Analysis of a soft bio-Inspired active actuation model for the design of artificial vocal folds

Azadeh Shariati
Research Fellow
Department of Mechanical Engineering
University College London
Torrington Place, London WC1E 7JE, UK
Email: a.shariati@ucl.ac.uk

Helge A. Wurdemann
Professor of Robotics
Department of Mechanical Engineering
University College London
Torrington Place, London WC1E 7JE, UK
Email: h.wurdemann@ucl.ac.uk

ABSTRACT

Phonation results from the passively-induced oscillation of the vocal folds in the larynx, creating sound waves that are then articulated by the mouth and nose. Patients undergoing laryngectomy have their vocal folds removed and thus must rely on alternative sources of achieving desired vibration of artificial vocal folds. Existing solutions, such as voice prostheses and the Electrolarynx, are limited by producing sufficient voice quality, for instance. In this paper, we present a mathematical analysis of a physical model of an active vocal fold prosthesis. The inverse dynamical equation of the system will help to understand whether specific types of soft actuators can produce the required force to generate natural phonations. Hence, this is referred to as the active actuation model. We present the analysis to replicate the vowels /a/, /e/, /i/, and /u/ and voice qualities of vocal fry, modal, falsetto, breathy, pressed, and whispery. These characteristics would be required as a first step to design an artificial vocal folds system. Inverse dynamics is used to identify the required forces to change the glottis area and frequencies to achieve sufficient oscillation of artificial vocal folds. Two types of Ionic polymer-metal composite (IPMC) actuators are used to assess their ability to produce these forces and the corresponding activation voltages required. The results of our proposed analysis will enable research into the effects of natural phonation and, further, provide the foundational work for the creation of advanced larynx prostheses.

1 Introduction

Phonation in humans is achieved through airflow generated by the lungs travelling through the larynx where vocal folds, ligaments covered by layers of mucous tissue, then begin to passively oscillate. This creates the oscillating airflow patterns of sound which are further shaped by articulation as they exit the mouth [1]. A number of health conditions can negatively impact this phonation ability as vocal folds might be damaged or required to be removed. For instance, severe cases of laryngeal cancer, the second most common head and neck cancer, might result in laryngectomy, a surgical intervention to remove both the tumor and the surrounding larynx [2]. This medical treatment will leave patients without vocal folds to oscillate, and with airflow being redirected away from the mouth through an introduced opening, a stoma, in the patients' neck for breathing. The two most common methods used to replace the subsequent loss of phonation include voice prostheses and Electrolarynxes [5]. A voice prosthesis is a one-way silicone valve placed in an opening that is created between the stoma windpipe and the oesophagus [6, 7]. Patients are able to speak by closing the stoma to redirect airflow through the voice prosthesis and into the back of the mouth for articulation. These devices generally have a short life-expectancy, needing replacements every six months, although earlier failure is common [7]. On the other hand, an Electrolarynx is a device that patients must press to the outside of their throat, above the stoma, that vibrates the remaining tissue to generate an

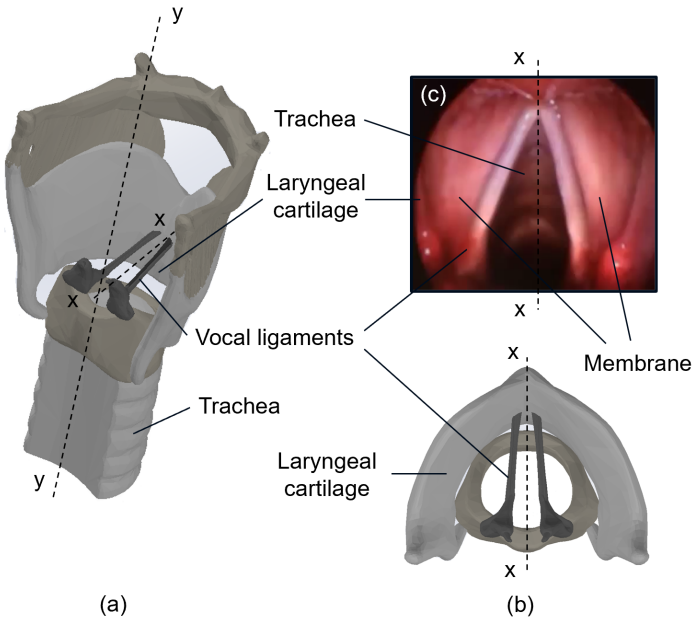

Figure 1. The layout of the natural larynx system: (a) A 3D CAD model and (b) top view of the larynx composed of laryngeal cartilage, vocal ligaments and trachea, used for our modelling approach [3]. (c) Image of intact human vocal folds [4].

oscillating airflow behind the mouth, independent of the airflow through the stoma for breathing. Most designs are only capable of producing vibrations at a singular fundamental frequency, resulting in monotonic speech [8]. Both approaches result in limited ranges of voice quality due to vibrating tissue with differing mechanical properties to the natural vocal folds and lack of fine motor control of the tracts used for airflow. These difficulties in varying voice quality and fundamental frequencies not only impair the use of stylistic effects in languages such as English, but even more significantly impair the ability to differentiate words in tonal languages, such as Mandarin Chinese, but also others such as Gujarati that rely on voice quality [9, 10].

Physical replicas of natural vocal folds have been created that imitate the mechanical properties involved in these effects [11, 12]. They have been primarily designed to account for the complexity involved in models of the airflow and the passive mechanical properties of the tissue layers. In fact, robotic models have been explored to imitate both healthy behaviour and study conditions such as vocal fold paralysis [13]. These physical models rely on airflow for passive actuation, airflow that has been redirected out of the stoma for breathing in laryngectomy patients. As these passive devices have limited potential use for implants, due to limited durability, for instance, active devices might be a feasible alternative. However, new models would be required to mathematically understand the behaviour and response of vocal folds to suggest design requirements for active implants.

In this paper, we propose a physical model of an active vocal fold prosthesis (a primarily physical model). The inverse dynamical equation of the system, which is referred to as the active actuation model, is determined to understand whether specific types of soft actuators can produce the required force to generate natural phonations. Natural displacements of intact human vocal folds which can generate specific vowels and sound qualities have been borrowed from the literature [1,2]. These displacements are used as input to the inverse dynamics, which then determine the required actuation forces. In some cases, the actuation system needs to be able to move the paralyzed vocal folds. Therefore, we have used viscoelastic coefficients of the human vocal fold in our dynamical model. We then examine whether a specific type of actuator can produce the required force and frequency within the particular operational domain. The results of our proposed analysis will enable research into providing the foundational work for the creation of advanced active larynx prostheses. In particular, the contribution of this paper are summarised below:

- A soft robotic mechanism is introduced to imitate the key physiological features of the larynx involved in producing phonation.
- Ionic polymer-metal composite (IPMC) actuators are explored as the active actuation mechanism due to low voltage requirements, small sizes and ability to generate sufficient forces and frequencies to produce phonation in the human larynx.
- A dynamic model of the mechanism including the response to a range of vocal qualities are analysed. In particular, the relationship between the required voltages and actuation forces is obtained.

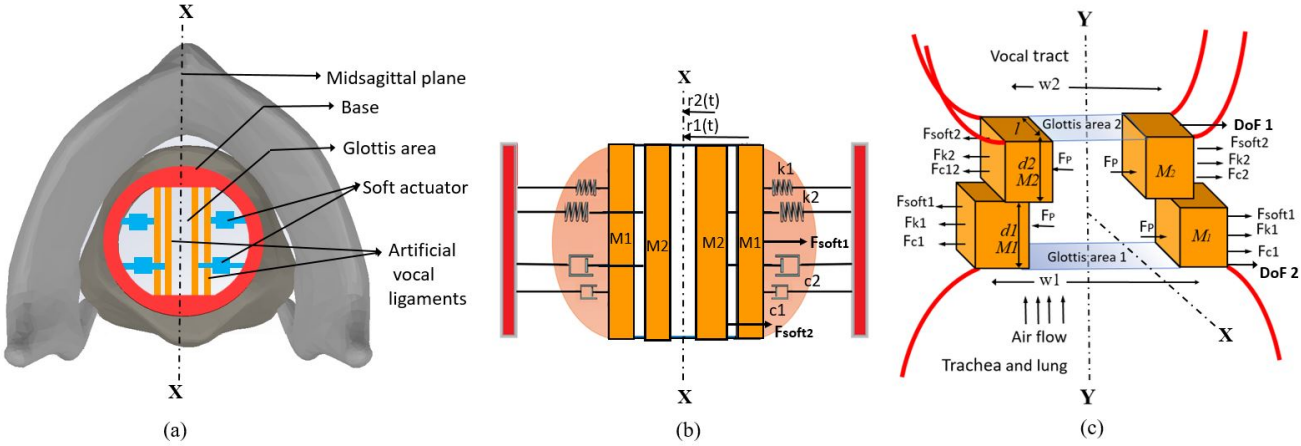

Figure 2. (a) Configuration of the proposed larynx mechanism with active soft actuators. (b) Schematic of a two-mass model of vocal folds with soft actuators (front view of the vocal fold model). (c) Forces applied to the two-mass vocal folds (top view of the vocal fold model).

Section 2 provides an overview of both the natural and proposed artificial vocal fold system, which is then modelled in Section 3. Our model is based on a lumped mass dynamical model to define the required parameters for active actuation. When considering the choice of actuators, different soft robotic techniques have been investigated. IPMC actuators are selected, due to their ability to produce required forces, displacements and frequencies, and further investigated. Section 4 then explores and analyses two types of IPMC actuators for imitating vowels and voice qualities. The results are presented in Section 5 and discussed, along with further design considerations. Finally, the conclusions are presented in Section 6.

2 System Overview

Figure 1 (a) shows the overview of an intact human larynx, located at the top of the neck. It encases the vocal folds and the soft tissue, which stretches across the larynx to produce phonation. The folds contain opposing ligaments, which form a constriction at the top of the trachea, where it joins to the lower vocal tract. When air is ejected at a sufficient velocity through the glottis, the folds vibrate and act as an oscillating valve, which interrupts the airflow into a series of pulses. These pulses of volume flow serve as the excitation source for the vocal tract in all voiced sounds, and the passive resonances of the vocal tract are excited by the glottal pulses. When the vocal folds vibrate, they move back and forth vertical along the x-axis shown in Figure 1 (c), opening and closing the glottis area, in the manner of a shear wave, travelling up the vocal fold tissue along the y-axis, also known as the mucosa wave.

To investigate voice synthesis, a wide range of models have been proposed such as the single-mass model [14]. Flanagan et al. introduced a single-mass model of the vocal folds, which moves perpendicular to the airflow imitating z-axis oscillation only [14]. Afterwards, Ishizaka and Flanagan suggested one of the most well-known vocal folds model consisting of two masses connected to the larynx base via two springs and two dampers, which are joined by a third spring. The two masses are positioned above and below each other to represent the additional motion of the mucosa wave [15]. Adachi et al. modelled the vocal folds with a single mass vibrating both perpendicular and parallel to the airflow [16]. Birkholz et al. used a two-mass model of the vocal folds, which are inclined instead of being parallel to the airflow. Their model is able to synthesise specific words in six voice qualities [17]. On the other hand, Story et al. proposed three-mass model [18] and Titze et al. a 16-masses model [19], whereas Alipour et al. introduced a numerical model through a finite element analysis [20].

The aforementioned one-mass, two-mass [15], and 16-mass [19] models aim to model human vocal folds behaviour. The aim of this paper, however, is to find a suitable actuator using inverse dynamics rather than modelling the human vocal folds to create future active implants. Hence, we incorporate an existing vocal fold model into our dynamical equations considering the possibility of incorporating the vocal fold tissue into the prosthetic mechanism. A two-mass model, similar to the one proposed by Ishizaka, has been proposed as their model includes the number of Degrees of Freedom required to generate the main features of human cord behaviour [15], i.e., their model is capable of articulating key voice qualities [17]. In addition, we employed the 2-DoF glottis area data found by Birkholz and Ishizaka as an input to our inverse dynamics.

Figure 2 (a) shows a resulting schematic of the artificial vocal folds presented and analysed in this paper. The artificial vocal folds contain a base, the red ring as illustrated in Figure 2 (a), representing the laryngeal cartilage (see Figures 1 (c)). Four bars, the orange vertical strips in Figure 2 (a), act as the artificial vocal fold tissue. The artificial vocal folds are driven by soft actuators, the blue horizontal strips in Figure 2 (a), allowing the glottis to open and close.

The actuators in this design must move artificial vocal folds or real paralysed vocal folds. Using a two-mass model, we have considered the vocal fold as a combination of m_1 and m_2 masses, located one behind the other in the vocal tract (Figure

2(c), see DoF 1 and DoF 2). The model includes two assumptions for the purpose of developing the design requirements and dynamic model. According to the first assumption, the same actuation forces are applied with respect to the midsagittal plane. Secondly, the design of the larynx mechanism is symmetrical with respect to the midsagittal plane (see the plane through the X-X and Y-Y lines in Figures 1 and 2). Therefore, the half-system can be modelled with two masses, m_1 and m_2, resulting in a model with two Degrees of Freedom (DoFs).

3 Dynamic Model of the Vocal Folds

In this section, we explore the forces resulting in vibration of the artificial vocal fold masses described in Section 2 so that the required human vowels/voice qualities, e.g., the vowels */a/*, */e/*, */i/*, and */u/* and voice qualities of *vocal fry*, *modal*, *falsetto*, *breathy*, *pressed*, and *whispery*, can be reproduced. These qualities are based on the literature [17], and will cover speaking and singing as well as controlled ways to signal paralinguistic information and phonological contrasts. To achieve this, we first derived the dynamic equations of our artificial vocal folds mechanism. Secondly, some glottis area and their respective vocal folds displacements were specified which have been previously identified in the literature as requirements to produce particular vowels and voice qualities. Finally, by using the inverse dynamics and desired displacement, the required actuation forces to produce vowels/ voice qualities are obtained. The obtained actuation force will be incorporated in Section 4 to find the corresponding soft actuator voltage, and will be examined to determine weather the voltage is within the operational domain (see Figure 3).

3.1 Dynamic Representation

Figures 2 (b) and (c) show a schematic mechanical two-mass model of the artificial vocal folds with actuators. The model includes two symmetrical halves, each containing two dampers, with damper constants of c_1 and c_2, and two springs, with spring constants of k_1 and k_2, connecting the base to two masses of m_1 and m_2. Masses m_1 and m_2 represent the masses of vocal folds in a two-DoF model. The free-body-force diagram of the mechanism is shown in Figure 2 (c). The external forces F_{soft_i} are applied to the masses by the actuators. F_{Pi} are forces acting on the vocal folds due to the air pressure. Finally, F_{ki} and F_{ci} are forces resulting in the inherent spring and damper characteristics of the membrane in the mechanism. We define the configuration variables of the mechanism as $[r_1(t), r_2(t)]^T$, where $r_1(t)$ and $r_2(t)$ are displacements of the masses (artificial vocal folds) from the symmetry line. Using $F_{ki} = k_i r_i(t)$, $F_{ci} = c_i \frac{dr_i(t)}{dt}$ and the Newton-Euler approach, i.e., $\Sigma F_i = m_i \frac{d^2 r_i(t)}{dt^2}$, the dynamic equations of the artificial vocal fold mechanism are obtained in Equation 1.

$$\frac{1}{2} m_i \frac{d^2 r_i(t)}{dt^2} + \frac{1}{2} c_i \frac{dr_i(t)}{dt} + \frac{1}{2} k_i r_i(t) = F_{p_i} + F_{soft_i}, i = 1,2 \tag{1}$$

where F_{p_i} are multiple forces resulting from the glottis pressure P_g acting on the area $A_{o_i} = l_i \times d_i$.

$$F_{p_i} = A_{o_i} \times P_g, i = 1,2 \tag{2}$$

The parameters m_i, $c_i = 2\frac{\sqrt{m_i k_i}}{q}$, k_i, q, l_i, and d_i in Equation 1 and 2 were chosen by taking the mean values of 53 intact male human vocal folds as found in [21], where $m_i = 0.018$ gr, $k_i = 160$ N/m, $q = 9.0$, $l_i = 0.0138$ m, and $d_i = 0.003$ m for $i = 1,2$. Parameter q is the quality factor [16], l_i is the length of each vocal folds, and d_i is the thickness of vocal folds. The pressure $P_g = 800$ Pa is chosen as reported in [16]. We have relied on parameter values from a male-only study as research findings conclude that males are five times more likely to be affected by laryngeal cancer [22] and, thus, are more probable users of artificial replacements.

3.2 Vocal Fold Displacements and Required Actuation Forces

Figure 3 presents a diagram of the process how to determine the soft actuator's working requirements based on our dynamic equations for reproducing specific vowels and voice qualities. We used the glottis areas estimated by experiments introduce by Ishizaka et al. from intact human larynx to synthesise four vowels including */a/*, */e/*, */i/*, and */u/* [15]. In addition, we adapted the glottis areas as calculated in [17], where Birkholz et al. adjusted the traditional two-mass model to identify data for not only three vocal registers of *vocal fry*, *modal* and *falsetto* but also three additional vocal qualities of *pressed* (also known as *harsh*) voices, *breathy* voices and *whispery* voices. In order to find the vocal folds displacement, i.e., $r_i(t)$, the glottis area is divided into the length l_i of the vocal folds as follows,

$$r_i(t) = \frac{A_i(t)}{l_i}, i = 1,2 \tag{3}$$

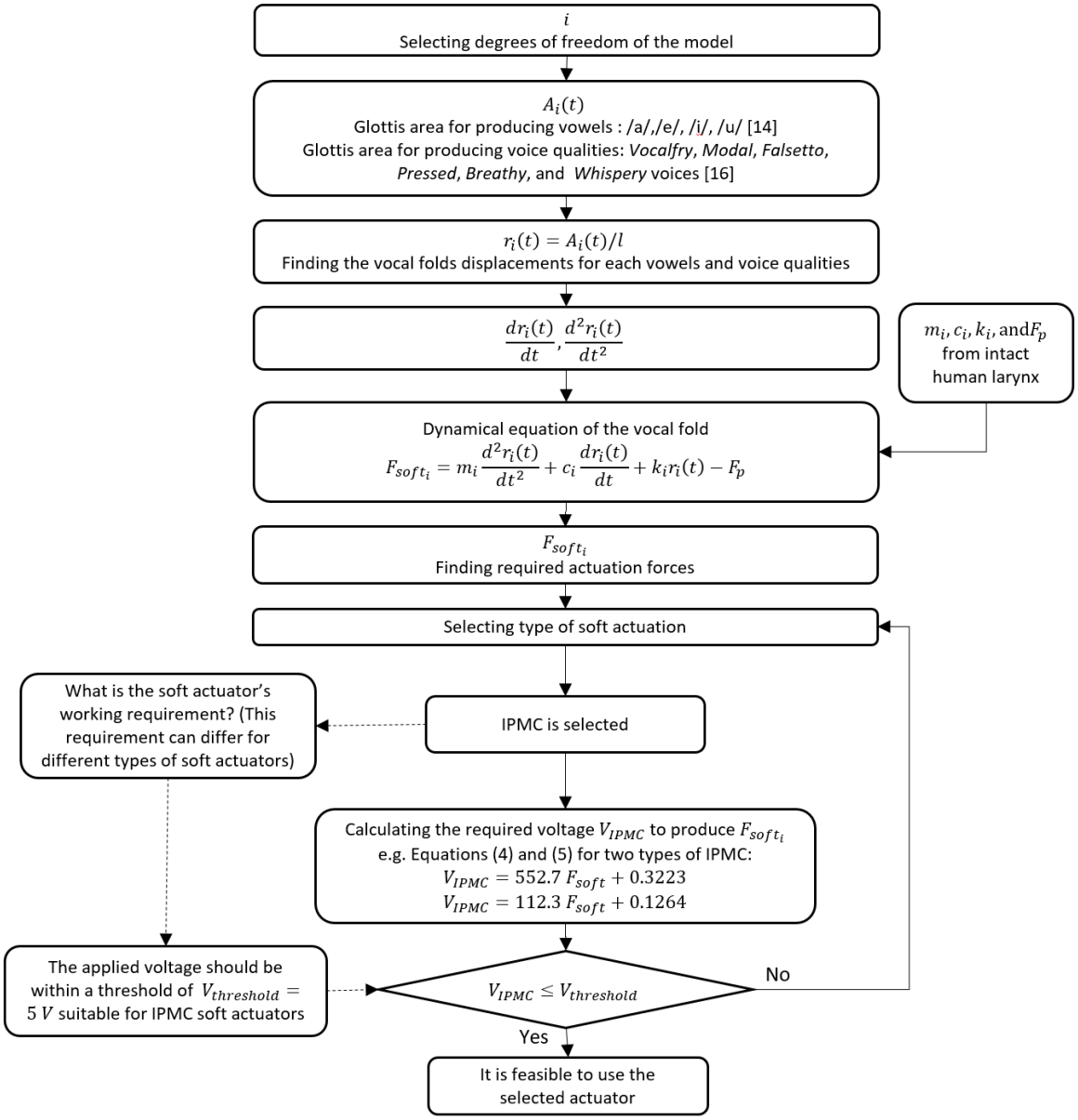

Figure 3. Flow chart of the process to obtain the soft actuator's working requirements (here: applied voltages), which are needed to drive, e.g., IPMC actuators, using our dynamic equations to produce specific vowels and voice qualities (e.g., */a/*, */e/*, */i/*, */u/*, *modal* voice, *breathy* voice, *pressed* voice, *vocal fry*, *whispery* voice, and *falsetto*).

The vocal folds displacement as suggested by Ishizaka et al. are presented in Figures 4 (a)-(d) with differences in displacement amplitude and frequencies across the vowels and voice qualities. Figures 4 (e)-(j) show the vocal folds displacements calculated from the data in Birkholz et al. [17].

The actuation forces F_{soft_1} and F_{soft_2} which are required to move the vocal fold masses to sufficient displacements, resulting in the four vowels and the six voice qualities, are calculated. The glottis area and corresponding displacements of the center of mass of the vocal folds, i.e., the results for $r_1(t)$ and $r_2(t)$ are used in Equation 1. Using first and second derivatives of $r_i(t)$ in Equation 1, F_{soft_1} and F_{soft_2} can be calculated as shown in Figure 4 (k)-(t). Following the flow chart in Figure 3, the desired voltages required to be applied to the soft actuators and the artificial vocal fold system in order to produce the vowels and voice qualities are obtained.

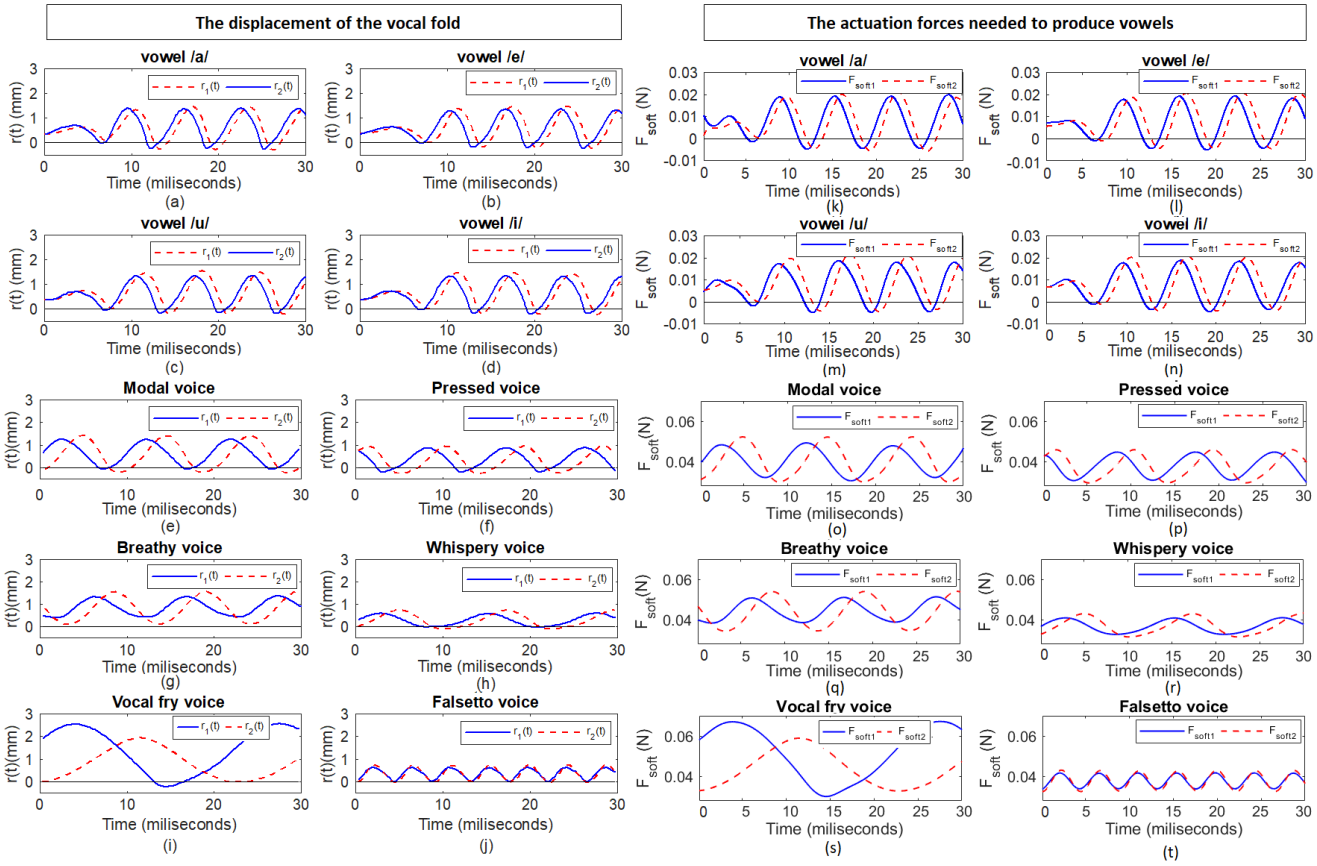

Figure 4. The displacement of the center of mass of the vocal fold $r(t)$ and the actuation forces F_{soft} which are required to producing vowels and voice qualities. The displacements are shown in: (a) Vowel /*a*/. (b) Vowel /*e*/. (c) Vowel /*i*/. (d) Vowel /*u*/. (e) *Modal* voice. (f) *Breathy* voice. (g) *Pressed* voice. (h) *Vocal fry* voice. (i) *Whispery* voice. (j) *Falsetto* voice. The actuation forces are shown in: (k) /*a*/, (l) /*e*/, (m) /*i*/, and (n) /*u*/, (o) *modal* voice, (p) *breathy* voice, (q) *pressed* voice, (r) *vocal fry*, (s) *whispery* voice, and (t) *falsetto*.

4 Actuator Selection and Model

4.1 Types of Soft Actuation

A number of soft robotic actuators could be considered for this medical application due to their inherent compliance suitable to be integrated with soft mucous tissue models of the folds, and biocompatibility requirements of materials for prosthetic vocal fold implants. In the field of soft robotics, a large range of actuation methods is covered [23]. Fluidic Elastomer Actuators (FEA) can be pneumatically or hydraulically driven achieving changes in pressure (for pneumatically driven) and volume (for hydraulically driven) that create actuation motion [24, 25]. Similarly, phase change actuation methods and chemical reaction techniques, such as combustion, also use volume change for actuation. However, their volume changes are due to the internal processes as opposed to material shifting. On the other side, shape memory materials create motion through changes in either size or shape in response to thermal changes or applied stresses due to the rearrangement of molecules within the same material at the same phase [26]. These actuators are often used in bimorph arrangements to form bending motions due relative change in lengths of the actuators. Electro-/Magneto-Active Polymers (E/MAP) also maintain material and phase but move in response to the application of an electric or magnetic field. Electroactive Polymers (EAP) are generally more common, although magnetic metal composite and magnetic rheological designs are developing [27]. The two main types of EAP are Dielectric Elastomer Actuators (DEA) and Ionic Polymer Metal Composites (IPMC).

The dimensions of any actuator chosen should ideally correspond to those discussed earlier in Section 2. For the actuator lengths, an extension and contraction of about 3.1 mm is required. Furthermore, the actuation should achieve sufficient forces to displace the masses outlined in Section 3 for the different glottal areas, i.e., maximum forces of 0.75 N at a displacement of 3.1 mm would be needed to achieve the lowest register of *vocal fry*. Another challenge of the applicability of soft robotic actuation is the large frequency range. Human phonation covers a large bandwidth of sound frequencies averaging around 120 Hz to 210 Hz for males and females, respectively, with frequency variation of about 7 semitones [28].

FEAs, shape memory, and phase change actuators are generally slow in their response requiring volumetric and thermal changes that could be problematic within the limited available space [26, 29, 30]. E/MAPs, on the other hand, are capable of achieving much greater frequencies on the relevant scale. Hence, DEAs are often chosen. However, high activation voltages

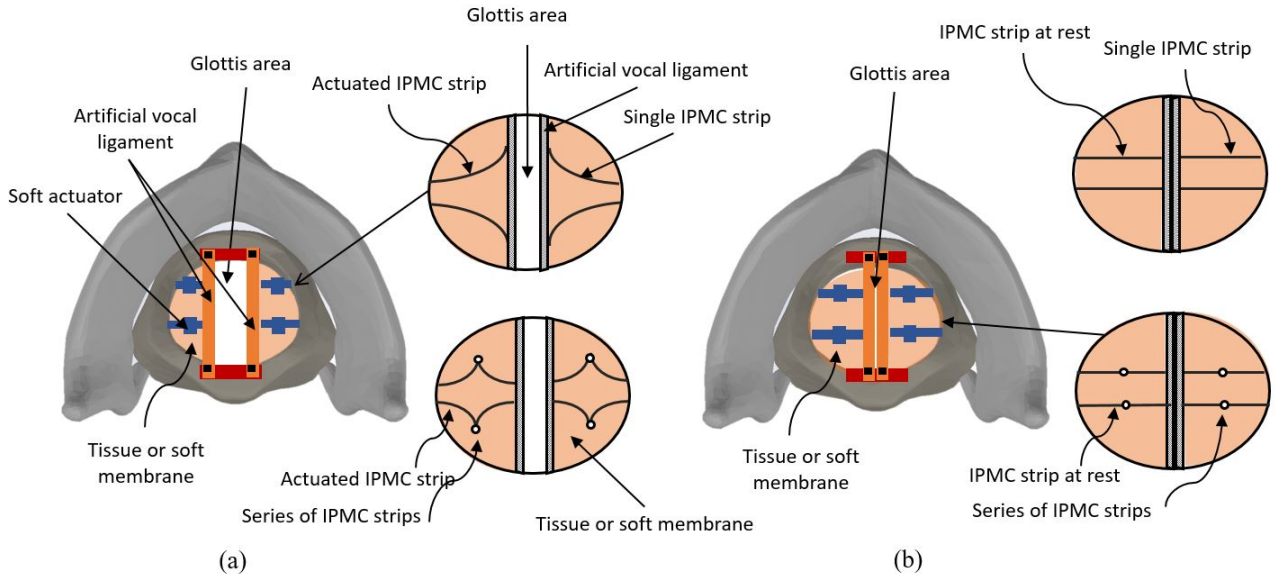

Figure 5. Configuration of IPMC actuators in the mechanism. (a) IPMC actuators excited with an activation voltage. (b) Configuration of IPMC actuators within the larynx mechanism while IPMC strips are at rest (no voltage is applied to the actuators). Figures includes two types of IPMC arrangement: a single IPMC strip and a series of IPMC strips.

(i.e., in the range of a few hundred volts) are required to achieve the required frequencies for the proposed application [31]. IPMCs though show large deformation in the presence of lower activation voltages (typically less than 5V) [32] whilst maintaining a fast response to the activation voltage [32]. In addition, they have been found to operate in wet environments, with high chemical stability and durability (it is possible to bend over 10^6 times), and are biocompatible with the human body [33]. Hence, IPMC actuators were chosen for further investigation.

4.2 Modelling IPMC Forces and Voltages

IPMCs consist of an ionomer core (typically Nafion), sandwiched by two conductive mediums such as a metal. Mobile ions can migrate due to an electric field from the anode to the cathode and cause charge imbalance which results in electrostatic forces and bending of the IPMC strip [32]. Their relationships between different activation voltages, forces and displacements of IPMC are nonlinear and IPMC designs are each distinct in material and mechanical specifications. Therefore, each design needs to be modelled individually to identify not only if they are capable of achieving the necessary blocking forces at the displacements required, but also if the magnitude of activation voltages required to create them are within reasonable limits or will cause damage to the IPMC. Two designs, labelled here as Type I and Type II, are chosen from the literature in line with the requirements for this medical application. Both configuration are then modelled to obtain the activation voltages required to produce the force and displacement relationships that were identified in Section 3 for the four vowels and six voice qualities. Figure 5 (a) shows the schematic of the mechanism while IPMC actuators are excited with an activation voltage, and Figure 5 (a) demonstrates the mechanism while IPMC actuators are at rest (no voltage is applied). It is shown that while the IPMC strips are activated they can move the artificial vocal ligaments, and change the glottis area. Figure 5 depicts two types of IPMC arrangement: a single IPMC strip and a series of IPMC strips. Using multiple IPMCs, as shown in the bottom row of Figure 5, can be applied to produce a greater force. However, space may impose limitations for using IPMCs in series.

4.3 IPMC Type I Design

Case study I investigates the IPMC strips developed by Biswal et al. in [34] using silver electrodes and a base polymer of Nafion-117 membrane (Ion Power, Inc.). The size of the IPMC strip is $20\,\text{mm} \times 5\,\text{mm} \times 0.2\,\text{mm}$ ($l \times b \times t$) with an elastic modulus of $E = 0.081877\,\text{GPa}$. In [34], the IPMC strip could be reduced to a cantilever beam with the tip angle of the beam, ϕ, and the tip displacement in x-direction. P_x, was found by applying six different voltage as listed in Table 1, column 1-3. As a cantilever beam, the relationship between blocking force in x-direction F_{soft}, the moment T_x, and tip displacement P_x equals $\frac{T_x}{P_x}$. To determine the moment T_x, the equation of calculating the bending moment of an Euler-Bernoulli beam $T_x = \frac{EI}{R}$ is used. Here, I is the moment of inertia $I = \frac{bt^3}{12}$, and R is radius of curvature, using $R = \frac{l}{\phi}$, the relative force is obtained in Equation 4.

$$F_{soft} = \frac{Ebt^3\phi}{12lP_x} \tag{4}$$

Using Equation 4, F_{soft} is obtained for 0.2V, 0.4V, 0.6V, 0.8V, 1V, and 1.2V summarised in Table 1, column 4.
By fitting a first-order polynomial function to the applied voltages and IPMC forces i.e. F_{soft}, a relationship between the applied voltage and IPMC force is obtained in Equation 5.

$$V_{IPMC} = 552.7F_{soft} + 0.3223 \tag{5}$$

where V_{IPMC} is in V and F_{soft} in N.

4.4 IPMC Type II Design

Case study II explores the IPMC strips used by Vokoun et al. in [35] made of platinum electrodes and a Nafion polymer. The size of the IPMC strip is 25 mm × 5 mm × 0.32 mm, the Young's modulus of the Nafion is 0.63 GPa, and the Young's modulus of platinum electrodes is 5 GPa. In [35], experimental work found that the blocked force increased as the Nafion thickness and Nafion Young's modulus increased, reaching up to 60 mN. In addition, the relationship between the blocked force F_{soft} and activation voltage V_{IPMC} was determined by a first-order polynomial function [35] given in Equation 6.

$$V_{IPMC} = 112.3F_{soft} + 0.1264 \tag{6}$$

where V_{IPMC} is in V and F_{soft} in N. This was then used in combination with Equation 2 to obtain the activation voltages needed to actuate the different vowels and voice qualities discussed earlier.

5 Results and Discussion

The activation voltages needed to produce the vowels and the voice qualities for IPMC types I and II are calculated and shown in Figure 6. The blue lines show required voltages to move m_1. The red dashed lines show the required voltages to move m_2. From Figures 6 (a)-(j), a voltage range of [−2V 12V] is required for IPMC type I for producing vowels */a/*, */e/*, */i/*, and */u/* with an approximate frequency of 150 Hz. In particular, voltages in the range of [17V 30V] with a frequency of 115 Hz, [17V 26V] with 100 Hz, [20V 30V] with 85 Hz, [15V 30V] with 100 Hz, [17V 37V] with 40 Hz and [18V 24V] 200 Hz are needed to produce *modal*, *pressed*, *breathy*, *whispery*, *vocal fry* and *falsetto* voices, respectively. From Figures 6 (k)-(t), it can be concluded that a voltage in the range of [−0.5V 2.6V] for the vowels and [3.6V 6V], [3.4V 5.3V], [4V 6.2V], [3.5V 5V], [3.6V 7.6V], and [3.8V 4.9V] (with the same frequency of type I) for the voice qualities of *modal*, *pressed*, *breathy*, *whispery*, *vocal fry* and *falsetto*, respectively, are needed to be applied for IPMC type II actuator. Comparing these voltage amplitudes with the literature [35–40], it can be observed that some of the magnitudes slightly exceed the maximum voltage threshold. This challenge could be solved by connecting IPMC actuators in a series configuration as shown in Figure 5 (a) and (b). The series configuration has been introduced in [41]. Hence, the required force is divided between 4 strips and, as a result, a quarter of the voltages presented in Figures 6 (a)-(j)) are needed to be applied to a single actuator. These new voltages would be 0.65V for vowels, 1.5V, 1.32V, 1.55V, 1.25V, 1.9V, and 1.25V for the six voice qualities, which are

Table 1. Tip angle, displacement and blocked force of the IPMC actuator in [34] for different activation voltages.

Applied voltage V_{IPMC} [in V]	Tip displacement in x-direction P_x [in mm]	Tip angle ϕ [in rad]	Blocked force F_{soft} [in mN]
0.2	19.8	0.11098	0.1
0.4	19.6	0.23362	0.2
0.6	19.0	0.52446	0.4
0.8	18.0	0.82714	0.6
1	16.0	1.20238	1.0
1.2	12.8	1.69516	1.8

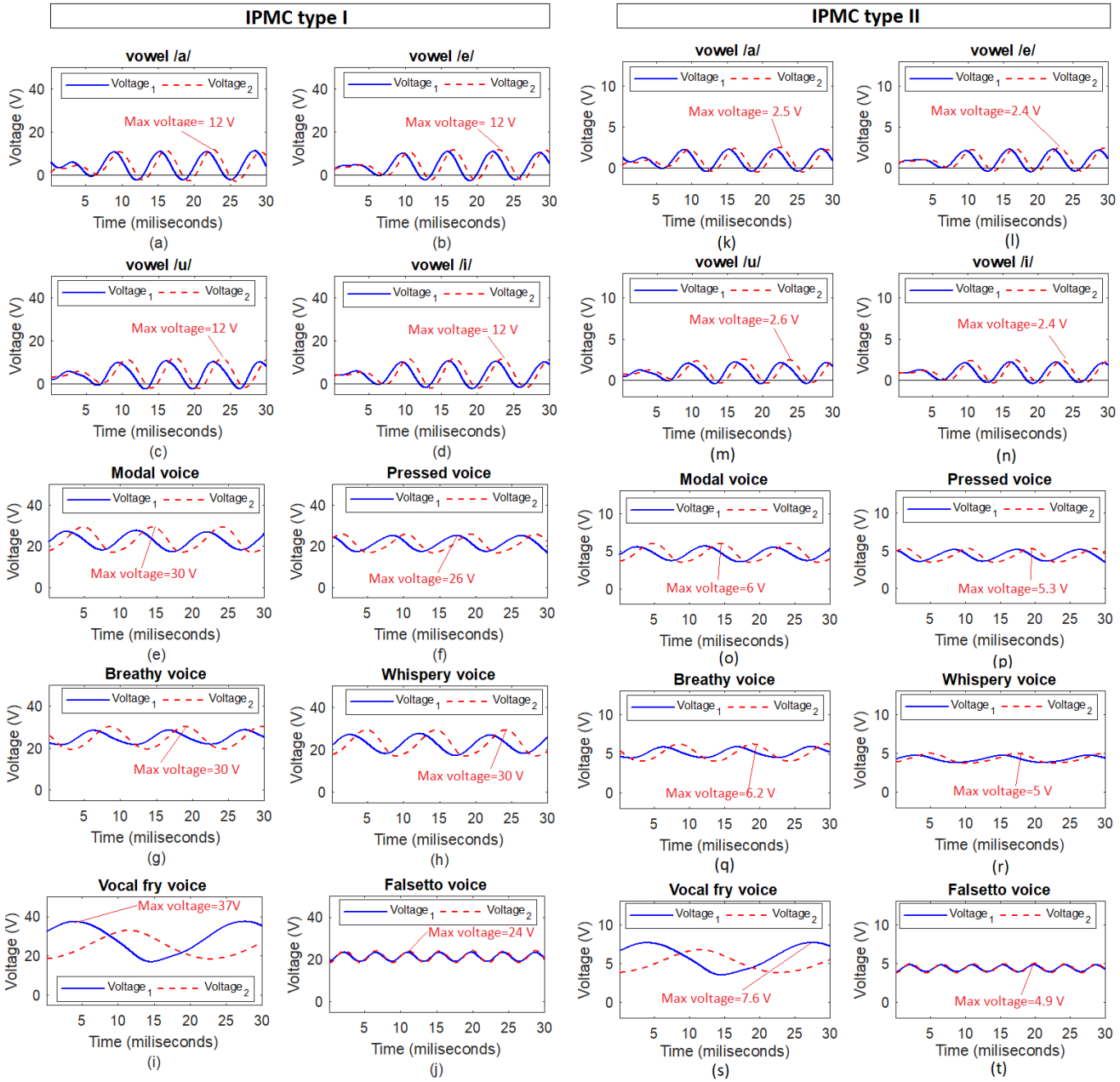

Figure 6. The determined voltages required to produce vowels and voice qualities using the type I and type II IPMC strips. Type I: (a) Vowel /*a*/. (b) Vowel /*e*/. (c) Vowel /*i*/. (d) Vowel /*u*/. (e) *Modal* voice. (f) *Breathy* voice. (g) *Pressed* voice. (h) *Vocal fry* voice. (i) *Whispery* voice. (j) *Falsetto* voice. Type II: (k) Vowel /*a*/. (l) Vowel /*e*/. (m) Vowel /*i*/. (n) Vowel /*u*/. (o) *Modal* voice. (p) *Breathy* voice. (q) *Pressed* voice. (r) *Vocal fry* voice. (s) *Whispery* voice. (t) *Falsetto* voice.

below the threshed and suitable for a single IPMC strip. With regards to the frequency analysis, Griffiths et al. has shown that IPMCs have the capacity to track high frequency bending excitation up to kHz [42]. Bonomo et al. described IPMC current sensing by adapting a piezoelectric coupling model and combining it with Euler-Bernoulli beam theory [43]. The model was demonstrated to be potentially usable for frequencies of up to 150 Hz. In practice, Ma et al. constructed a bionic high-frequency wing flapping at 19 Hz, based on IPMC actuators under an AC voltage of 5V [44].

6 Conclusions

In this paper, a design analysis was proposed to imitate natural phonation mechanisms using actively oscillated artificial vocal folds. The dynamic equations of the artificial vocal folds were used to calculate the actuation forces required to oscillate the vocal folds in the glottal areas to produce the four vowels /*a*/, /*e*/, /*i*/, and /*u*/ and six voice qualities of *whispery*, *breathy*,

pressed, *vocal fry*, *modal* and *falsetto*. Two types of IPMC were chosen to actively oscillate the vocal fold mechanism. The relationship between required voltages and resulting forces was obtained of these two IPMC designs investigating their ability to provide relevant blocked forces and displacements, in addition to activation voltages. Our results will enable further research into the effects of natural phonation and provide the foundational work for the development of advanced larynx prostheses available to post-laryngectomy patients.

The methodology presented in this paper can be used for future control system design to produce different phonations. In future work, we will combine simulations involving the IPMC as well as the mechanics of the artificial vocal fold. The desired voltage can be obtained using the inverse dynamics of the mechanism and the equations of the actuator, as presented in this paper. Then this desired voltage can be used as an input to the IPMC which can control the dynamical system of the artificial vocal folds to generate the desired phonation. We will also create prototypes of larynx prostheses with artificial vocal folds based on our proposed design with the two integrated IPMC actuators. The goal will be to reproduce the results that we determined based on our analysis with regards to displacements and frequencies.

Acknowledgement

This work has been supported by the Engineering and Physical Sciences Research Council (grant number: EP/V01062X/1).

References

[1] Gunderson, L. L., and Tepper, J. E., 2015. *Clinical radiation oncology*. Elsevier Health Sciences.

[2] Mirghani, H., Mure, C., and Mlecnik, B., 2019. "High immunoscore is associated with good response to neo-adjuvant chemotherapy and prolonged survival in advanced head and neck cancer patients". *Annals of Oncology,* **30**(Supplement 5), pp. mdz252–006.

[3] Larynx 3D model by University of Dundee and BodyParts3D,for life science, howpublished = https://sketchfab.com/3d-models/anatomy-of-the-larynx-a00bc73a303c46248db6a13a88b23404, note = Accessed: 2022-2-29.

[4] Rosen, C. A., and Simpson, C. B., 2008. *Operative techniques in laryngology*. Springer Science & Business Media.

[5] Li, W., Zhaopeng, Q., Yijun, F., and Haijun, N., 2018. "Design and preliminary evaluation of electrolarynx with f0 control based on capacitive touch technology". *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* **26**(3), pp. 629–636.

[6] Verkerke, G. J., and Thomson, S., 2014. "Sound-producing voice prostheses: 150 years of research". *Annual review of biomedical engineering,* **16**, pp. 215–245.

[7] Brownlee, B., Ahmad, S., Grammer, T., and Krempl, G., 2018. "Selective patient experience with the blom-singer dual valve voice prosthesis". *The Laryngoscope,* **128**(2), pp. 422–426.

[8] Fuchs, A. K., Hagmüller, M., and Kubin, G., 2016. "The new bionic electro-larynx speech system". *IEEE journal of selected topics in signal processing,* **10**(5), pp. 952–961.

[9] Keating, P., . E. C., 2006. "Linguistic voice quality".

[10] Kuang, J., 2017. "Covariation between voice quality and pitch: Revisiting the case of Mandarin creaky voice". *The Journal of the Acoustical Society of America,* **142**(3), pp. 1693–1706.

[11] , 2014. "Sound-producing voice prostheses: 150 years of research". *Annual review of biomedical engineering.,* **16**, pp. 215–245.

[12] Manti, M., Cianchetti, M., Nacci, A., Ursino, F., and Laschi, C., 2015. "A biorobotic model of the human larynx". In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 3623–3626.

[13] Giannaccini, M. E., Hinitt, A., Gough, E., Stinchcombe, A., and Rossiter, J., 2019. "Robotic simulator of vocal fold paralysis". In Conference on Biomimetic and Biohybrid Systems, Springer, pp. 134–145.

[14] Flanagan, J., and Landgraf, L., 1968. "Self-oscillating source for vocal-tract synthesizers". *IEEE Transactions on Audio and Electroacoustics,* **16**(1), pp. 57–64.

[15] Ishizaka, K., and Flanagan, J. L., 1972. "Synthesis of voiced sounds from a two-mass model of the vocal cords". *Bell system technical journal,* **51**(6), pp. 1233–1268.

[16] Adachi, S., and Yu, J., 2005. "Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing". *The Journal of the Acoustical Society of America,* **117**(5), pp. 3213–3224.

[17] Birkholz, P., Kröger, B. J., and Neuschaefer-Rube, C., 2011. "Articulatory synthesis of words in six voice qualities using a modified two-mass model of the vocal folds". In First International Workshop on Performative Speech and Singing Synthesis, Vol. 370.

[18] Story, B. H., and Titze, I. R., 1995. "Voice simulation with a body-cover model of the vocal folds". *The Journal of the Acoustical Society of America,* **97**(2), pp. 1249–1260.

[19] Titze, I. R., 1973. "The human vocal cords: a mathematical model". *Phonetica,* **28**(3-4), pp. 129–170.

[20] Alipour, F., and Titze, I., 1995. "Combined simulation of two dimensional airflow and vocal fold vibration". *Status and Progress Report, National Center for Voice and Speech,* **8**, pp. 9–14.

[21] Eckel, H., Sittel, C., Zorowka, P., and Jerke, A., 1994. "Dimensions of the laryngeal framework in adults". *Surgical and Radiologic Anatomy,* **16**(1), pp. 31–36.

[22] Kim, S.-Y., Joo, Y.-H., et al., 2019. "Metabolic syndrome and incidence of laryngeal cancer: A nationwide cohort study". *Scientific reports,* **9**(1), p. 667.

[23] Boyraz, P., Runge, G., and Raatz, A., 2018. "An overview of novel actuators for soft robotics". *High-Throughput,* **7**(3).

[24] Shariati, A., Shi, J., Spurgeon, S., and Wurdemann, H. A., 2021. "Dynamic modelling and visco-elastic parameter identification of a fibre-reinforced soft fluidic elastomer manipulator". In 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, pp. 661–667.

[25] Shi, J., Frantz, J. C., Shariati, A., Shiva, A., Dai, J. S., Martins, D., and Wurdemann, H. A., 2021. "Screw theory-based stiffness analysis for a fluidic-driven soft robotic manipulator". In 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, pp. 11938–11944.

[26] Jani, J. M., Leary, M., Subic, A., and Gibson, M. A., 2014. "A review of shape memory alloy research, applications and opportunities". *Materials & Design (1980-2015),* **56**, pp. 1078–1113.

[27] Jayaneththi, V., Aw, K., and McDaid, A., 2020. "Nonlinear displacement control of magnetic material actuators". *Smart Materials and Structures,* **29**(3), p. 035010.

[28] Verkerke, G., and Thomson, S.", 2014.". *Annual Review of Biomedical Engineering,* **16**, pp. 215–245.

[29] Feng, H., Sun, Y., Todd, P. A., and Lee, H. P., 2020. "Body wave generation for anguilliform locomotion using a fiber-reinforced soft fluidic elastomer actuator array toward the development of the eel-inspired underwater soft robot". *Soft Robotics,* **7**(2), pp. 233–250.

[30] van Laake, L. C., de Vries, J., Kani, S. M., and Overvelde, J. T., 2022. "A fluidic relaxation oscillator for reprogrammable sequential actuation in soft robots". *Matter,* **5**(9), pp. 2898–2917.

[31] Franke, M., Ehrenhofer, A., Lahiri, S., Henke, E.-F., Wallmersperger, T., and Richter, A., 2020. "Dielectric elastomer actuator driven soft robotic structures with bioinspired skeletal and muscular reinforcement". *Frontiers in Robotics and AI,* **7**, p. 510757.

[32] Shahinpoor, M., 2015. *Ionic polymer metal composites (IPMCs): Smart multi-functional materials and artificial muscles*, Vol. 2. Royal Society of Chemistry.

[33] Shahinpoor, M., and Kim, K. J., 2001. "Design, development, and testing of a multifingered heart compression/assist device equipped with ipmc artificial muscles". In Smart Structures and Materials 2001: Electroactive Polymer Actuators and Devices, Vol. 4329, International Society for Optics and Photonics, pp. 411–420.

[34] Biswal, D. K., Bandopadhya, D., and Dwivedy, S. K., 2013. "Investigation and evaluation of effect of dehydration on vibration characteristics of silver-electroded ionic polymer–metal composite actuator". *Journal of Intelligent Material Systems and Structures,* **24**(10), pp. 1197–1212.

[35] Vokoun, D., He, Q., Heller, L., Yu, M., and Dai, Z., 2015. "Modeling of ipmc cantilever's displacements and blocking forces". *Journal of Bionic Engineering,* **12**(1), pp. 142–151.

[36] Yılmaz, O. C., Sen, I., Gurses, B. O., and Altinkaya, E., 2019. "The effect of gold electrode thicknesses on electromechanical performance of nafion-based ionic polymer metal composite actuators". *Composites Part B: Engineering,* **165**, pp. 747–753.

[37] Bonomo, C., Fortuna, L., Giannone, P., and Graziani, S., 2006. "A circuit to model the electrical behavior of an ionic polymer-metal composite". *IEEE Transactions on Circuits and Systems I: Regular Papers,* **53**(2), pp. 338–350.

[38] Hao, L., Sun, Z., Li, Z., Su, Y., and Gao, J., 2012. "A novel adaptive force control method for ipmc manipulation". *Smart materials and structures,* **21**(7), p. 075016.

[39] Luqman, M., Lee, J.-W., Moon, K.-K., and Yoo, Y.-T., 2011. "Sulfonated polystyrene-based ionic polymer–metal composite (ipmc) actuator". *Journal of Industrial and Engineering Chemistry,* **17**(1), pp. 49–55.

[40] Hubbard, J. J., Fleming, M., Palmre, V., Pugal, D., Kim, K. J., and Leang, K. K., 2013. "Monolithic ipmc fins for propulsion and maneuvering in bioinspired underwater robotics". *IEEE Journal of Oceanic Engineering,* **39**(3), pp. 540–551.

[41] Shariati, A., Meghdari, A., and Shariati, P., 2008. "Intelligent control of an ipmc actuated manipulator using emotional learning-based controller". In Metamaterials: Fundamentals and Applications, Vol. 7029, International Society for Optics and Photonics, p. 70291J.

[42] Griffiths, D. J., 2008. "Development of ionic polymer metallic composites as sensors". PhD thesis, Virginia Tech.

[43] Bonomo, C., Fortuna, L., Giannone, P., Graziani, S., and Strazzeri, S., 2006. "A model for ionic polymer metal composites as sensors". *Smart materials and structures,* **15**(3), p. 749.

[44] Ma, S., Zhang, Y., and Liang, Y., 2019. "High-performance ionic-polymer–metal composite: Toward large-deformation fast-response artificial muscles". *Advanced Functional Materials*, p. 1908508.