + R_7}{R_5} \tag{5}$$

With $R_6 = R_5$ and $R_8 = R_7$, combining Eq. (4) and (5),

$$V_{OUT} = \frac{R_7}{R_5} \times (V_{04} - V_{03}) \tag{6}$$

Here R_1, R_2, and R_4 are 33k Ω, R_3 is the flex sensor, R_5, R_6 are 10k Ω, and R_7 and R_8 are 100k Ω each. The waveform generator consists of a microcontroller, digital to analog converter, and power amplifier. The microcontroller is programmed to acquire signals from the sensor in real-time and convert them to digital values by the inbuilt analog to digital converter. Initially, the voltage is captured with the mouth closed and stored as a preset value. A control signal is made HIGH when the mouth opening activity is captured indicated by the values which are lower than the preset value. The stored excitation signal in the microcontroller memory is converted to analog signals by an R-2R digital to analog converter. The amplified signals through a class AB power amplifier vibrate the transducer which is in contact with the neck surface. A circuit board with surface mounted devices is constructed as shown in Figure 6.

Figure 7 shows the sizing of the transducer used in the present electrolarynx. A coil pulsating at the fundamental frequency is made to strike a coupler which is pressed against the throat. The magnet, coil, and piston are positioned inside a nylon shaft of 2.5 cm diameter by making appropriate holes. A circular-shaped neodymium magnet of 2.5 cm in diameter and 0.3 cm in thickness acts as the base of the housing. Another cylindrical magnet of size 1.4 cm in diameter and 1.0 cm in height is placed

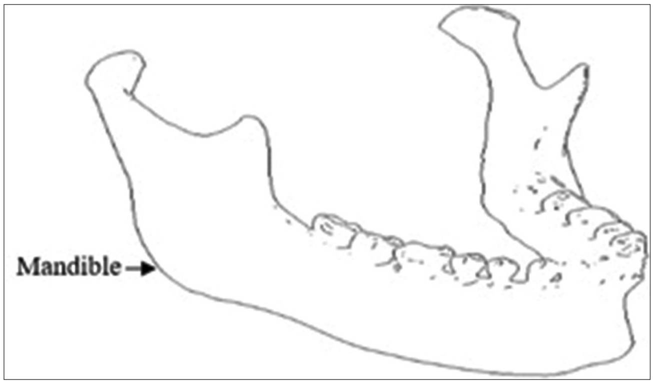


Figure 2: Placement of the flex sensor (Mandible Image Courtesy[26])

Figure 4: Block diagram of the prototype

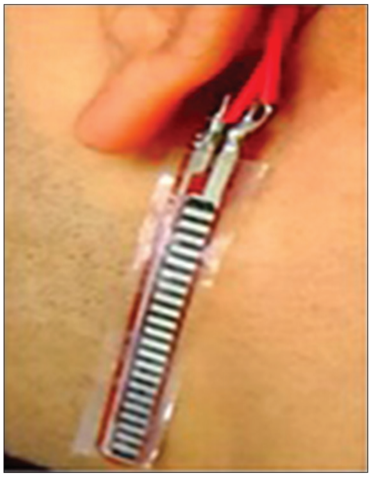


Figure 3: Placement of the Flex sensor

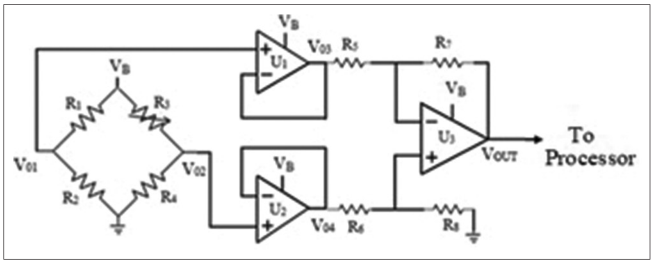


Figure 5: Sensor Interface

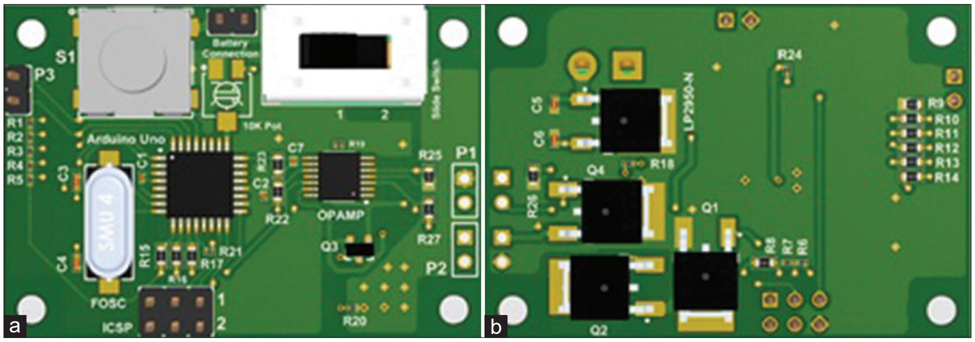


Figure 6: Circuit board with components mounted on (a) Top layer (b) Bottom layer

at the center of the circular magnet. An air core piston wound with a coil to form resistance of 8 Ω is used as the electromagnet. A circular-shaped flexible card of 2.5 cm diameter is used as the coupler

Construction of wearable electrolarynx

A wearable electrolarynx prototype is prepared to make the hands-free operation of the device. A customized thin and flexible steel plate with 0.1 cm thickness is used as the platform. With magnets being readily attached to the steel, the base of the transducer fits firmly onto it. The printed circuit board and the battery are glued on the remaining portion of the metal plate. One such combination with the board, battery, and transducer is represented in Figure 8. This combination allows the prototype to be a wearable electrolarynx as shown in Figure 9.

Results

A high capacity less weight Li-ion battery is used to power the circuit. Two batteries of 3.7 V, 860 m Ah ratings are connected in series. A positive voltage regulator (LP2905) is used to convert the battery voltage into 5 V and to supply the processor. Initially, the sensor is calibrated and obtained minimum and maximum voltages and found to be 0.10 V and 4.25 V respectively. The conversion time required is thirteen clock cycles. With an external clock frequency of 16M Hz and a prescaler value of 16, the duration for the conversion is 13 µs

Voltage variation during mouth movements

When the mouth opens, the mandible moves forward and the flex sensor straightens up. This causes the resistance value to decrease. From the circuit configuration, it can be observed that the V_{OUT} value increases when the flex sensor resistance decreases. The variation of voltage at V_{OUT} to time during different phases of mouth is plotted as shown in Figure 10, using the Arduino Integrated Development Environment. Table 1 represents the voltage variation and corresponding digital equivalents during mouth closing, mastication, and opening for speech. An activation signal is generated for the interval corresponding to the mouth open for speech as shown in Figure 11.

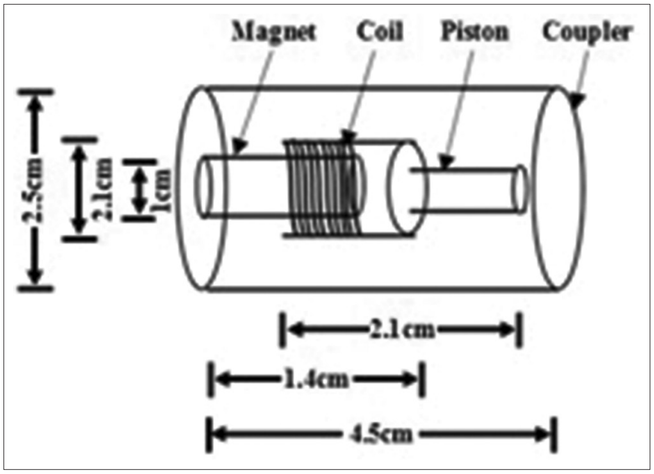


Figure 7: Internal parts of the transducer

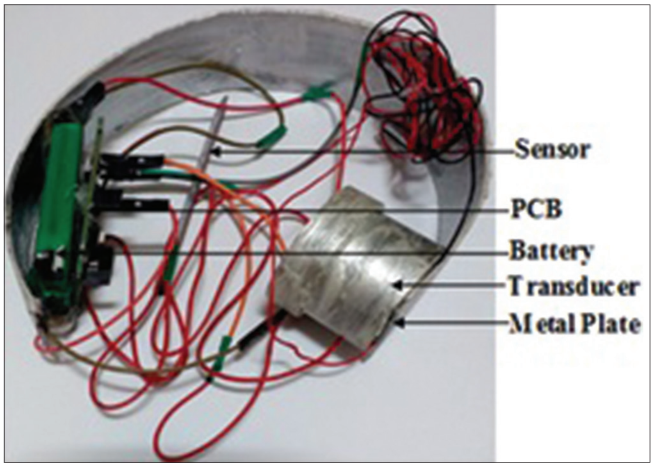


Figure 8: Prototype electrolarynx

It can be inferred from Figure 10 that the output voltage variation causes the circuit configuration to detect mouth movements during speech. The corresponding enable signal is used to automatically turn-on or turn-off the electrolarynx.

Voice initiation and termination duration

The ability of the automatically controlled electrolarynx for its time taken to respond to the mouth opening and closing

Figure 9: Wearable prototype in use

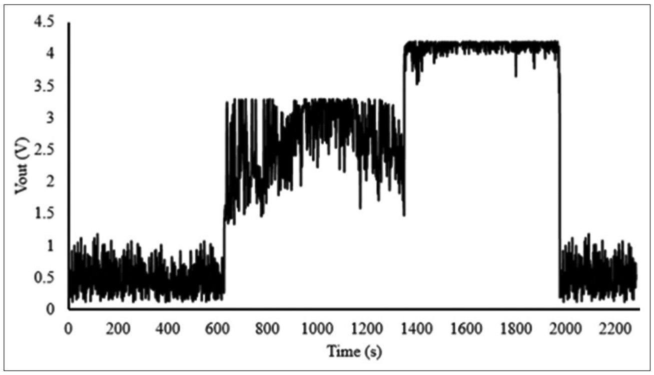


Figure 10: Output voltage as a function of time during different phases of mouth

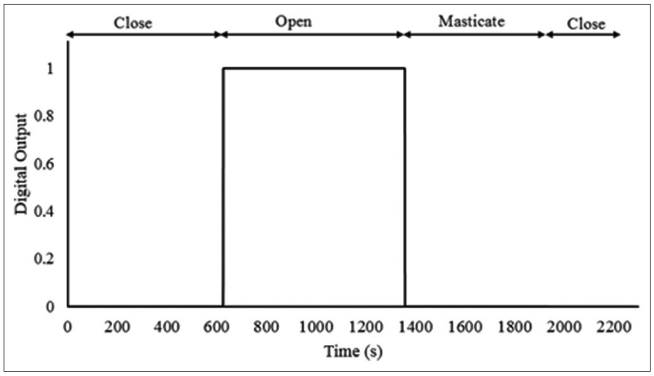


Figure 11: Transducer activation signal as per mouth movement

actions is measured when in the system in use as shown in Figure 9. These durations are also compared with that of manually controlled electrolarynx (Servox Digital Speech Aid) and normal voicing methods. A visual cue is provided to the subject to begin the speech and the time taken by the electrolarynx to activate is noted down. Similarly, time taken to deactivate is also measured after the cue for stopping the speech. A standard convention is followed to record the audio outputs and calculation of reaction durations.[33-35]

Table 1: Voltage range for corresponding mouth positions

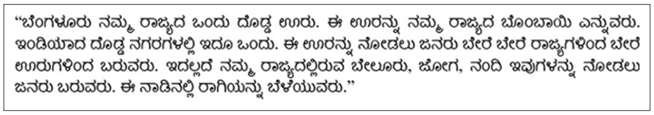
Mouth position	Voltage level (V)
Close	0.12 to 1.18
Open	1.34 to 3.29
Mastication	3.40 to 4.19

A total of 20 measurements were taken from the healthy individual of age 37 years with a body mass index of 22.9 kg/m^2 for measuring the duration required for voice initiation (DVI) and duration required for voice termination (DVT) using normal, Servox, and prototype methods. The voice samples are collected at different times of 4 days' period. Each experiment is started with a relaxation duration of 15 s, followed by a duration of 2 s for "be ready." It is followed by a reaction cue such as sustained vowel/i/,/a/or/u/. The vocalization has to be stopped on the stop signal. All the voice samples are recorded using Multi-Speech Model 4500 sampled at 44.1k Hz and stored in a Laptop. Table 2 lists these durations of vocal reaction time in milliseconds.

Figures 12 and 13 show the box plot of the DVI and DVT for all trials along with minimum, maximum, median, and average values. The average value of DVI and DVT for the prototype is 215.68 m s and 231.41 m s respectively. The values are very close to that of normal values of the subject (211.69 m s and 228.71 m s) as indicated in Table 2. The average values of Servox are 325.72 m s and 319.18 m s which are higher than that of both automatically controlled and normal reaction times. A *t*-test with $P < 0.05$ is performed on the duration of voice initiation and termination. It can be inferred from the DVI and DVT test that, the ability of the automatically controlled electrolarynx is comparable to that of the normal method of speaking. The test also reveals that Servox has significantly larger reaction durations than that of both automatic electrolarynx and normal methods.

Subjective evaluation

The subjective evaluation is carried out by the speech therapists listening to the individuals reading a phonetically balanced Kannada paragraph. The speech is recorded from five volunteers aged 34 ± 5 years who are native speakers of Kannada (4 males, 1 female) as they withhold their breath and used Servox and prototype one after the other as the voice substitution method. A Multi-Speech Model 4500 is used as the recording instrument and sampled at 44.1k Hz to store in a Laptop. Following is the passage with 41 words used for perceptual evaluations.[36]



“ಬೆಂಗಳೂರು ನಮ್ಮ ರಾಜ್ಯದ ಒಂದು ದೊಡ್ಡ ಊರು. ಈ ಊರನ್ನು ನಮ್ಮ ರಾಜ್ಯದ ಬೊಂಬಾಯಿ ಎನ್ನುವರು. ಇಂಡಿಯಾದ ದೊಡ್ಡ ನಗರಗಳಲ್ಲಿ ಇದೂ ಒಂದು. ಈ ಊರನ್ನು ನೋಡಲು ಜನರು ಬೇರೆ ಬೇರೆ ರಾಜ್ಯಗಳಿಂದ ಬೇರೆ ಊರುಗಳಿಂದ ಬರುವರು. ಇದಲ್ಲದೆ ನಮ್ಮ ರಾಜ್ಯದಲ್ಲಿರುವ ಬೇಲೂರು, ಜೋಗ, ನಂದಿ ಇವುಗಳನ್ನು ನೋಡಲು ಜನರು ಬರುವರು. ಈ ನಾಡಿನಲ್ಲಿ ರಾಗಿಯನ್ನು ಬೆಳೆಯುವರು.”

At first, the mean fundamental frequency of the individuals was measured by recording their normal conversation.

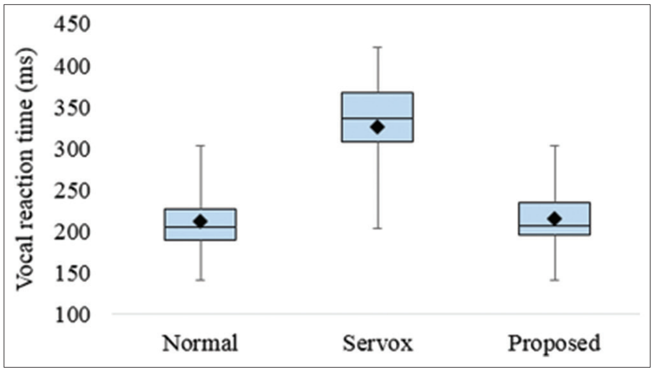


Figure 12: Box plot of duration required for voice initiation

Table 2: Vocal reaction time under different methods in milliseconds

Serial number	Normal		Servox		Prototype	
	DVI	DVT	DVI	DVT	DVI	DVT
1	207.58	241.66	385.68	295.95	282.23	218.36
2	222.12	248.66	347.93	294.69	157.87	297.71
3	195.02	212.99	422.54	199.59	176.50	184.50
4	165.81	327.20	418.47	202.30	303.90	182.94
5	282.34	236.92	204.44	283.52	206.57	156.89
6	156.87	192.71	323.21	246.71	233.99	126.41
7	176.49	184.48	329.75	307.98	237.18	236.92
8	303.90	218.36	359.77	180.56	205.53	192.71
9	176.72	297.71	331.52	191.93	206.57	184.48
10	195.32	184.50	217.82	292.14	223.08	238.32
11	224.76	182.94	397.06	410.88	195.02	328.27
12	141.43	212.99	267.80	341.17	175.72	291.22
13	216.20	327.20	360.96	330.89	196.23	241.66
14	203.92	281.45	406.01	540.57	223.67	248.66
15	264.58	156.89	341.09	315.00	141.43	212.99
16	247.97	126.41	322.13	379.50	215.10	327.20
17	237.18	281.45	206.54	499.86	203.86	281.45
18	204.53	211.05	339.55	353.45	274.58	211.05
19	206.57	230.23	207.92	400.72	247.97	230.23
20	204.62	218.34	324.22	316.24	206.62	236.24
Average	211.69	228.71	325.72	319.18	216.68	231.41

DVI - Duration required for voice initiation;
DVT - Duration required for voice termination

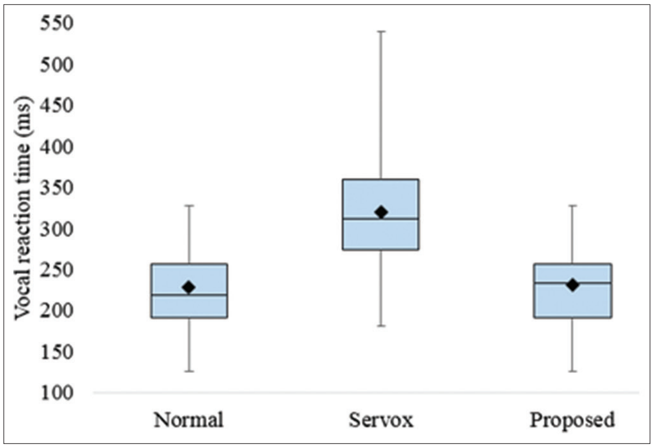


Figure 13: Box plot of duration required for voice termination

Followed by this, audios were obtained by using Servox with three different frequencies (80 Hz, 250 Hz, and average fundamental frequency of the user). As a next step, the audio recording of subjects when using the proposed electrolarynx with a frequency equal to their average fundamental frequency is recorded

More than 800 words were presented to the evaluators as a group of four paragraphs per subject (5 individuals × 41 words × 4 excitation sources). Since the average fundamental frequency of the female was coinciding with 250 Hz, only three samples were obtained from her. Thus, a total of 779 words from 19 passages and an additional 20% of random samples were used. To find out the fundamental frequency, the speech analysis software, PRAAT, is used.[37] The average fundamental frequency for the subjects A, B, C, D, and E are found to be 150 Hz, 160 Hz, 120 Hz, 130 Hz, and 250 Hz, respectively.

Rating procedure

Each listener was seated in a sound-treated room and was instructed to evaluate the recorded speech quality between 0 and 3 according to the overall impression (I), unintended additive noise (N), together called as IN scale.[38] Score 0 is the best, and 3 is the worst in the category. The following attributes of the scaling are employed.[39]

- I–Impression of the speech quality
- It is the general impression produced by speech including intelligibility, noise, and fluency.
- N– Noise.
- Caused by the audibility of all sorts of uncontrolled sounds, such as inherent electrolarynx noise.

Listeners

A group of 10 normal hearing postgraduate students of the School of Allied Health Sciences, Manipal, who knew the perceptual scoring system is requested to be as Speech-Language Therapists (SLT) evaluators. The obtained data are first analyzed for intra-evaluator agreement using exact agreement statistics.[40]

The intra-evaluator agreement for I and N is 0.96 and 0.88 respectively with an average of 0.92. Intra-class correlation coefficients (ICC) from analysis of variance two factor without replication were used to analyze inter-evaluator reliability. The ICC was 0.69, and 0.67 respectively for impression and noise, with an average of 0.68. Table 3 represents the average and standard deviation of SLTs.

The *post hoc t*-test is conducted to analyze whether the usage of the proposed device has significantly improved the overall grading. Table 4 shows the result of the *t*-test between the grading of Servox and the prototype.

Table 3: Perceptual evaluation by speech-language therapists

Subjects	Source	Impression (I)		Noise (N)	
		μ	σ	μ	σ
A	S_{80}	2.30	0.32	2.35	0.45
	S_{250}	2.75	0.32	2.60	0.62
	S_{150}	2.75	0.32	3.0	0
	M_{150}	2.65	0.37	2.95	0.15
B	S_{80}	2.50	0.48	2.65	0.45
	S_{250}	2.85	0.22	2.95	0.15
	S_{160}	2.70	0.57	2.95	0.15
	M_{160}	2.75	0.57	2.90	0.30
C	S_{80}	0.50	0.52	1.15	0.74
	S_{250}	0.65	0.60	0.90	0.44
	S_{120}	1.15	0.64	1.15	0.63
	M_{120}	0.50	0.48	0.85	0.74
D	S_{80}	1.45	0.54	1.90	0.54
	S_{250}	1.55	0.40	2.15	0.39
	S_{130}	1.20	0.57	1.85	0.55
	M_{130}	2.0	0.56	2.45	0.47
E	S_{80}	1.75	0.61	2.45	0.52
	S_{250}	1.60	0.56	2.10	0.66
	M_{250}	1.10	0.36	1.45	0.27

Table 4: *P* value of *t*-test evaluations

Subjects	Fundamental frequency (Hz)		
	80	250	Average F0
A	0.007	0.169	0.015
B	0.002	0.002	0.271
C	0.013	0.002	0.006
D	0	0.04	0.001
E	0.002	-	0

The *P* value from Table 4, indicates that for subjects A and B, the values are not significant for the frequency 250 Hz and Avg F0 respectively. It is also evident from Table 3 that their grading is lesser than the remaining subjects. Hence, it indicates that the SLT evaluators did not find the difference between Servox and prototype when the perception is too poor. In all other cases, the values are significantly improved by using the prototype. Therefore, it can be inferred from the subjective evaluation that, the proposed electrolarynx is better than Servox with respect to the overall intelligibility of the speech.

Discussion

The purpose of this study was to design a wearable and automatically controlled electrolarynx. The design of automatic control is validated by comparing the voice reaction times that of normal and Servox methods. The current findings suggest that the proposed electrolarynx is better than the manually controlled electrolarynx, Servox.

The higher durations for voice termination than that of voice initiation in all cases are corroborated with works of.[13,22,41-43] In Goldstein *et al.*,[13] EMG activated electrolarynx is compared with that of normal and manually controlled devices. It was observed that there is no significant improvement from the latter two methods. There was no comparison made among the different activation methods except EMG activation in.[41-43] In Wu *et al.*,[22] the reaction time comparison is made between normal, manual, video controlled along with the ANN controlled electrolarynges. It observes that ANN controlled electrolarynx is significantly better than the video-controlled one. However, there is no significant improvement from the manually controlled electrolarynx.

In a survey conducted on the age of laryngeal cancer incidence between 30 and 80 years of age, the median is found to be 60 years.[44] The present work is carried out with a 37-year-old normal subject. The age factor and the skill to improve the voice reaction time are not available in the literature. Despite this, the present work could be extended to have the measurements with a laryngectomy.

The reliability study of flex sensor in goniometer reveals the lack of accuracy when measuring small joints with less degree of freedom.[45] The repeated experiments of about 10,000 times on strain sensors had reliable sensor signals under strain below 100%.[46] The wearable cloth with flex sensors at elbow, wrist, and knee positions is a good candidate to measure the body posture by the variation in their angle.[47] Therefore, when combined with an electronic sensor interface, such as the one proposed in this article would give a good correlation for the degree of bending.

In this study, automatic control of the prototype is tested with a single subject. An electrolarynx is a single person with reusable equipment. Most often, the patient requires training by a professional speech therapist for its appropriate use. At this point, the threshold levels could be set by the therapist to suit the specific patient. The current work will be extended to provide a mechanism for fine-tuning the threshold level. This feature could be helpful for the patients who directly buy electrolarynx from the store as per the prescription of a doctor.

The movements of the mandibular are also observed in humans during chewing, hopping, brisk walking, and running. To avoid triggering these undesired situations, a manual switch to turn the device off completely is provided in this work. In future work, a three-axis sensor will be used to accommodate both head and mouth movements during speech. To make the system power-efficient, switching voltage regulators will be used as an alternative to the linear voltage regulators used in the present system.

Conclusions

In this study, we proposed, developed, and assessed an electrolarynx with an automatic turn-on and turn-off mechanism providing easy control over the device. The novel approach of detecting the movement of the lower jaw during mouth opening is sufficient to control the device. To evaluate the effectiveness of the proposed system, the voice reaction duration is compared with a manually controlled electrolarynx as well as using a person's normal voice. The proposed device was found to be significantly improved from the manual electrolarynx. The automatic switching coupled with a wearable style electrolarynx has the potential to make the patient's life better.

Financial support and sponsorship

Self financed.

Conflicts of interest

There are no conflicts of interest.

References

1. Tucker HM, Wanamaker J, Trott M, Hicks D. Complications of laryngeal framework surgery (phonosurgery). Laryngoscope 1993;103:525-8.
2. Imura M, Nakashima Y, Kawaguchi R, Terashima M, Yamada S. Verbal communication of an orally intubated patient in the intensive care unit: A case report. Prog Rehabil Med 2019;4:20190016.
3. Cox SR, Raphael LJ, Doyle PC. Production of vowels by electrolaryngeal speakers using clear speech. Folia Phoniatr Logop 2019;72:250-6.
4. Pawar PV, Sayed SI, Kazi R, Jagade MV. Current status and future prospects in prosthetic voice rehabilitation following laryngectomy. J Cancer Res Ther 2008;4:186.
5. Christensen JM, Fletcher SG, McCutcheon MJ. Esophageal speaker articulation of/s, z: A dynamic palatometric assessment. J Commun Disord 1992;25:65-76.
6. Eadie TL, Doyle PC, Hansen K, Beaudin PG. Influence of speaker gender on listener judgments of tracheoesophageal speech. J Voice 2008;22:43-57.
7. Lal GJ, Gopalakrishnan E, Govind D. Glottal activity detection from the speech signal using multifractal analysis. Circuits Syst Signal Process 2020;39:2118-50.
8. Takahashi H, Nakao M, Kikuchi Y, Kaga K. Alaryngeal speech aid using an intra-oral electrolarynx and a miniature fingertip switch. Auris Nasus Larynx 2005;32:157-62.
9. Takahashi H, Nakao M, Kikuchi Y, Kaga K. Intra-oral pressure-based voicing control of electrolaryngeal speech with intra-oral vibrator. J Voice 2008;22:420-9.
10. Kikuchi Y, Kasuya H. Development and Evaluation of pitch Adjustable Electrolarynx. In: Speech Prosody 2004. Nara, Japan: International Conference; 2004.
11. Wan C, Wang E, Wu L, Wang S, Wan M. Design and evaluation of an electrolarynx with tonal control function for Mandarin. Folia Phoniatr 2012;64:290-6.
12. Wang L, Feng Y, Yang Z, Niu H. Development and evaluation of wheel-controlled pitch-adjustable electrolarynx. Med Biol Eng Comput 2017;55:1463-72.
13. Goldstein EA, Heaton JT, Kobler JB, Stanley GB, Hillman RE. Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity. IEEE Trans Biomed Eng 2004;51:325-32.
14. Yu S, Lee T, Ng ML. Surface electromyographic activity of extrinsic laryngeal muscles in cantonese tone production. J Signal Process Syst 2016;82:287-94.
15. Yu S, Lee T, Ng ML. Surface electromyographic activity of non-laryngeal neck muscles in Cantonese tone production. In: Chinese Spoken Language Processing (ISCSLP), 2014 9th International Symposium On. 2014: IEEE; 2014. p. 304-7.
16. Rameau A. Pilot study for a novel and personalized voice restoration device for patients with laryngectomy. Head Neck 2020;42:839-45.
17. Prakash A, Sharma S, Sharma N. A compact-sized surface EMG sensor for myoelectric hand prosthesis. Biomed Eng Lett 2019;9:467-79.
18. Stepp CE, Heaton JT, Rolland RG, Hillman RE. Neck and face surface electromyography for prosthetic voice control after total laryngectomy. IEEE Trans Neural Syst Rehabil Eng 2009;17:146-55.
19. Eskes M, van Alphen MJ, Smeele LE, Brandsma D, Balm AJ, van der Heijden F. Predicting 3D lip movement using facial sEMG: A first step towards estimating functional and aesthetic outcome of oral cancer surgery. Med Biol Eng Comput 2017;55:573-83.
20. Heaton JT, Robertson M, Griffin C. Development of a wireless electromyographically controlled electrolarynx voice prosthesis. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. Massachusetts, USA: IEEE; 2011. p. 5352-5.
21. Kubert HL, Stepp CE, Zeitels SM, Gooey JE, Walsh MJ, Prakash SR, *et al.* Electromyographic control of a hands-free electrolarynx using neck strap muscles. J Commun Disord 2009;42:211-25.
22. Wu L, Wan C, Wang S, Wan M. Development and evaluation of on/off control for electrolaryngeal speech via artificial neural network based on visual information of lips. J Voice 2013;27:259.e7-16.
23. Wu L, Wan C, Xiao K, Wang S, Wan M. Evaluation of a method for vowel-specific voice source control of an electrolarynx using visual information. Speech Commun 2014;57:39-49.
24. Matsunaga Y, Matsui K, Nakatoh Y, Kato YO. Development of hands-free speech enhancement system for both el-users and esophageal speech users. In: International Symposium on Distributed Computing and Artificial Intelligence. Porto, Portugal: Springer; 2017. p. 334-41.
25. Nochino T, Ohno Y, Kato T, Taniike M, Okada S. Sleep stage estimation method using a camera for home use. Biomed Eng Lett 2019;9:257-65.
26. McCarthy J, Schreiber J, Karp N, Thorne C, Grayson B. Lengthening the human mandible by gradual distraction. Plast Reconstr Surg 1992;89:1-8.
27. Gabriel AC. Some anatomical features of the mandible. J Anat 1958;92:580.
28. Ostry DJ, Munhall KG. Control of jaw orientation and position in mastication and speech. J Neurophysiol 1994;71:1528-45.
29. Daegling DJ. The human mandible and the origins of speech. J Anthropol 2012;2012:1-4.
30. Sazonov E, Schuckers S, Lopez-Meyer P, Makeyev O, Sazonova N, Melanson EL, *et al.* Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. Physiol Meas 2008;29:525.
31. Sazonov ES, Fontana JM. A sensor system for automatic detection of food intake through non-invasive monitoring of chewing. IEEE Sens J 2012;12:1340-8.
32. Franco S. Design with Operational Amplifiers and Analog Integrated Circuits. 3rd ed. New York, USA: McGraw-Hill; 2002.
33. Shipp T, Izdebski K, Morrissey P. Physiologic stages of vocal reaction time. J Speech Hear Res 1984;27:173-8.
34. Watson BC. Foreperiod duration, range, and ordering effects on

acoustic LRT in normal speakers. J Voice 1994;8:248-54.

35. Murphy M, Baumgartner JM. Voice initiation and termination time in stuttering and nonstuttering children. J Fluency Disord 1981;6:257-64.
36. Kacker S, Basavaraj V. Indian Speech, Language and Hearing Tests-The ISHA Battery-1990. Mysore: Indian Speech and Hearing Association; 1990.
37. Boersma P. Praat, a system for doing phonetics by computer. Glot Int 2002;5:341-5.
38. Moerman M, Martens JP, Crevier-Buchman L, de Haan E, Grand S, Tessier C, *et al.* The INFVo perceptual rating scale for substitution voicing: Development and reliability. Eur Arch Otorhinolaryngol 2006;263:435-9.
39. Moerman M, Martens JP, Van der Borgt M, Peleman M, Gillis M, Dejonckere P. Perceptual evaluation of substitution voices: Development and evaluation of the (I) INFVo rating scale. Eur Arch Otorhinolaryngol 2006;263:183-7.
40. Kreiman J, Gerratt BR, Kempster GB, Erman A, Berke GS. Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. J Speech Hear Res 1993;36:21-40.
41. Ooe K, Villagran CR, Fukuda T. Development of the compact control system using of neck EMG signal for welfare applications. In: 2010 International Symposium on Micro-NanoMechatronics and Human Science. Nagoya, Japan: IEEE; 2010. p. 127-32.
42. Ooe K. Development of controllable artificial larynx by neck myoelectric signal. Proc Eng 2012;47:869-72.
43. Fuchs AK, Amon C, Hagmüller M. Speech/non-speech detection for electro-larynx speech using EMG. In: Biosignals. Lisbon, Portugal; 2015. p. 138-44.
44. Nahavandipour A, Jakobsen KK, Grønhøj C, Hebbelstrup Jensen D, Kim Schmidt Karnov K, Klitmøller Agander T, *et al.* Incidence and survival of laryngeal cancer in Denmark: A nation-wide study from 1980 to 2014. Acta Oncol 2019;58:977-82.
45. Henderson J, Condell J, Connolly J, Kelly D, Curran K. Reliability and validity of clinically accessible smart glove technologies to measure joint range of motion. Sensors 2021;21:1555.
46. Kim S, Oh J, Jeong D, Park W, Bae J. Consistent and reproducible direct ink writing of eutectic gallium-indium for high-quality soft sensors. Soft Robot 2018;5:601-12.
47. Abro ZA, Yi-Fan Z, Nan-Liang C, Cheng-Yu H, Lakho RA, Halepoto H. A novel flex sensor-based flexible smart garment for monitoring body postures. J Ind Text 2019;49:262-74.