2017

NEUROPHYSIOLOGICAL MUSCLE ACTIVATION SCHEME FOR CONTROLLING VOCAL FOLD MODELS

MANRÍQUEZ PERALTA, RODRIGO SEBASTIÁN

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTMENT OF ELECTRONIC ENGINEERING

Neurophysiological Muscle Activation Scheme for Controlling Vocal Fold Models

Thesis submitted by

Rodrigo Sebastián Manríquez Peralta

in partial fulfillment of the requirement for the professional title of

Electronic Engineer

and the degree of

Master of Science in Electronic Engineering

Thesis Supervisor
Matías Zañartu, Ph.D.

Valparaíso, 2017.

THESIS TITLE:

Neurophysiological Muscle Activation Scheme for Controlling Vocal Fold Models.

AUTHOR:

Rodrigo Sebastián Manríquez Peralta

THESIS PROJECT, submitted in partial fulfillment of the requirements for the professional title of Electronic Engineer and the degree of Master of Science in Electronic Engineering from Department of Electronic Engineering at Universidad Técnica Federico Santa María.

Matías Zañartu, Ph. D. _____

María José Escobar, Ph. D. _____

Sean Peterson, Ph. D. _____

Valparaíso, Octubre de 2018.

A mi madre y padre,

por su incondicional apoyo

A mis amigos,

por la confianza y complicidad.

TABLE OF CONTENTS

List of Figures

List of Tables

RESUMEN

El presente trabajo de tesis propone un modelo matemático y computacional de fluctuaciones naturales en la activación muscular. Este modelo busca caracterizar el ruido y la variabilidad asociada a la activación de los músculos encargados de la fonación, de forma de poder evaluar su impacto en un modelo numérico de las cuerdas vocales de orden reducido. La formulación matemática del modelo es descrita usando como base distintos supuestos fisiológicamente válidos respecto a la generación de los impulsos eléctricos a nivel neuronal y de fibras musculares. Para evaluar el esquema propuesto, se realiza un anális paramétrico en conjunto con un modelo de cuerdas vocales conocido como *body-cover model* (BCM), el cual es controlado por una serie de parámetros dependientes de la activación muscular. Las fluctuaciones inherentes en la activación muscular se caracterizan por tener un ancho de banda que varía con la frecuencia de disparo, además de presentar componentes de baja frecuencia en cada nivel. Al aplicar el mencionado esquema de activación al modelo producción vocal, se observan cambios dinámicos en el comportamiento de las cuerdas, en que el coeficiente de variación no presenta un comportamiento uniforme con la activación muscular. Por otra parte, las perturbaciones que se observan en las salidas del modelo de cuerdas vocales se encuentran dentro de rangos saludables. Los componentes aleatorios en el modelo muscular ejercen cambios tanto a nivel de estructura fina como en la capacidad de controlar la frecuencia fundamental de oscilación de las cuerdas vocales. El modelo propuesto constituye un novedoso avance para controlar modelos de cuerdas vocales de orden reducido en condiciones normales de voz, y también se puede extender a condiciones neuropatológicas.

ABSTRACT

A physiologically-based scheme that incorporates inherent neurological fluctuations in the activation of intrinsic laryngeal muscles into a lumped-element vocal fold model is proposed. Herein, muscles are activated through a combination of neural firing rate and recruitment of additional motor units, both of which have stochastic components. The mathematical framework and underlying physiological assumptions are described, and the effects of the fluctuations are tested via a parametric analysis using a body-cover model of the vocal folds. The inherent muscle activation fluctuations have a bandwidth that varies with the firing rate yielding both low and high frequency components. When applying the proposed fluctuation scheme to the voice production model, changes in the dynamic behavior of the vocal folds are observed, where the coefficient of variation is not uniform with muscle activation. Perturbations arising in the vocal fold model output using the proposed stochastic muscle activation scheme are within the normal range. The stochastic components of muscle activation influence both the fine structure variability and the ability to achieve a target value for pitch control. These components can have a significant impact on the vocal fold parameters, as well as the outputs of the voice production model. The proposed model constitutes a novel and physiologically-based approach for controlling lumped-element models for normal voice production and can be extended to explore neuropathological conditions.

Chapter 1

INTRODUCTION

1.1 Overview

Phonation is the primary physiological process of speech production, in which the coordinated activation of breathing and laryngeal muscles control the interaction between airflow, sound, and the vibratory activity of the vocal folds (VF). As a result, phonation determines the distinctive features of the speech production, defining the fundamental frequency (f_0), amplitude, and quality among others.

A significant amount of data describing voice production from research and clinical perspectives have been collected in the last decades using imaging and signal recording techniques. These efforts have allowed for the development of mathematical models able to reproduce different phonatory mechanism in normal and altered conditions. Of particular value has been the development of lumped-element models of the VF, since they can efficiently represent a wide range of gestures and voice quality, including the self-oscillating behavior and the modal response of the vibrating VFs [7] [8] [9]. Lumped-element VF models can be coupled with models of aerodynamic interactions and acoustical loads (sub and supraglotally), which results in a complete framework able to simulate tthe physical phenomena of speech production. [10].

Reduced order VF models can also mimic complex pathological conditions, including incomplete glottal closure [11] and nerve paralysis [12] [13], which opens the possibility of using these models in the study, diagnosis, and/or treatment of VF pathologies [14]. However, a number of gaps need to be filled before modeling can become a well-established clinical tool. One of the most critical points that lumped-element models are lacking is their relationship with the nervous system, which ultimately determines the contractile activity of the muscle. Efforts have been made to unfold the physiological relationship between laryngeal muscle activation and VF configuration for reduced order

models of the VF [15]. However, there are numerous assumptions in this initial study that need to be revisited. For instance, the effect of antagonistic muscles is overly simplified and the number of intrinsic laryngeal muscles that the scheme effectively controls is reduced to two. More importantly for the present study, the way the muscles are activated in [15] does not have a neural basis, which in turn results in fixed deterministic muscle activations. These limitations reduce the physiological relevance of reduced order VF models as well as their potential clinical impact.

In this thesis, a new neurophysiological scheme for muscle activation is proposed. This approach significantly extends prior efforts with the aim of introducing a neurophysiological description in the control of muscle behavior for a reduced order model of the vocal folds. The scheme features interspike interval variability [16], interactions between different types of muscle fibers, muscular recruitment [17] [18] using recruitment of motor units (MU) [19], and it is constructed using electro-physiologically valid parameters measured in laryngeal muscles. The proposed approach intends to capture more faithfully the main characteristics of the muscles, and therefore generates a more realistic representation of the muscle activation signal.

1.2 Motivation

This study focuses on developing a scheme that accounts for neurological characteristics of muscles into a numeric model of the VFs.

Several models have been developed to describe the neuro basis of force output during the muscle activation in the human body. For this purpose, the electrical and contractile properties of the muscle fibers were used in combination with the spatio-temporal electrical patterns of the neural population innervating the muscle. These models mimic several physiological processes of a generic muscle, which consider an arbitrary size, number of fibers, etc. [20] [21] [22], as well as the behavior of specific muscles, (e.g. the ones located in the arm) [23] [24], including the laryngeal musculature [4]. The latter study could be considered the closest approach to model neurological laryngeal muscle behavior. In spite of being a pioneer study, there are various limitations in this approach [4] that need to be pointed out, including the disregard of the intricated relationships between the VFs driving forces, VF configuration, and vibration patterns that led to an overly simplified relationship between force and perturbation in f_0, the lack of muscle recruitment, and the oversight

of the effect of other laryngeal muscles and the interactions among them. On the other hand, there is a scheme that relates muscular activation with parameters of a specific numerical model known as the body-cover model [15]. This scheme is based on a set of equations and physiological rules that define the parameters such as VF length, spring constant and masses, among others. However, the activations described in the work mentioned above are fixed constants, and therefore do not consider any natural muscular fluctuations introduced by the central nervous system. Therefore, a model that considers natural neurological fluctuations of muscular activation over time is lacking. In addition, this model needs to be based upon valid physiological assumptions, hence reflecting neurological characteristics of muscles. Although there are some models of force generation based on neural activity [20], these models have arbitrary parameters and are not adequate to descrube laryngeal muscles behavoir.

1.3 Objectives and expected contribution

The overall objective of this thesis is to develop a novel framework for a muscle activation signal with natural random fluctuations. The proposed framework will be used to model two laryngeal muscles which are responsible of phonation and will be applied in a voice production model, in a way that inherent neurological fluctuations of laryngeal muscles are incorporated into the voice modeling dynamics. With this in mind, we proceed to do a parametric analysis of the vocal fold model to evaluate the variability in the voice output given different muscle configurations. As such, the following specific aims are proposed:

1) To propose a laryngeal muscular activation model based on physiologically valid assumptions. The model should be able to generate a muscle activation signal that has random fluctuations over time.

2) To characterize the noise of the activation signal, and its variability at different levels of activation. This will be used to describe the effects produced on the VF model.

3) To evaluate the impact of the muscle activation model on a numerical model of VF, based upon its dynamic, aerodynamic, and acoustic behavior.

The work focuses mainly on laryngeal muscles, although the same principles could be used to model other muscles involved in speech production. The complete framework, with the laryngeal muscle activation model and the numerical model of VF, is expected to simulate conditions that have never been explored before. In addition, it is expected that the approach can be used to investigate the voice modeling of relevant pathologies like muscle tension dysphonia and Parkinson's disease.

Chapter 2

BACKGROUND

In this chapter, a brief description of the voice production process is presented. Likewise, relevant aspects of muscle physiology, and particularly of laryngeal muscles, are presented. In addition, a review of the typical numerical models of vocal folds and muscle activation methods are presented, along with a description of the models used in the present work.

2.1 Fundamentals of voice production

The primary process that involves the production of voice is *phonation*. Phonation is a technical term used to describe the oscillation of the VF [25]. It includes a complex interaction between physiological, biomechanical, physical, and neurological processes, which together determine voice characteristics like pitch, intensity, etc. Phonation occurs in the larynx, an organ located in the neck, where vocal folds are housed. Vocal folds (or vocal chords) are two membranes stretched from back to front across the larynx. They form a space between them known as glottis, that allows for air to pass. By vibrating, vocal folds modulate or pulsate the flow of air expelled from the lungs, thus resulting in a dipole sound source that accounts for most of the acoustic energy during speech production [25].

The most accepted theory of phonation is the myoelastic theory [26], which describes the phenomenon as a system that oscillates as a result of the interaction between aerodynamic flow and pressures coming from the lungs. In this system, the lungs act as the primary source of energy, storing it in the form of air through inhalation. During inhalation, lung capacity increases, producing a negative pressure that drives air into the lungs. While this occurs, vocal folds are separated (abduction process), allowing air to pass freely through the glottis. To induce phonation, air is expelled from the lungs while the vocal folds close together (adduction process), creating a narrow space in which air

can pass. The adduction of vocal folds as air is expelled producing an increasing pressure below the glottis, which leads to a self-oscillatory movement of the vocal folds. The oscillation of vocal folds induces a pulsating airflow that is later filtered by the vocal tract and finally radiated by the mouth [25].

The oscillation cycle of vocal folds is produced as a result of a recursive process during exhalation. When the adduction of vocal folds is produced, a convergent geometry takes place. This geometry allows for a Bernoulli flow phenomenon below the glottis, which drives the vocal folds apart. This induces a change in the geometry, up to a divergent configuration. In this configuration, a jet flow regime is established, which reduce pressure. The reduction in pressure combined with the elastic forces of the tissue restores the initial convergent configuration, thus completing the cycle [26].

Figure 2.1 shows a view of the coronal (frontal) angle of the larynx. As it can be seen, the vocal folds are located in the narrowest portion of the airway. Above the folds, the ventricular folds can be seen. The ventricular folds, also known as false chords, correspond to a muscular structure that usually does not participate in the phonation process. However, it can oscillate due to various reasons. An unhealthy use of the voice (like muscle tension dysphonia) can lead to a vibrational effect of the false chords [27]. Similarly, in some particular singing techniques, like belting, the ventricular folds are used to narrow the air space, allowing for diverse acoustical properties [28] . Similarly, in metal singing techniques like growls [29], the false chords are used to produce the characteristic harsh sound of the voice.

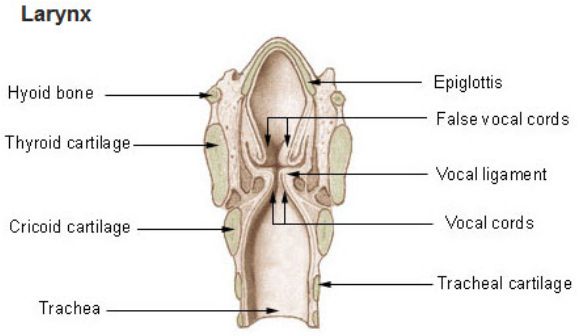


Figure 2.1: Coronal section of the larynx, showing the inner structure of the larynx. Vocal folds are located at the narrowest section of the airway. When the air passes through, the vocal chords oscillate as a result of the aerodynamic interaction of the strcuture with the air flow. Figure from National Cancer Institute. [1]

The vocal folds are comprised of multiple layers, each with different characteristics that affect the pattern of vibration. Figure 2.2 shows the different layers that compose them. The inner layer is formed of muscular tissue, which corresponds to the fibers of the *thyroarytenoid* muscle. This corresponds to a significant portion of the vocal folds, with a thickness of approximately 7 to 8 mm [30]. Outside the muscle is the *lamina propia*, a layered system that can be sectioned into three sublayers: the superficial, middle, and deep layers. The deep layer is made up primarily of collagen fibers, which have a protein-based structure that limits elongation. The intermediate layer on the other side is made up mostly of elastin fibers. Elastin fibers are made of a particular type of protein structure that allows for ample elongation, similar to a rubber band. The superficial layer is also composed of elastin fibers, surrounded by interstitial fluids, being also less uniformly oriented than the ones in the intermediate layer [30]. Finally, the outermost layer is a thin skin made up of a stratified squamous epithelium [31]. The epithelium encapsulates a softer, fluid-like tissue. Together, these five layers structure a scheme known as the five-layer division [32].

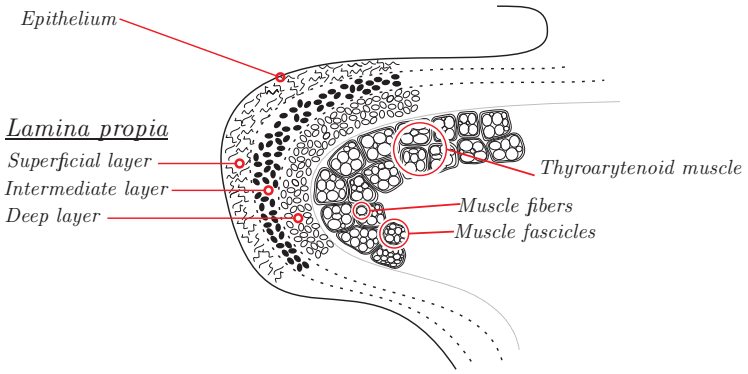


Figure 2.2: Schematic of a coronal section through a vocal chord, showing the different tissue layers. The outermost layer is the epithelium, which act as a sheat for the *lamina propia*. The lamina propia is comprised of three sub-layers, the superficial, middle, and deep layers. Finally, the innermost layer is the thyroarytenoid muscle itself.

Different layering paradigms are also used, depending on the physiological and physical properties that need to be described. For example, in a three-layer scheme, the epithelium and the superficial layer of the lamina propia are grouped in a layer labeled as mucosa, whereas the remaining layers of the lamina propia, the intermediate and deep layers, are arranged in the ligament

layer. Finally, the muscle layers refer to the thyroarytenoid layer of the vocal folds. Another paradigm of grouping layers is the two-layer configuration. In this scheme, biomechanical properties of the layers are taken into account, grouping the muscular and deep layer of the lamina propia, given their limited elongation properties. This inner layer is known as the *body*. The remaining layers, the intermediate, superficial and epithelium layers, more elastic than the former ones, are grouped in the *cover* layer. The two-layer scheme will be mentioned again in a further section to describe a mathematical model that is based in this configuration [10].

2.2 Brief review of vocal fold models

A typical simulation of VF vibration is carried out for two reasons. Either it is used for voice synthesis, or to model a specific aspect of the voice production. Depending on the level of detail required, the kinematic description of the VF oscillation can vary from simple models to highly complex and detailed models, each having its different advantages and disadvantages. According to Titze's model classification [26], numerical models of voice production can be classified into four groups. These are: (1) Low-dimensional models, (2) high-dimensional models, (3) continuum models and (4) finite models. Low-dimensional models use a lumped-element representation of vocal folds, intending to capture the fundamental aspects of phonation. These kinds of models lack a correct description of the glottal opening and other biological elements. Despite this limitation, they capture most the energy and relevant physics of voice production and have a low computational cost, this allowing for parametric analyses and optimization procedures [33]. High-dimensional models are a natural extension of low-dimensional representation. They increase the number of lumped-element models, which allows for a better description of the VF oscillation at the expense of a higher computational cost. The latter is usually implemented by finite difference methods, in contrast with the next methods. Continuum models extend the number of lumped-elements to create a solid body, therefore being able to represent an elastic continuum that is solved for analytically.. Finally, finite element models provide another approach to solve a continuum mechanics problems. They have the advantage of being able to handle tissue elements of variable sizes and shapes. By using interpolation techniques, the models become more stable and can better represent the biological structure, at the expense of a very large computational

load.

The lumped-element representation of VF has received particular attention, given the simple but complete framework that can efficiently represent a wide range of phonation scenarios. These models are structured as a collection of mass-spring-damper systems that interact with aerodynamic flows or acoustical loads. These models are coupled with a representation of the vocal tract, resulting in a complete and straightforward framework that can simulate the transmission and propagation of acoustic waves within the vocal tract, the subglottal system, and the biological tissue. As said before, the complexity of the system can be as simple as one mass, going to up to three masses for the reduced-order models, or as complex as multiple masses distributed in different layers, for the high-order descriptions. With more lumped elements, the degrees of freedom increase, better capturing the diverse aspects of the phenomenon, but also increasing the computational cost and overall complexity. For more details on specific configurations of VF models, refer to the review by Erath *et al.* [7].

Low dimensional models have proven to be a handy tool to understand the fundamental mechanisms of speech. They have been used for describing the VF production phenomena [34] [9] [7] and synthesizing natural sounding voices [8] [35] [11], including sustained vowels [36] and running speech [37]. These models have also been used to model pathological voice conditions, including incomplete glottal closure [38] and nerve paralysis [12] [13]. Modeling pathologies allow for VF models to be used in clinical assessment of vocal folds [14].

For the purposes of this research, a particular type of low-dimension lumped element model will be used: The Body-Cover Model (BCM). The BCM [10] considers a two-layer configuration: the body with one mass, and the cover with two masses (superior and inferior). The model considers a set of parameters like vibratory masses (m_u, m_l, m_b), their initial positions (x_{u0}, x_{l0}), the spring constants (k_u, k_l, k_b, k_c), and vocal fold length and thickness (L, T), among others. A diagram of the BCM is presented in Figure 2.3.

As mentioned earlier, the body-cover can be used to describe the vocal fold structure. The principal assumption in the model is that the vocal fold can be divided into two tissue layers with different mechanical properties. As said before, these two layers are the body and the cover layers (see Figure 2.3). In his work, Hirano [39] suggested that the vocal fold should be treated as a double-structured vibrator whose stiffness parameters should be based on the

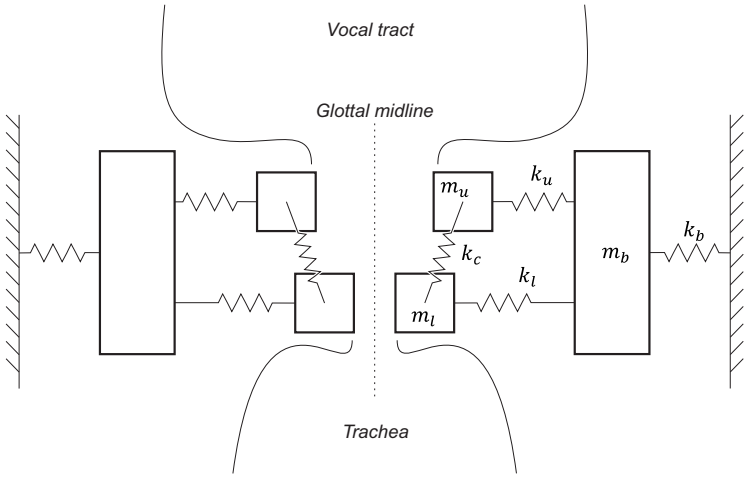


Figure 2.3: Body-Cover Model schematic. m_u and m_l represent the superior and inferior masses of the cover layer, whereas m_b correspond to the body mass.

relative actions of the thyroarytenoid (TA) and cricothyroid (CT) muscles. Based on this concept, the body-cover model was conceived by Titze and Story [10]. The body-layer was represented as a mass (m_b) attached to the thyroid wall through a spring (k_b), whereas the cover layer was described as two masses (m_u, m_l), each attached to the body mass with springs (k_u, k_l). Both cover masses are also attached between themselves through a coupling spring (k_c), which allows for vertical motion. A brief mathematical description of the body-cover model is summarized next:

- Equations of motion For the three masses, the equations of motion are written in terms of the coupling forces and other external driving forces that are applied on each mass.

$$F_u = m_u \ddot{x_u} = F_{ku} + F_{du} - F_{kc}F_{eu} + F_{uCol} \tag{2.2.1}$$

$$F_l = m_l \ddot{x_l} = F_{kl} + F_{dl} + F_{kc}F_{el} + F_{uCol} \tag{2.2.2}$$

$$F_b = m_b \ddot{x_b} = F_{kb} + F_{db} - [F_{ku} + F_{dl} + F_{kl} + F_{dl}] \quad , \tag{2.2.3}$$

  where F_u, F_l and F_b are the forces of the upper, lower (cover) and body masses respectively, written in terms of the derivatives of their positions.

Also, F_{du}, F_{dl} and F_{db} are the forces due to damping, F_{ku}, F_{kl} and F_{kb} are the lateral spring forces. F_{uCol} and F_{lCol} are the forces generated only during collision with the opposite vocal fold, F_{eu} and F_{el} are the external forces generated by the glottal flow, and F_{kc} corresponds to the spring force due to the coupling of m_u and m_l. The equations are written in terms of the derivatives of their displacements x_u, x_l and x_b, considering the initial positions x_{u0}, x_{l0} and x_{b0}. For more details on each of the driving forces, please refer to the original paper [10].

- Pressure equations

  In the BCM, the three-mass system interacts with driving forces from aerodynamic source via the glottal area. Intraglottal pressure exerts a force on the upper and lower cover masses, producing oscillation. Generally speaking, the modeling of the interaction between the masses and intraglottal pressure considers the following assumptions:

  1) The flow detaches at the minimum glottal diameter.
  2) A Bernoulli-type flow exists from the subglottal region to the minimum glottal diameter.
  3) A constant diameter jet exists from the minimum diameter to the glottal exit. Pressure is considered to be constant in this region.
  4) Pressure recovery after glottal exit follows equations derived by Ishizaka and Matsudaira (1972) [8].

  Therefore, there are two principal equations for describing the two possible flow regimes. For the Bernoulli regime:

$$P = P_s - (P_s - P_i) \times \frac{a^{-2} - a_s^{-2}}{a_m^{-2}(1 - k_e) - a_s^{-2}} \quad , \tag{2.2.4}$$

  and for the Jet regime:

$$P = P_i - (P_s - P_i) \times \frac{k_e a_m^{-2}}{a_m^{-2}(1 - k_e) - a_s^{-2}} \quad , \tag{2.2.5}$$

  where P is the pressure at any point of the upstream area and P_s and P_i are the subglottal and supraglottal pressures respectively. Also, a is the cross-sectional area of the channel, a_s is the subglottal duct area, a_m

is the minimum cross-sectional area within the glottis, and k_e is the exit pressure coefficient. During the oscillation of the VF, the flow can be blocked if collisions were to occur between the masses. During collision, pressures are computed as:

$$P = P_s \qquad \text{below the collision} \tag{2.2.6}$$

$$P = 0 \qquad \text{within the collision} \tag{2.2.7}$$

$$P = P_i \qquad \text{above the collision} \tag{2.2.8}$$

For more details on the pressure equations, refer to the original paper [10].

- Flow equations

  A vocal tract is also coupled into the three-mass model, including models for the subglottal, pharyngeal, oral, and nasal sections. This means that a wave reflection is produced, changing the characteristics of the flow. Reflected pressures p_s^+ and p_s^- are computed as follows:

$$p_s^- = p_s^+ - (\rho c/A_s)u \tag{2.2.9}$$

$$p_s^+ = p_s^- + (\rho c/A_i)u \quad , \tag{2.2.10}$$

  in which p_s^+ and p_s^- are the incident pressures above and below the glottis, respectively. ρ is the density of air, c the speed of sound, and A_s and A_i are the areas of the first sections of the subglottal and supraglottal ducts, respectively. u corresponds to the flow through the glottis, which can be obtained upon equations given by Titze [40].

The BCM was chosen for this work mainly for two reasons. First, although it is a simple model, it allows for very realistic reproduction of the dynamical behavior of the vocal folds. This simplicity translates into a low computational cost, which is ideal for working with a parametric analysis. Second, its model parameters (masses, spring constants, etc.) can be determined through a series of equations and rules depending on laryngeal muscle activation [15]. These rules are crucial to study how muscle variability affect the vocal folds dynamics.

To reduce the number of control parameters, the normalized activations of the *cricothyroid* activity (a_{CT}), the *thyroarytenoid* activity (a_{TA}), and the *lateral cricoarytenoid activity* (a_{LCA}) were considered in [10]. The effect of the *posterior cricoarytenoid* muscles was included by allowing a_{LCA} to become

negative, and finally, the effect of the *interarytenoid* muscle was neglected. A summary of the physiological rules is presented bellow:

- Elongation Rule: The elongation rule expresses the vocal fold length L in terms of three activations. The longitudinal vocal fold strain ϵ is also presented in equation 2.2.11, as further parameters also depend on this value.

$$\epsilon = G(Ra_{CT} - a_{TA}) - Ha_{LCA} \tag{2.2.11}$$

$$L = L_0 [1 + \epsilon] \tag{2.2.12}$$

  In equation 2.2.11 G, R and H are the gain of elongation, torque ratio and adductory strain factor respectively. In equation 2.2.12 L_0 is the resting length. All these coefficients are assigned according to previous studies. See [15] for more details.

- Nodal Point Rule: The nodal point z_n controls upper and lower amplitudes of vibration. It depends only on activation of TA:

$$z_n = (1 + a_{TA}) \cdot T/3 \tag{2.2.13}$$

  where T is vocal fold thickness.

- Thickness Rule: Vocal fold thickness T increases with its shortening, which implies that, in a sense, it is inversely proportional to VF length L. In this case, the relationship is better defined using VF strain ϵ.

$$T = \frac{1 + a_{TA}}{1 + 0.8\epsilon} \tag{2.2.14}$$

  where T_0 is vibrating thickness at resting state. As with VF length, T depends on activations of all three laryngeal muscles.

- Depth Rule: The depth rule D_b for the body layer and D_c for the cover layer are defined according to:

$$D_b = \frac{a_{TA} D_{mus} + 0.5 D_{lig}}{1 + 0.2\epsilon} \tag{2.2.15}$$

$$D_c = \frac{D_{muc} + 0.5 D_{lig}}{1 + 0.2\epsilon} \tag{2.2.16}$$

  in which D_{mus}, D_{muc} and D_{lig} are the muscular, mucosa and ligament depths respectively. These parameters are obtained according to clinical observations.

- Adduction Rule: The adduction rule controls the glottal half-width ξ_{02} which establishes how close the vocal process is. It has been deduced using fibroscopic measurements.

$$\xi_{02} = 0.25L_0(1 - 2a_{LCA}) \tag{2.2.17}$$

  For a value of $a_{LCA} = 0.5$, vocal folds are "just touching". Thereby, if $a_{LCA} > 0.5$ then the vocal process is pressed together (actually, chords are overlapping), and if $a_{LCA} < 0.5$ it implies that folds are separated. This is one of the most sensitive parameters in the whole system, where in little variations on the LCA activation can produce significant changes in the behavior of the system.

- Convergence Rule: The convergence rule establishes a relationship for the separation of the lower portion of the vocal process.

$$\xi_c = \xi_{01} - \xi_{02} = T(0.05 - 0.15a_{TA}) \tag{2.2.18}$$

  To reach self-sustained oscillations in the BCM, a nearly vertical medial surface is necessary. This produces a "nearly rectangular" section for the glottis, which is ideal for phonation in a physical model [41].

For details referring to the model and the activation rules, refer to Titze's original papers [10] [15].

2.3 Muscle physiology

2.3.1 Basic concepts

There are three major muscle groups: cardiac, smooth and skeletal. *Cardiac muscles* are found in the walls of the heart, and are in charge of involuntary contractions in this organ. Their most relevant characteristic is that they contract in single twitches. On the other hand, *smooth muscles*, which are found in organs such as the stomach, intestines, and uterus. They have different mechanical properties depending on their function and type. Finally, *skeletal muscles*, which are responsible for most voluntary contractions on the body. They differ from cardiac muscles in the sense that voluntary nerves control them, generating multiple twitch contractions depending on the frequency of stimulation. Continuous contractions can build up to a maximum tension,

known as *tetanization*. Movement is produced by the action of skeletal muscles, being therefore bound into bones to provide support. Biceps, extra-ocular muscles and laryngeal muscles are examples of skeletal muscles [42].

Muscles are structurally composed of fibers. Fibers within a muscle are grouped in bundles known as *fasciculi*, of various sizes depending on the muscle. Space between fasciculi is filled with a strong connective tissue that keeps together the whole structure. Fibers within fasciculi are also filled with another type of connective tissue. A fiber itself is a specific type of cell, composed of many nuclei and a longitudinal structure called *myofibril*. Myofibrils are long chains of *sarcomeres*, which corresponds to the contractile unit of the fiber [42].

The smallest functional element in a muscle is the Motor Unit (MU). A MU is comprised of a motor neuron and all the muscle fibers innervated by its axon [43]. Muscles develop force through the combined effect of the contractions of individual fibers, which are triggered as a response to trains of action potentials (or spikes) transmitted to the muscle fibers by motor axons. Motor Units operate as an ensemble, working in conjunction to create stronger contractions and force.

Not all MU are the same. First, the number of fibers per motor neuron vary in MU in the same muscle, and differ even more from muscle to muscle; thereby defining the *innervation ratio*, which is the number of mean fibers per motor unit. As a general rule, muscles with low innervation ratio have more control over the force output. For example, biceps muscles in the arms have over thousand fibers per motor neuron, while extra-ocular muscles have a mean of ten. On the other hand, MUs in a muscle can also be different depending on the type of fibers innervated by the motor axon. In an arbitrary MU, all the fibers innervated are of the same kind, leading to a classification of motor units based on the fatigability and contractile speed [44] [45]. If an MU does not exhibit "sag" (decline in force after the initial increase during unfused tetanic stimulation) then it is classified as slow contracting, or type S. However, if an MU does exhibit sag, then it is denoted as fast contracting, or type F. Additionally, type F motor units can also be subclassified depending on fatigue properties, being fast to fatigue (type FF) or fatigue-resistant (type FR). It is also possible to classify motor units based on their histochemical properties, but this is not of real interest for this study.

The central nervous system controls the muscle contraction through MU by two primary mechanisms: (1) successive action potentials (spike train) which can be measured as a discharge or fire rate, and (2) the recruitment and

derecruitment processes which regulate the number of active MU [46] [47] [48]. By modifying the firing rate, the nervous system can control the number of successive twitches that are summed at a single fiber level; and by adding more active MU, then more fibers are incorporated in the contraction process, increasing the resulting force.

The motor unit recruitment is not arbitrary, and it follows an orderly sequence based on the fiber classification previously mentioned. For low force outputs, slow-twitch motor units are recruited first; and as more force is required, fast-twitch motor units begin to be recruited. As a general rule, this means that type S motor units are activated before type F motor units [19] [17]. This is known as the size principle for motor unit recruitment [18].

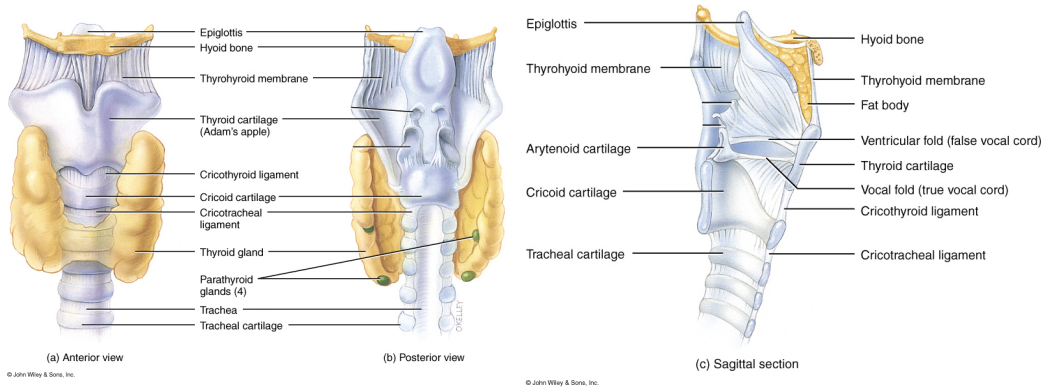
As it has been previously mentioned, motor unit firing rate determines the contraction level at single fiber level. There are many factors like, conduction velocity and post-synaptic potentials, that affect the timing in which action potentials are generated. This variability in neuronal discharge is usually described in terms of the times between successive action potentials, or Inter Spike Intervals (ISI). For motoneurons in muscles, the coefficient of variance (CV) for the ISI ranges from 10% to 30% during voluntary contractions according to some experiments [49] [50]. The term *synaptic noise* is commonly used to refer to the random fluctuations in membrane potentials, whereas tiny variations on the spike trains that arrive at a motoneuron population are described as low-frequency oscillations of the synaptic input. Ultimately, these neuronal fluctuations produce variations in the force output, which are partially associated to ISI variability [16], and also with low-frequency oscillations presented in the synaptic input [51] [52].

2.3.2 Laryngeal Muscles

Regarding voice production, the main anatomical elements that determine the sound output, are the chest, neck and the head. Therefore, muscles in lungs, vocal tract, larynx, pharynx and oral cavity determine the different characteristics of the voice. From an oscillatory perspective, the most important muscles are located in the larynx (intrinsic laryngeal muscles), in which vocal folds are housed. Vocal folds vibrate as result of the interaction of the tissue with the air flow coming from the lungs. The vibration of the vocal folds is the one which gives the voice the oscillatory and periodicity properties, such as pitch; hence the importance of the larynx in the voice production process.

The larynx is an organ situated in the neck, connecting the inferior part

of the pharynx with the trachea and it is involved in actions like breathing, phonation and also protecting the trachea against food aspiration. The larynx is composed by 6 cartilages, three single: *thyroid*, *cricoid* and *epiglottis*; and three paired: *arytenoid*, *corniculate* and *cuneiform*. On the other side, muscles in the larynx can be classified into two groups, *extrinsic* and *intrinsic*. Extrinsic muscles connect the larynx to other external structures and organs, like the hyoid bone; whereas intrinsic muscles interconnect the cartilages of the larynx. Figure 2.4 presents a lateral, anterior and posterior view of the larynx. There are five intrinsic laryngeal muscles, each responsible for connecting two of the previously mentioned cartilages [32].



(a) (b)

Figure 2.4: Anterior, posterior and sagital planes of the larynx. Figure from *Anatomy of Larynx*, University of Liverpool website [2]

The *thyroarytenoid* (TA) is a paired muscle that connects the thyroid with the arytenoid cartilages. They comprise most of the structure of the vocal chord itself, and the resulting contraction of this muscle draw the arytenoid cartilages forward. By doing so, the muscle stiffens and the VF shorten. Interestingly, the TA shows a great difference in phenotype compared with limb muscles [53].

The *cricothyroid* (CT) is also a paired muscle, and is responsible of vocal chords tension, producing elongation in anteroposterior direction. By controlling the length of the cords, the CT is considered one of the primary pitch control muscles [54]. The cricothyroid is the only one of the intrinsic muscles that is innerved by the superior laryngeal nerve (SLN).

The *lateral cricoarytenoid* (LCA) is another paired muscle that functions as an adductor by drawing the arytenoids forward and medially. It brings the

Table 2.1: Reported contraction times (in miliseconds) for the thyroarytenoid (TA), cricothyroid (CT), lateral cricoarytenoid (LCA) and posterior cricoarytenoid (PCA) muscles of various species

Species	TA (ms)	CT (ms)	LCA (ms)	PCA (ms)	Reference
Rat	-	7.2	-	3.4	Hinrichsen & Dulhunty (1982) [60]
Rabbit	6.5	23-30	-	-	Hall-Craggs (1968) [63]
Cat	22	52.8	-	-	Hast (1967) [58]
	21	44	-	22	Hirose *et al* (1969) [59]
Dog	14	35	16	30	Martensson & Skoglund (1964) [3]
	14	39	-	-	Hast (1966) [55]
	-	-	-	33.2	Cooper*et al* (1994) [56]
	24	-	-	33.5	Perlman & Alipour-Haghighi (1988) [57]
Squirrel monkey	13.2	18.8	-	-	Hast (1969) [62]
Rhesus macaque	14	36.4	-	-	Hast (1969) [62]
Gibbon	16	39	-	-	Hast (1969) [62]

VF together, and also produce rotation on the arytenoids.

The *posterior cricoarytenoid* (PCA) is a paired muscle that functions as the primary abductor of the VF. It brings the VF away from the midline, opposing the action of the LCA. Both muscles -LCA and PCA- work as an agonist-antagonist pair, which means that their actions oppose each other.

Finally, the *interarytenoid* (IA) is a muscle that is comprised of two parts: The traverse component (unpaired) and The oblique part (paired). The principal function of the IA is to aid the LCA in his adduction function, closing the glottis. In particular, the IA is responsible for sealing off the posterior glottis.

Contraction properties of laryngeal muscles have been studied in mammals like dogs [3] [55] [56] [57], cats [58] [59], rats [60], and some primates [61] [62]. Mean contraction times for intrinsic laryngeal muscles are reported and summarised in Table 2.1.

From this table, we observe that contraction times of these muscles are influenced by the body mass of the animal. For smaller species, contraction times tend to be shorter. Second, contraction times for the TA muscle tend to be shorter that the other muscles, being in the range of vast extraocular

Table 2.2: Percentages of type I (slow twitch oxidative) fibers in the thyroarytenoid (TA), cricothyroid (CT), lateral cricoarytenoid (LCA) and posterior cricoarytenoid (PCA) muscles of various species, including humans. For some species, the proportion in the TA muscle is reported considering the lateral and medialis division respectively

Species	TA (ms)	CT (ms)	LCA (ms)	PCA (ms)	Reference
Rat	0	-		10	DelGaudio *et al.* (1995)
	-	16	-	2.7	Hinrichsen & Dulhunty (1982) [60]
Rabbit	0	35	-	40	Asmussen & Wohlrab (1972) [67]
Cat	10	40	-	40	Edstrom *et al.* (1974) [68]
	1/16	37	-	32	Mascarello & Veggetti (1979) [69]
	9	45		25	Yokoyama *et al.* (1995) [70]
Dog	13/21	45	-	37	Mascarello & Veggetti (1979) [69]
	5	40	-	40	Braund *et al.* (1988) [71]
	-	-		35	Sanders *et al.* (1993) [72]
	20	45		-	Perlman & Alipour-Haghighi (1988) [57]
Human	37	43	-	65	Happak *et al.* (1989) [65]
	35	47	40	67	Teig *et al.* (1978) [64]

muscles of the same species. In contrast, contraction times for the CT muscle are from two to four times longer, in the range of fast limb muscles of the same species. Meanwhile, PCA has an intermediate contraction time values, whereas information for LCA is scarce.

Studies in humans are uncommon for laryngeal muscles, because they can only be done in cadaveric bodies which are usually not ideal to study dynamical properties, like contraction times. However, fiber proportion for CT, TA, and PCA can be measured. Data for humans [64] [65] [66] and other species are summarized in Table 2.2.

From this Table, it can be seen that small mammals (like rats) are entirely devoid of slow fibers in the TA muscle, incrementing the proportion as the size of the animal increases. Also, it can be seen that type I fiber content correlates with contraction time. For example, in the cat, TA contains 10% of

slow fibers with a contraction time of 22[ms], whereas CT contains 40%-45% with a contraction time of 52.8[ms]. This correlation holds true for the rabbit, in which TA has very few type I fibers (contraction time of 6.5[ms]) while CT (contraction time 24–30 ms) has 35% type I fibers.

For their experiment, Martensson & Skoguld [3] determined the frequency range of stimulation at which maximum tension is obtained. Using direct nerve stimulation on laryngeal muscles in dogs, he established that for CT and TA maximum tension occurs at 150[Hz] (see Figure 2.5). At higher frequencies, a decrease in maximum tension can be observed, which can be partly explained as a reduction in action potential amplitude.

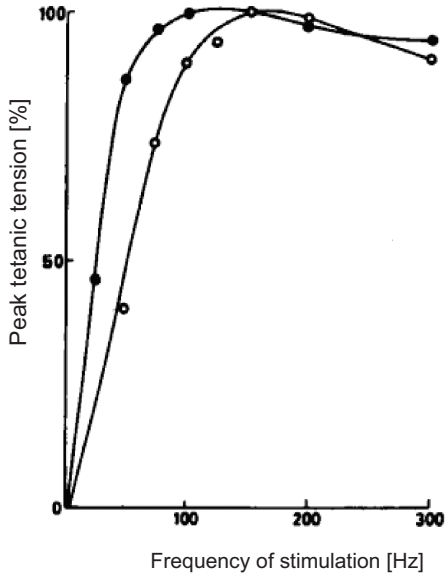


Figure 2.5: Relation between peak of tetanic tension and frequency of stimulation. Full circles: cricothyroid (CT); open circles: thyroarytenoid (TA). Peak tetanic tension (Y-axis) is presented in per cent of maximum tension obtained. Figure from Martensson & Skoguld (1964) [3]

2.4 Models of muscle activation

Mathematical and numerical models of muscles describe muscle force through a series of parameters such as muscle length, neuronal activation, shortening velocity and muscle architecture [73] [74] [75] [76] [77]. Normally, these studies are based on Hill's muscle model, which relates tension to the velocity with

regard to the internal thermodynamics.

$$(v + b)(T + a) = b(T_0 + a) \tag{2.4.1}$$

Where T is the tension (or load) in the muscle, T_0 is the maximum isometric tension, v is the velocity of contraction and a, and b are constants. This state equation is typically used in biomechanical studies but does not offer a force representation of the neurophysiological source.

As it has previously mentioned, the nervous system controls the motor drive through motoneurons that trigger fibber contraction in muscles. Therefore, most muscular models that seek to represent muscle force based on neuronal activity incorporate the concept of a pool of motor units, structuring the model around some combination of the effect of individual units. The model developed by Fuglevand *et al.* is one of the most important in this regard [20]. Fuglevand's model is comprised of three elements: a motoneuron model, a motor-unit force model, and a model of the surface EMG. The motoneuron model considered an excitatory drive function as an input, which represents the net synaptic input during voluntary muscle contraction. The twitch force model was estimated as an impulse response to a critically damped system of second order, with twitch amplitudes being assigned according to rank in the recruitment order. Finally, the EMG was synthesized as the sum of all motor-unit action potential trains generated in the motor units. This model also considers two recruitment conditions, which were tested to establish which one suits a better EMG-force relationship. Firing rate and MU recruitment were used as the primary mechanisms of drive control, based on their importance regulating muscular force [46].

Other models take Fuglevand's scheme as a starting point, working around aspects like recruitment and firing rate to configure the most complex and accurate models [21] [22]. Despite this, it has been stated by some studies that this is insufficient to describe the considerable variability on twitch forms in real muscles [78]. Thereby, Song *et al* [79] propose that by including different type of fibers -slow and fast twitches- this variability can be addressed. Other models also state that firing rate is difficult to measure empirically, so they determinate this as a parameter estimated employing genetic algorithms and optimization processes [80]. Other models include different mechanisms regarding FF and FR fibers so that fatigue can be included in these muscular models [81] [82]. By considering the principal mechanisms regarding muscle force production, models for specific muscles like the gastrocnemius [83] [24],

dorsal interosseus [84], and bicep brachii [23] have been developed, which have been validated using data of the force profile, EMG measurements or ultrasound images.

Although there have been attempts to model the behavior and mechanics of intrinsic laryngeal muscles, these studies usually refer to other aspects like stiffness [85], stress [86] or posture [87]. Modeling laryngeal muscle force (or activation) is a difficult task because validation usually requires data that is hard to acquire, and the process is too invasive too have normal conditions of phonation. Despite this, Titze [4] developed a quantitative model of the ripple of vocal fold tension, based on neurological properties of the *thyroarytenoid* laryngeal muscle. He considered data from a previous study (Alipour *et al.* 1987) in which canine TA was stimulated in vitro at various activation frequencies. By measuring the force in the muscle, he deduced that average force increases with activation frequency, although not in a linear fashion. The results of this are presented in Figure 2.6.

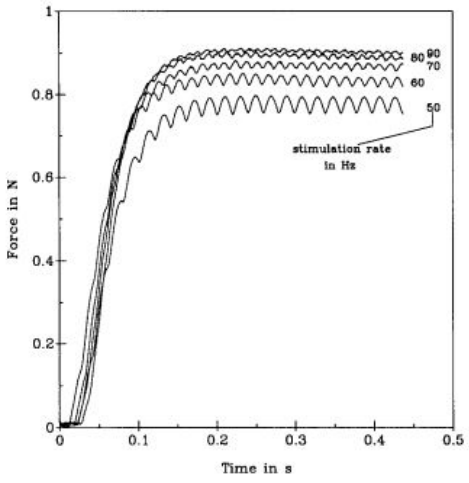


Figure 2.6: Summation of multiple twitches produced by periodic stimulation at various rates. Tetanus is produced at about 90[Hz]. Data are from the canine TA muscle stimulated in vitro (after Alipour, Titze, & Durham, 1987). Figure from Titze 1991 [4]

Based on the electrical and contractile properties of the muscle fibers, and in combination with the spatio-temporal electrical pattern of the neural population innervation, he developed a muscular model that can represent muscular fluctuations. Titze based his model in a periodic twitch summation of multiple

motor units. An example of a simulation from his model with one motor unit can be seen in figure 2.7

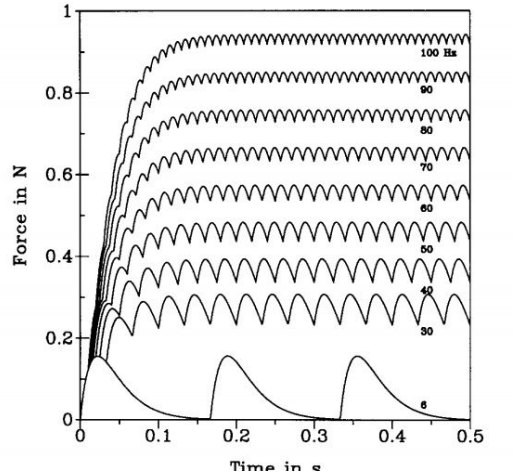


Figure 2.7: Summation of multiple twitches produced by periodic excitation at various rates with the model. Figure from Titze 1991 [4].

To the twitch summation, Titze added random variability to the periods between spikes, or Inter-Spike Interval. Moreover, he added amplitude variability to the model be considered a random number of motor units for each simulation. With this, he created a basic model of activation fluctuations. This variability is then used to predict perturbations on fundamental frequency (f_0) of the vocal folds oscillation using a linear relationship between muscular force and f_0. Titze deduced from his simulations that perturbations measured in f_0 (coefficient of variance and jitter) are highly dependent on the contractile dynamic of the *thyroarytenoid* muscle. His results are presented in Figure 2.8:

Perturbations are calculated as a function of the size (motor unit count) and firing rate. Predicted perturbations range from 0.2 percent to 1.2 percent, which corresponds to the range of healthy voices [88]. In spite of being a pioneer study, there are various limitations in [4] that need to be pointed out, including the disregard of the intricated relationships between the driving forces, VF configuration, and vibration patterns that led to an overly simplified relationship between force and perturbation in F_0, the lack of muscle recruitment that is known to control muscle contraction [46], and the oversight of the effect of other laryngeal muscles and the interactions among them.

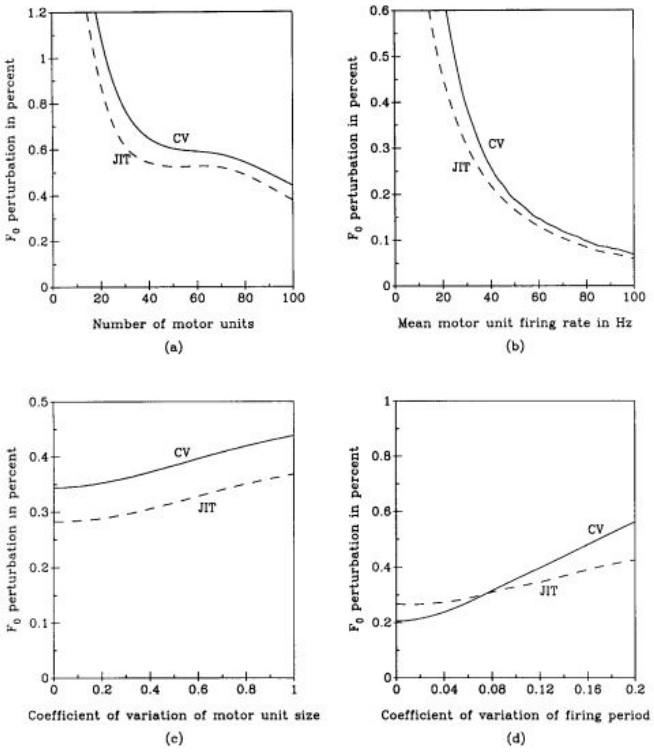


Figure 2.8: Predicted fundamental frequency F_0 perturbations on the basis of (a) number of motor units, (b) mean motor unit firing rate, (c) coefficient of variation of motor unit twitch amplitude, and (d) coefficient of variation of the ISI. In all figures, CV is the coefficient of variation of the F_0 contour and JIT is the jitter. Figure from Titze 1991 [4]

Chapter 3

METHODS

3.1 Physiological and Morphological Aspects of Muscle Activation

Activation of the laryngeal muscles comprises two major physiological processes responsible for muscle force production, namely, the temporal and the spatial summations of the muscle contraction [6, 19]. Temporal summation was modeled at the level of individual motor units, which are composed by an *alpha motor neuron* (spinal motor neuron) and the muscle fibers that it innervates [6, 43]. The spatial summation consists of the successive activation of additional motor units with increasing strength of voluntary muscle contraction; –*i.e.*, motor unit recruitment [19] [89].

Fibers forming an individual motor unit respond synchronously to every action potential (AP) arriving to the neuronal pre-synaptic terminal, producing a motor unit action potential (MUAP), that is, the algebraic sum of the APs generated in all the fibers corresponding to a given motor unit. In turn, MUAPs lead to muscle contraction, the extent of which depends on the firing rate of the motor unit. A single MUAP leads to a simple twitch (single contraction), allowing the fibers to return to a relaxed baseline before a subsequent contraction is elicited.

Addition of motor units to the force production occurs in a precise sequence concurrently with the temporal summation of twitches. Typically, the first motor units to fire are those which generate the slowest and the smallest twitches, producing relatively small and slow contractions (type I motor units). As a considerable force is required, high threshold MUs generating faster and larger twitches begin to respond (type IIa and IIb motor units) [17] [18]. Figure 3.1 shows a sketch of the time course of slow and fast twitches, highlighting the differences in both timescale and amplitude of the responses.

Twitches superimpose as the discharge rate increases, leading to stronger

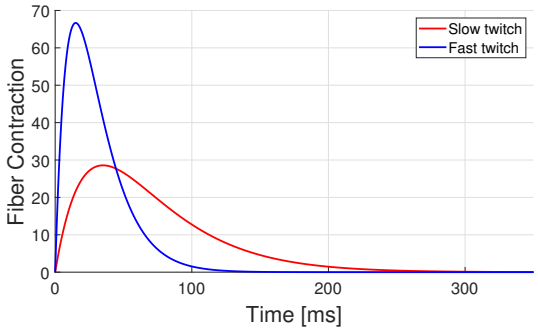


Figure 3.1: Examples of fast and slow twitch waveforms. Slow and fast twitches are normalized by the area under the curve, so the contribution in terms of energy is the same for both fibers [5].

muscle contraction. Linear superposition is referred to as the wave summation model [6]. Figure 3.2 shows the contractile force as a function of MUAP firing rate for a single motor unit. At low firing rates, a given twitch almost completely relaxes before the next twitch occurs, leading to low frequency undulations and a low net force of contraction. Conversely, at high firing rates the superposition of twitches leads to a fast rise and larger "steady state" contractile force magnitude, with small high frequency fluctuations. At a sufficiently high firing rate, a MU will cease to increase its contractile force with further increases in firing rate, referred to as tetanus.

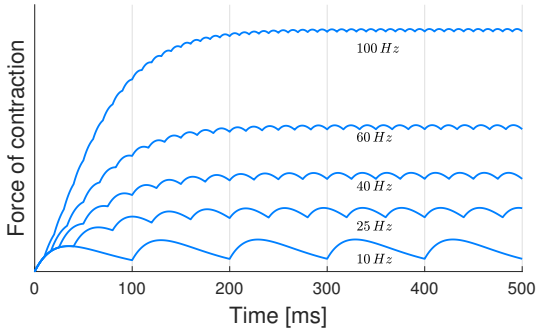


Figure 3.2: Wave summation model for different neuron firing rates. Each time a MU fires, a twitch is generated. The sum of successive twitches is the basis of the wave summation model [6].

Figure 3.2 is an idealized representation of the wave summation process, wherein the MUAP interval is a constant (deterministic) value. In actuality, biological systems exhibit some stochasticity, with the interspike intervals (ISI)

for MUAPs being no exception. In addition, recruitment of subsequent MUs, while exhibiting an overarching structure, also demonstrates some randomness in the process.

3.2 Muscle activation model

Human skeletal muscles typically comprise hundreds of MUs, with both the MUAP frequency and the number of recruited MUs dictating the total contractile force of the muscle. A single MU contraction (twitch) can be described using

$$\ddot{\alpha}(t) + \frac{2}{\tau}\dot{\alpha}(t) + \frac{1}{\tau^2} = u(t), \tag{3.2.1}$$

where $u(t)$ represents the MUAP as input, τ stands for the system time constant, and $\alpha(t)$ represents the resulting contraction force of the fibers. Consequently, the impulse response of the system is represented by

$$\alpha(t) = \frac{t}{\tau}e^{-(t-\tau)/\tau}, \qquad t \geq 0 \tag{3.2.2}$$

which characterize how an MU responds to an electric impulse (spike) [5]. Equation 3.2.2 is known as *alpha synapse function*. Herein, type I and II fibers can be differentiated by their time constants τ_s (slow) and τ_f (fast), respectively. All MUs for a given muscle are assumed to have the same number of fibers, independent of their type. Equation 3.2.2 should be scaled according to the magnitude of the response of the different fibers. However, due to the lack of available data on the laryngeal muscle fibers, a normalization by the area is performed to approximate the differences in amplitude of the slow and fast fibers. The normalized version of the alpha function results in the following equation

$$\alpha_\tau(t) = \frac{t}{\tau^2}e^{-t/\tau}, \qquad t \geq 0 \tag{3.2.3}$$

Figure 3.1 presents a plot of Equation 3.2.3 for both slow and fast twitch fibers. This waveform will serve as the foundation for the muscle activation scheme proposed herein to capture both the temporal and spatial summation processes.

To describe the *spatial summation*, MU recruitment is modeled via the *Rule of Five* (ROF), wherein additional MUs are recruited when currently activated MUs experience an approximately 5 Hz increase in MUAP firing

rate [90]. To facilitate modeling of the recruitment of MUs, we assume the MUs to be functionally bundled into clusters, herein referred to as a Group of Motor Units (GMU). GMUs can consist of both fast and slow MUs, the proportions of which will dictate the overall contraction speed of the GMU. GMUs are assumed to follow the ROF for recruitment.

GMUs are composed as follows: A fixed number of GMUs N is first defined for a given muscle. Slow-fiber MUs are assigned to the first GMUs, until all slow MUs are assigned. Fast-fiber MUs are then assigned to the remaining GMUs. Note that depending on the proportion of slow and fast fibers in a muscle, there could be a GMU with mixed fibers. GMUs then are recruited by the ROF from first to last (slow fibers to fast fibers), allowing for the recruitment of all slow fibers first.

To implement the ROF, we employ a parameter F to control the firing rate of the GMUs. The firing rate for a given GMU $j \in \{1, \ldots, N\}$ is governed by

$$\boldsymbol{F}_j = \min\left\{\max\left\{F - 5(j-1) + \boldsymbol{\eta}, 0\right\}, F_{\max}\right\}, \tag{3.2.4}$$

where $F_{\max}$ is the maximum firing rate that a GMU can physically sustain, *i.e.*, the firing rate at which the GMU tetanizes. Note that if a GMU has a firing rate $\boldsymbol{F}_j$ of zero, it is considered to be inactive, with no force contribution to the output. The parameter $\boldsymbol{\eta} \sim \mathcal{N}(0, \sigma_F)$ is a random noise term to capture the inherent variability in the ROF; that is, subsequent MUs may not be recruited at exactly a 5 Hz increase in F. Herein, bold font is used to indicate stochastic parameters and functions.

The stochasticity inherent in the arrival of a MUAP is captured in the *temporal summation* process by incoporating a random component into the interspike interval (the interval between any two subsequent AP spikes). For a given GMU j with firing rate F_j sampled from Equation 3.2.4, we construct an impulse train $\mathbf{III}_i(t, F_j)$ with interspike interval drawn from $\mathcal{N}\left(1/F_j, \mathrm{CV}_e(F_j)/F_j\right)$ for each MU i. The coefficient of variation (*i.e.*, standard deviation divided by mean), $\mathrm{CV}_e(F_j)$, derived from the experimental data of Mortiz [16], is given as

$$\mathrm{CV}_e(F_j) = \left(1 + e^{-F_j/50}\right)/10, \tag{3.2.5}$$

which captures the observed change in behavior across firing rates. This implementation implies that the CV_e ranges from 0.2 for lower activation frequencies to 0.1 at higher firing rates.

The pulse train for a single MU is given by the convolution of the alpha function (equation 3.2.3) with the corresponding impulse train :

$$\boldsymbol{p}_{MU_i}(t) = (\mathbf{III}_i(t, F_j) * \alpha_\tau)(t) \tag{3.2.6}$$

where τ is determined whenever the fibers in the MU are slow or fast. Given that for modeling purposes each GMU is composed of slow or/and fast MU, the contribution of a single GMU is determined by the sum of all the MU comprising the group. If the GMU is comprised of only one type of MU (slow or fast), then the resulting activation for that given GMU is ruled by the following equations:

$$\boldsymbol{p}_s(t, F_j) = \frac{1}{M_s} \sum_{i=1}^{M_s} (\mathbf{III}_i(t, F_j) * \alpha_{\tau_s})(t) \tag{3.2.7}$$

$$\boldsymbol{p}_f(t, F_j) = \frac{1}{M_f} \sum_{i=1}^{M_f} \left(\mathbf{III}_i(t, F_j) * \alpha_{\tau_f}\right)(t) \tag{3.2.8}$$

where M_s and M_f are the number of slow and fast MUs in the GMU respectively. Equation 3.2.7 is used for GMUs that are first activated, with $j = 1, 2, \ldots$; whereas equation 3.2.8 holds true for the last-recruited GMU, with $j = \ldots, N-1, N$, given the ROF for recruitment. Depending on the proportion of slow/fast fiber for the muscle, a specific GMU may have a combination of slow and fast MUs. For this specific GMU, the following equation holds true.

$$\boldsymbol{p}(t, F_j) = \frac{1}{M_s} \sum_{i=1}^{M_s} (\mathbf{III}_i(t, F_j) * \alpha_{\tau_s})(t) + \frac{1}{M_f} \sum_{i=1}^{M_f} \left(\mathbf{III}_i(t, F_j) * \alpha_{\tau_f}\right)(t) \tag{3.2.9}$$

Values for M_s and M_f are specific for each GMU, although indexes are omitted for simplicity.

Finally, muscle activation, which is a normalized representation of the contractile force exerted by a given muscle [15], is given by

$$\boldsymbol{a}_m(t) = \frac{\sum_{j=1}^{N} \boldsymbol{p}(t, F_j)}{\mathrm{E}\{\sum_{j=1}^{N} \boldsymbol{p}(t, F_{\mathrm{tet}})\}} \tag{3.2.10}$$

where $\mathrm{E}\{\cdot\}$ is the expectation operator as $t \to \infty$ and

$$F_{\mathrm{tet}} = F_{\max} + 5(N - 1) \tag{3.2.11}$$

is the firing rate for a fully tetanized muscle (all GMUs fully activated). In this manner, a fully tetanized muscle is given by $\mathrm{E}\{\boldsymbol{a}_m\} = 1$, whereas $\mathrm{E}\{\boldsymbol{a}_m\} = 0$ represents a fully relaxed muscle. We highlight the fact that $\boldsymbol{a}_m$ is a function of our firing rate control parameter F introduced in Equation 3.2.4. The nonlinear mapping between these parameters will be discussed in subsequent sections.

3.3 Model parameters

Two intrinsic laryngeal muscles are considered in this study due to their importance in pitch control during phonation [54]: *thyroarytenoid* (TA) and *cricothyroid* (CT). Table 3.1 presents the model parameters employed in this study. This includes experimental data on muscle morphology [65] [64] [3], as well as modeling assumptions, such as the number of GMUs per muscle and the number of MUs per GMU.

Table 3.1: Parameters for CT and TA muscles in the proposed model

Muscle	**TA**	**CT**
GMU per muscle (N)	10	10
MU per GMU (M)	35	44
Fibers per MU	10	20
Percentage of slow fibers	35%	47%
Percentage of fast fibers	65%	53%
Slow fibers time constant (τ_s)	35 ms	
Fast fibers time constant (τ_f)	15 ms	
Maximum firing rate for GMU ($F_{\max}$)	150 Hz	
Standard deviation for ROF ($\boldsymbol{\eta}$)	2 Hz	

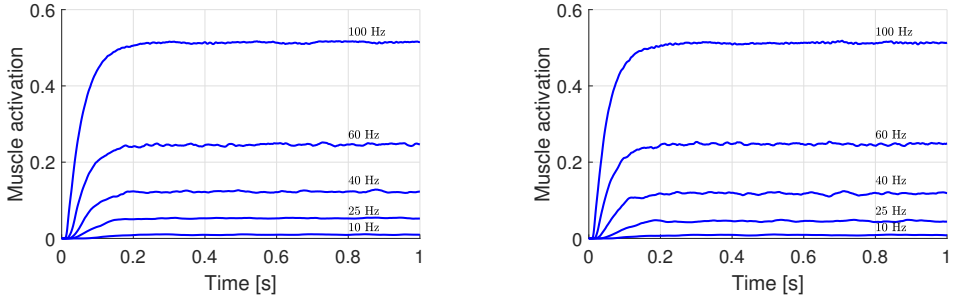
For this study, the body-cover model developed by Titze and Story [10] was employed. This low-dimensional model was chosen due to its simplicity and the physiologically-based relationship established between model parameters and muscle activation established in [15]. Glottal aerodynamics were modeled following [91], and no vocal tract was included to facilitate comparisons with results presented in [15].

Chapter 4

RESULTS

4.1 Muscle Activation Model Description

Figures 4.1a and 4.1b show examples of CT and TA muscle activation signals obtained using the proposed model for the same set of firing rates shown in Figure 3.2. At the lowest firing rate, only the first 2 GMUs are nominally recruited, while for the remaining firing rates all 10 GMUs may be active. In all cases shown, none of the GMUs are tetanized. The time series shown in the figure have transient portions that last for approximately 0.2 s, which represents the time required for the muscle to transition from the fully relaxed to a contracted state. In comparison with the traditional wave summation model shown in Figure 3.2, we observe that the muscle activation signal generated using the stochastic model lacks periodic structure, thus more closely resembling actual muscle behavior [16].

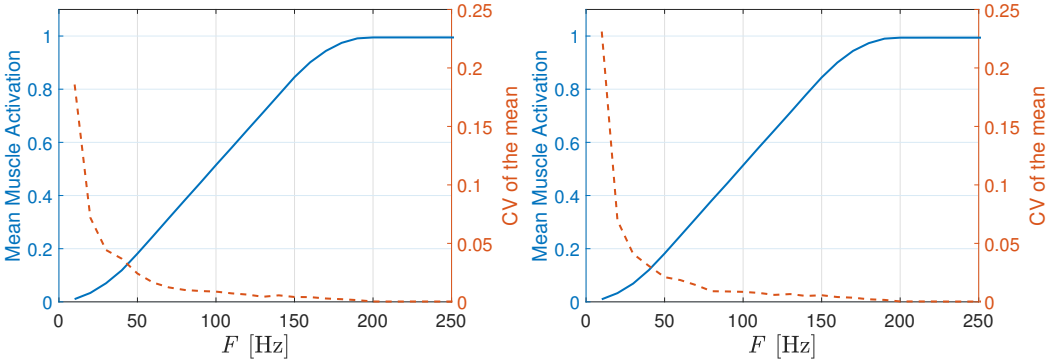


(a) CT muscle activation for a range of firing rates

(b) TA muscle activation for a range of firing rates

Figure 4.1: Example of muscle activation for the CT (left) and TA (right) muscles, firing at 10, 25, 40, 60, 100 Hz for 1 s. The time series have transient portions that last for approximately 0.2 s, which represents the transition from a relaxed to a contracted state.

Note that Figures 4.1a and 4.1b correspond to one realization of the proposed muscle activation scheme. In order to characterize its general behavior, we need to run statistics on many realizations of the signal. Therefore, 40 simulations of the activation signal were computed for each value of the firing rate, which spans from 10 Hz to 250 Hz, in steps of 10 Hz, *i.e.*, 10000 simulations are computed. Note that, $F_{\max}$ is set at 150 Hz, so the tetanization frequency F_{tet} is approximately 200 Hz given the ROF. At this frequency, nominally all 10 GMUs should be firing at $F_{\max}$, barring stochastic variability in the ROF in Equation 3.2.4. Simulations are performed up to 250 Hz to account for the latter.



(a) figure (b) figure

Figure 4.2: Mean activation (left axis, solid line) and CV (right axis, dashed line) versus firing rate, for CT (left figure) and TA (right figure) muscles. In this case, F_{max} is set at 150$[Hz]$, so the tetanization frequency F_{tet} is approximately 200$[Hz]$, due to the ROF. The mean muscle activation saturates at a value of 1 during tetanus, as expected.

The average muscle activation (average of the mean signal values for all realizations) of the 40 signal realizations for each firing rate F for both the TA and CT muscles is shown in Figure 4.2 (blue line). We note that in the range of 40 Hz $\leq F \leq$ 180 Hz the relationship between firing rate and mean activation is linear. Below 40 Hz there are inactive GMUs, while above 180 Hz the effect of saturated GMUs begins to be noticeable. Typical values of muscle activation employed in reduced order models range between approximately 0.1 and 0.5, which falls within the linear region of the mapping, which is amenable to simple control strategies.

The average CV of the 40 signal realizations for each firing rate F for both the TA and CT muscles is shown in Figure 4.3. The average CV is an

estimate of the variability within the activation signals across firing rate. The more considerable variability in low firing rates is a result of Equation 3.2.5, and the different responses between muscles is a product of the morphological construction of the muscles, as shown in Table 3.1. Specifically, the differences between the two muscles are confined to lower firing rates due to the different proportion of slow-small fibers, which changes the properties of the temporal filtering in the muscles. We note that CV is substantially uniform for $F > 180\,\text{Hz}$ as more and more GMUs become tetanized, and thus no longer change behavior with increasing firing rate.

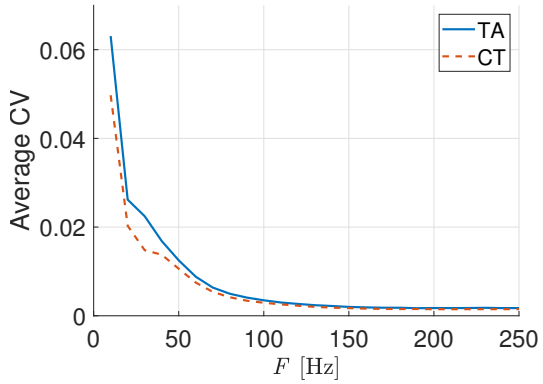


Figure 4.3: CV estimation of $a_m(t)$. For each one of the 40 iterations, CV was computed for $a_m(t)$. For each frequency, the mean (left) and standard deviation (right) is presented considering the 40 measures of CV.

In addition to differences in signal variability between individual realizations, the mean of the signal can also change. That is, each realization may have a different steady state mean muscle activation value due to the stochastic nature of the model. To capture this, Figure 4.2 also presents the coefficient of variation of the mean (standard deviation of the mean values divided by the average muscle activation) for the TA (Figure 4.2b) and CT (Figure 4.2a) muscles at each firing rate (yellow dashed line).

Comparing the CV of the mean in Figure 4.2 with the average CV in Figure 4.3 shows that the variability of the mean is on the order of the variability of an individual realization, which has implications for pitch control. To begin to establish a relationship between the mean activation behavior and pitch control, we posit that mean activation represents a neurological target. Therefore, the standard deviation of the mean is associated with the target variability. The CV of the mean decreases exponentially with the firing rate due mainly

to the increasing mean; the standard deviation of the means remains relatively constant with firing rate, except at very low firing rates (Figure 4.4). Neurologically, this translates into a better pitch control at higher activations.

The behavior of the average CV and the CV of the mean support the idea that variability in discharge rates influences force fluctuations at lower levels of activation. This is consistent with previous findings [52], which report that variability at lower levels is due to low-pass filtering of the neuronal drive. Most of the higher frequency components that are present in the input signal are damped out, leading to low-frequency oscillations manifesting in the muscle activation output. It is uncertain if low-frequency variations are due to ISI variability or low-frequency oscillations in MU discharge [51]; this is particularly true in the specific case of laryngeal muscles, for which information is scarce.

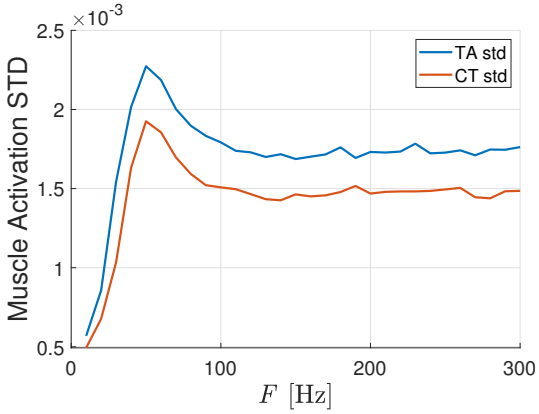


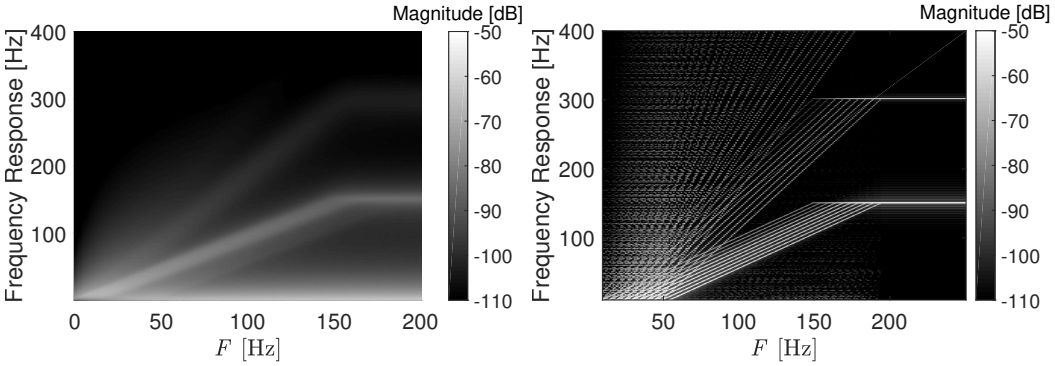
Figure 4.4: Standard deviation of mean activation. The peak produced at 50 [Hz] is produced when all GMUs become active. This estimation can be understood as a measure of relative variability.

Figure 4.4 presents the estimation of the standard deviation for the mean activation at each firing rate. Increasing standard deviation of the mean at lower frequencies is due to the addition of new active GMUs. With each new GMU, more fibers are recruited, and variability in the estimation of the mean also increases. At $50\,Hz$, all GMUs are active, with no new GMUs added for further firing rates. For higher frequencies, standard deviation remains relatively constant. Differences between muscles can be explained by the different fiber composition and the total amount of MU (See Table 3.1)

4.2 Spectral Analysis

To further characterize the properties of the muscle activation model, the spectral content is analyzed as a function of firing rate. The power spectral density (PSD) is computed as the average periodogram of the 40 signal realizations. Figure 4.5 presents the resulting PSD for the TA muscle as a function of firing rate. Color intensity indicates the magnitude of the frequency components. Two conditions are presented, Figure 4.5a presents the model with random components in the interspike interval and ROF, whereas Figure 4.5b shows the results for the deterministic model, without any random variability.

In Figure 4.5a, a strong energy band is present, centered around the firing rate, which has a slope of 1; this saturates at a firing rate of 150 Hz due to tetanization, see Table 3.1. The width of the high energy band is approximately 50 Hz that arises due to the ROF distributing energy between GMUs. Higher harmonics are present due to the quasi-periodic content in the signals.



(a) Spectogram for stochastic activation model.

(b) Spectogram for deterministic activation model.

Figure 4.5: PSD of the stochastic (left) and deterministic (right) muscle activation model versus firing rate for the TA muscle. Color intensity shows the magnitude of the PSD.

Figure 4.5b presents the PSD for the deterministic muscle activation model, which occurs when $\boldsymbol{\eta} = 0$ in Equation 3.2.4 and $\mathrm{CV}_e = 0$ in Equation 3.2.5, is completely deterministic. Herein, the PSD has a similar structure, but with the high energy bands resolved into clear tonal components. The contribution of the different slow and fast GMUs can also be appreciated, having a slight difference in intensity given the different amplitudes of both twitches.

The other salient feature in Figure 4.5a is a low frequency component in

the response below approximately 20 Hz. This arises from a cross-spectral DC component that is inversely proportional to the standard deviation of the random variable that models the ISI in the activation signal [92]. This is directly related to the non-zero CV of the mean observed in Figure 4.2; that is, there is variability in the mean muscle activation that arises as a direct result of the ISI variability. A mathematical deduction for this result is presented next.

Considering equation 3.2.10, then the power spectrum $G_{a_m}(f)$ for a given muscle m with muscle activation a_m is:

$$G_{a_m}(f) = \frac{G_p(f)}{\mathrm{E}\{\sum_{j=1}^{N} \boldsymbol{p}\,(t, F_{\mathrm{tet}})\}} \tag{4.2.1}$$

$$G_p(f) = \sum_{j=1}^{N} G_{GMU_j}(f) \quad . \tag{4.2.2}$$

This shows that the resulting power spectrum is given by the combined effect of the spectra of each GMU, denoted by G_{GMU_j}. These expressions consider all the active motor units. In the same way, the effect of each GMU depends on the effect of an individual slow or fast MU. Considering equation 4.2.2, Tetzlaff *et al* [92] describes the spectrum for a GMU as follows:

$$G_{GMU_j}(f) = \sum_{i=1}^{M} G_{\mathrm{III}_{i,i}}(f) A_i(f) A_i^*(f) + \sum_{i,l,i\neq l}^{M} G_{\mathrm{III}_{i,l}}(f) A_i(f) A_l^*(f) \quad , \tag{4.2.3}$$

where A_i corresponds the power spectrum of the alpha contraction function for the i MU (see equation 3.2.2) and M is the number of MU for each GMU. $G_{\mathrm{III}_{j,j}}$ and $G_{\mathrm{III}_{j,l}}$ are, respectively, the auto-spectra and cross-spectra of the spike trains. Equation 4.2.3 describes the resulting spectrum as the summation of the auto and cross spectra of the spike trains filtered by the transfer function of the alpha twitches. Taking a slow or a fast twitch affects the bandwidth of the transfer function, but for this specific case, the difference is minimal given the values considered. Now, the expression for the auto-spectra can be written as:

$$G_{\mathrm{III}_{j,j}}(f) = \frac{1}{T_j}\left[1 - |Q_j(f)|^2\right] + \frac{1}{T_j^2}|Q_j(f)|^2 \sum_{k=0}^{+\infty} \delta\left(f - k/T_j\right) \quad , \tag{4.2.4}$$

where $Q_j(f)$ is the Fourier transform of the probability density function of the ISI. The cross-spectra is composed of two elements: a continuous spectrum and a line spectrum. The continuous spectrum has values on a specific interval of frequencies (first term), whereas the line spectrum has power concentrated at discrete frequencies (second term). The line spectrum components have energy at multiples of the mean firing rate, with an amplitude proportional to $1/T_j^2$, but the continuous component limits the overall bandwidth of the spectrum. Equation 4.2.4 shows that the bandwidth of the continuous spectra is inversely proportional to the standard deviation of the random variable that models the ISI. Therefore, a spike train with low discharge variability has a greater bandwidth, compared with a spike train with high ISI variability. This means that the bandwidth of this component is proportional to the ISI variability. In the specific case in which the spike train has no variability, being entirely regular, then $|Q_j(f)|$ is equal to 1, and therefore the continuous spectrum is 0. On top of this, equation 4.2.3 shows that the auto-spectra is low-pass filtered by the transfer function of the twitch contraction.

On the other hand, Tetzlaff *et al.* [92] shows that the cross-spectra can be written as follows:

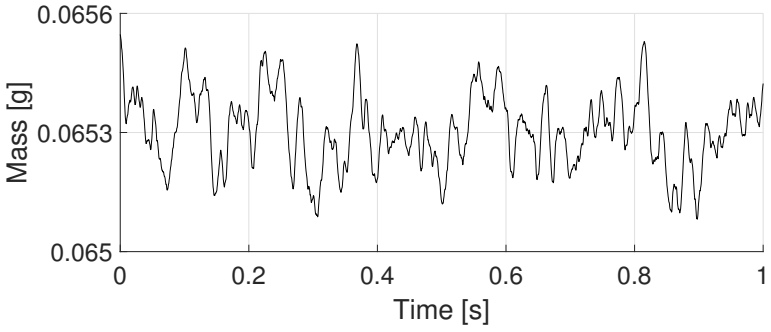
$$G_{\mathrm{III}_{i,l}} = \frac{\delta(f)}{T_i T_l} \quad . \tag{4.2.5}$$

Equation 4.2.5 shows that the total contribution of the cross spectra occurs at DC, and this value increases with the number of spike trains.

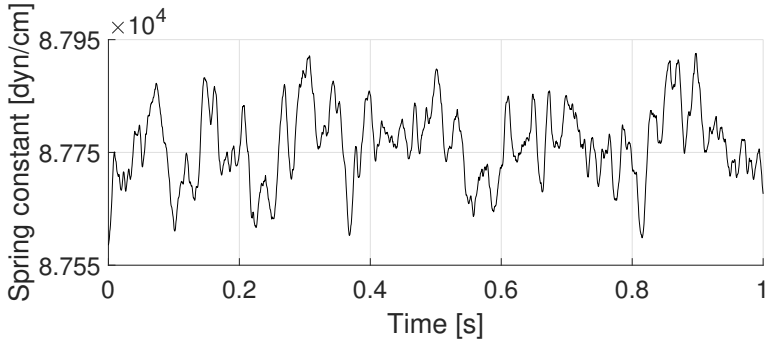
In the spectograms 4.5a and 4.5b all the previously mentioned effects can be appreciated. In the deterministic case (fully regular spike train), there is no low-frequency components, because the continuous spectrum is zero. Therefore, only the DC components and the line spectrum (tonal components) are present. The line spectrum components are presented as harmonics of the firing rate of each GMU, being composed only of deltas, separated each by $5[Hz]$. On the other hand, if there is ISI variability, then the low-frequency components are present in the spectogram. Also, the line spectrum components have a width that depends on the ISI variability. Low-frequency components are essential for the resulting activation profile (or force output), as they have been related to force steadiness at low activation values [52]. In the deterministic case in Figure 4.5b, there is no variability in the ISI; thus the DC component does not appear, and the mean muscle activation parameter is independent of realization, as expected.

4.3 Body-Cover Model Integration

The BCM of the vocal folds (VFs) is typically configured using physiological rules of muscle activation [15], that allow for a meaningful construction of the model parameters. We explore the impact of the developed stochastic muscle activation model by implementing it into the BCM. The BCM model parameters are functions of the *thyroarytenoid*, *cricothyroid*, and *lateral cricoarytenoid* (LC) muscle activations, and as such, the output of the BCM, which is impacted by these parameters, embeds the complex interactions between the various muscles. The proposed stochastic variability incorporated into the muscle activation will thus propagate through the BCM in a non-trivial manner.



(a)



(b)

Figure 4.6: Effect of the proposed stochastic TA muscle activation model on the (a) lower cover mass m_1, and (b) lower cover spring k_1 of the BCM. Both parameters vary in time due to the temporal variability in the TA muscle activation parameter.

Figure 4.6 shows an example of how the proposed stochastic model of muscle activation produces variability in the lower cover layer mass and spring constant in the BCM with time. This specific realization employs a firing

rate for the TA muscle of 70 Hz with the CT and LC activations assumed fixed at 0.2 and 0.5, respectively. Temporal variations in the model parameters in Figure 4.6 arise due to the stochasticity embedded in the TA muscle activation by the proposed model. The mean (standard deviation) are $6.53 \times 10^{-2}\,(9.21 \times 10^{-5})$ g and $8.78 \times 10^{4}\,(67.3)$ dyn/cm for the mass and spring constant, respectively.

To more thoroughly evaluate the impact of the proposed stochastic muscle activation model, we perform a parametric analysis of CT and TA activations. An evenly spaced grid of 20×20 firing rates for CT and TA ranging from 0 to 200 Hz was utilized with 40 simulations performed for each parameter combination. To facilitate comparison with the deterministic muscle activation rules established by Titze [15], we extend his Muscle Activation Plots (MAPs) to include variability in BCM output from our stochastic representation. MAPs allow for an explicit representation of the BCM parameters as functions of firing rate of each muscle.

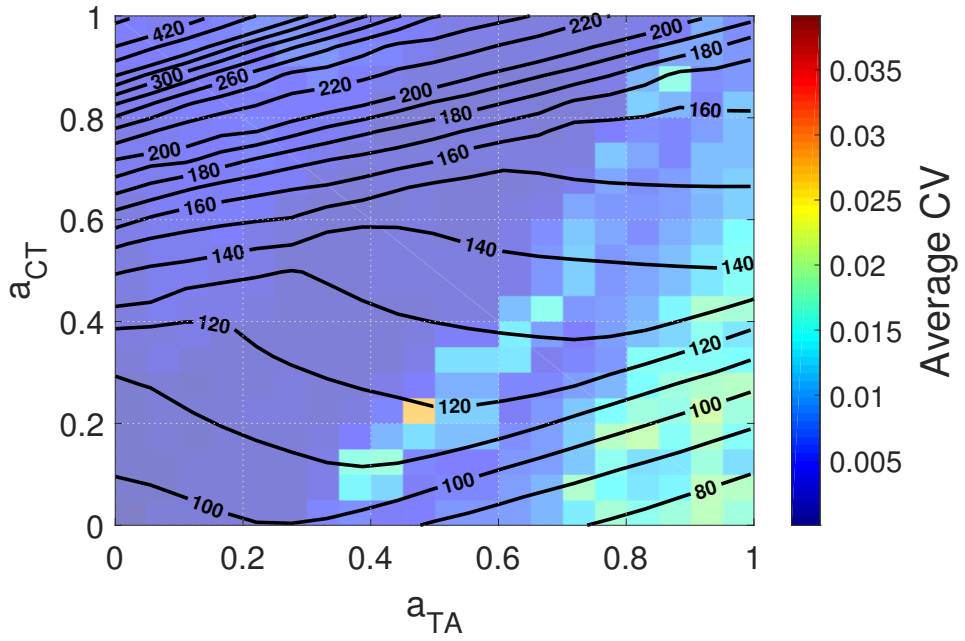


Figure 4.7: Mean fundamental frequency (iso-lines) and average CV (flood contour) for the BCM as functions of TA and CT muscular activations. All units are in Hz.

Figure 4.7 presents a contour MAP of fundamental frequency as a function of CT and TA firing rates. The estimation of f_0 was obtained using the RAPT algorithm [93] on the glottal area waveform. The contour lines in Figure 4.8

display the mean value of f_0, while the flood contour indicates the average CV of the realizations. In this MAP, activations for CT and TA vary from 0 to 1 linearly, which is not valid if we consider a fixed step for the firing rate input vector. Also, inherent stochasticity in the muscular model that affects the activation values prevents a correct visualization of the data. This can be corrected by selecting the firing rate F for CT and TA as X and Y axes. Although this alters the F_0 contour (especially near the edges), the values for the flood contour can be displayed correctly. In the case of Figure 4.7, a rearrangement of the data is performed to show the MAP as Titze initially presented it.

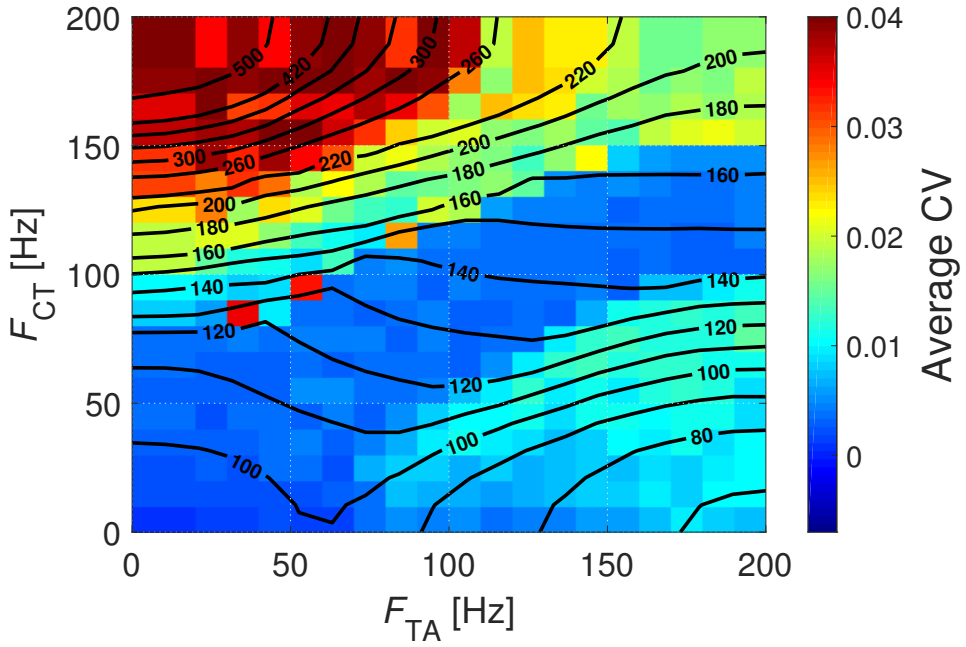


Figure 4.8: Mean fundamental frequency (iso-lines) and average CV (flood contour) for the BCM as functions of TA and CT firing rates. All units are in Hz.

Figure 4.8 shows the contour MAP considering the firing rate of the muscles as parameters. The range of displayed firing rates nominally corresponds to mean muscle activation parameters ranging from 0 to 1, barring the mapping presented in Figure 4.2. With this in mind, we note that the distribution of f_0 with muscle activation parameters displays similarities to the MAP presented by Titze [15]. Specifically, f_0 generally increases with increasing CT and decreasing TA firing rates and vice versa. The CV distribution is somewhat more complex, with the highest variability occurring when F_{CT} is high and

F_{TA} is low. A slight increase is also observed when F_{TA} is high and F_{CT} is low. Interestingly, CV is not elevated when both firing rates are high. Thus, there is not a direct relationship between the variability in f_0 and that of a particular muscle. This is in contrast with the results from the simpler model presented in [4] that ascribed all f_0 variation to the TA muscle.

To further investigate the trends observed in Figure 4.8, we explore the details of the important BCM parameters influencing f_0, namely the lower cover spring k_1 and mass m_1. Sample time series for a specific case was previously presented in Figure 4.6. Figure 4.9(a) presents the mean spring stiffness and average CV for the full range of CT and TA firing rates. In general, the mean value of k_1 is a strong function of CT, increasing rapidly as F_{CT} increases. It is a much weaker function of F_{TA}. The opposite is true for m_1, shown in Figure 4.9(b), which increases with F_{TA} while remaining virtually unchanged with F_{CT}. The average CV distribution for k_1 shows relatively higher values for the extremes of F_{CT}, whereas the density for m_1 is highest at low values of F_{TA} and relatively invariant otherwise. The combination of these two average CV distributions largely explains the average CV map for f_0 in Figure 4.8.

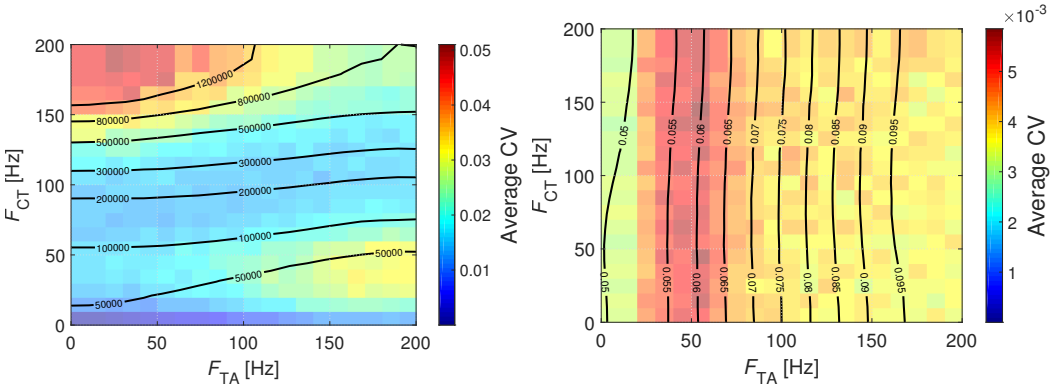


Figure 4.9: Mean BCM parameters (iso-lines) and average CV (flood contour) as functions of TA and CT firing rates. (a) Lower cover spring constant k_1 (in dyn/cm); and (b) lower cover mass m_1 (in g) of the BCM.

Figure 4.10 presents the MAPs for the remaining parameters of the BCM. With this figures, a similar analysis can be performed. The behavior of the lower mass presented in figure 4.9b is similar to the upper and body masses in Figures 4.10a and 4.10b, remaining almost unaltered with changes in F_{CT}. On the other hand, density for m_2 is highest for high values of F_{TA}, while

m_b presents the opposite behavior. The remaining coupling constants k_2, k_c and k_b in Figures 4.10c, 4.10d and 4.10e respectively present all a very non-homogeneous behaviour, increasing or decreasing their density on different levels of F_{CT} and F_{TA}.

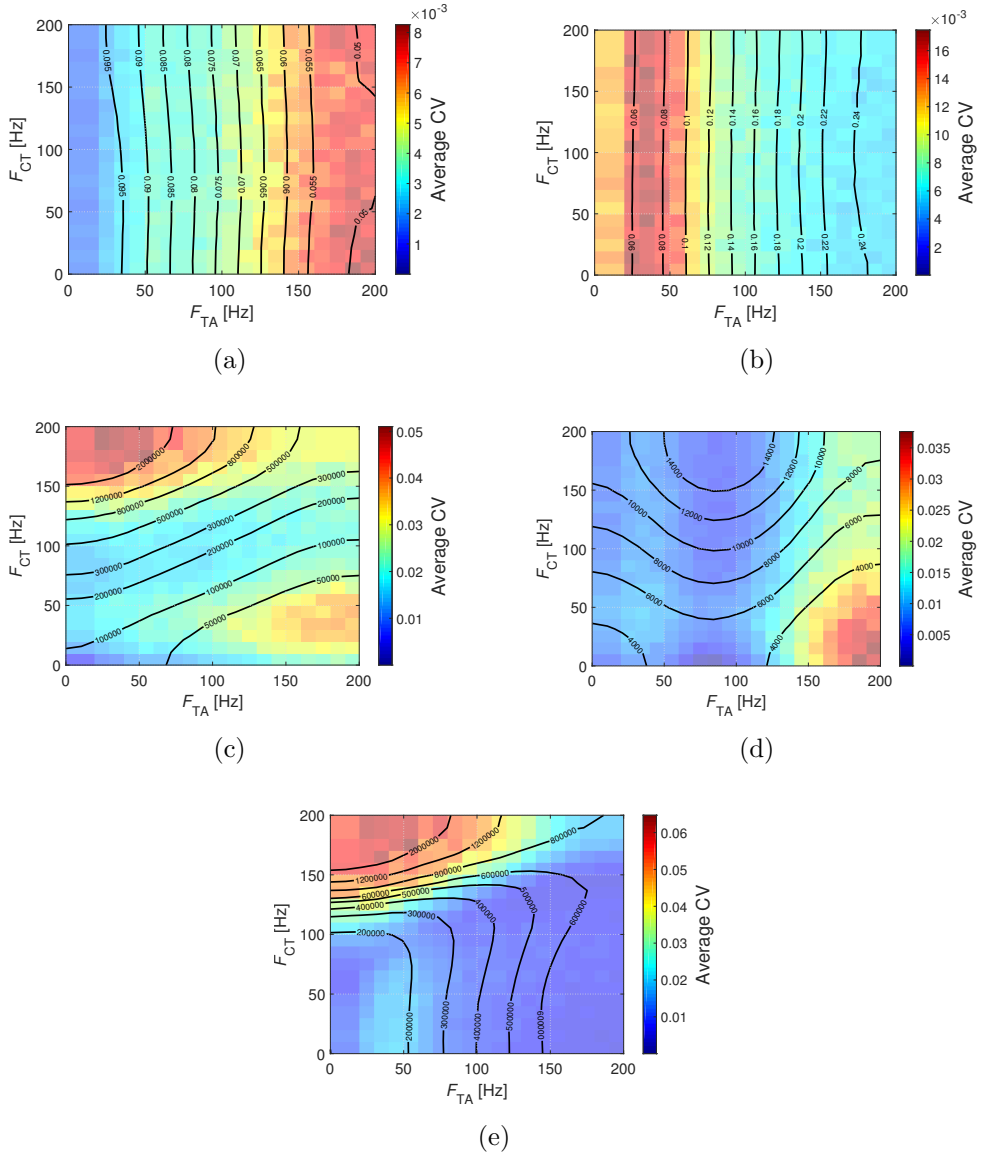


Figure 4.10: Mean BCM parameters (iso-lines) and average CV (flood contour) as functions of TA and CT firing rates. (a) upper cover mass m_2 and (b) body mass m_b (in g); (c) upper cover spring constant k_2, (d) body spring constant k_c and (e) cover coupling spring constant k_b (in dyn/cm).

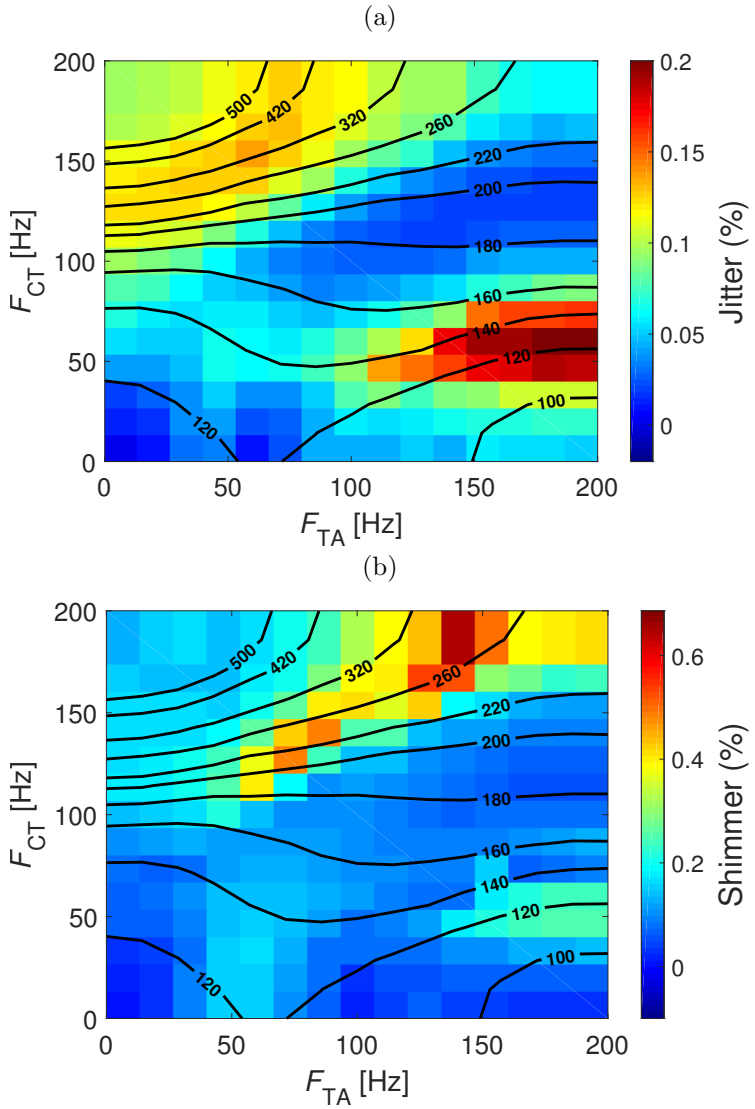


Figure 4.11: Mean f_0 iso-lines overlaid with mean (a) jitter, and (b) shimmer flood contours as functions of TA and CT firing rates.

To place the previous discussions regarding CV in a more clinical context, we now present the described variabilities in terms of jitter and shimmer [32]. While we acknowledge that these measures are not the most clinically relevant, they do enable comparisons with prior studies on the impact of muscle activation variability [4]. Furthermore, these parameters have been widely reported in healthy human subjects [88], thus providing a measure for indirect validation of our proposed stochastic muscle activation model. Jitter and shim-

mer were computed from the simulations previously described using PRAAT scripts [94].

Figure 4.11 presents jitter and shimmer as functions of CT and TA firing rates. The mean fundamental frequency iso-lines are included for reference. For jitter, presented in Figure 4.11(a), there are cases yielding jitter above 1%, which is typical for a normal voice [88]. Therefore, it can be inferred that although the model introduces neuronal variability, this variability is within the "normal" range. Similar observations and conclusions can be drawn from the plot of shimmer in Figure 4.11(b), though these simulations do not include a vocal tract, which can influence this measure.

Figure 4.11(a) displays another interesting result; considering a specific f_0 contour, 120 Hz for example, it can be seen that different combinations of firing rates produce the same f_0, but exhibit varying levels of jitter. If the objective is to phonate at a specific target of f_0, then it is clear that depending on the configuration, different values of jitter can be obtained for the same case. We note that the range of jitter values is in agreement with those reported in [4], wherein it was found that jitter increased with increasing TA firing rate. Our results show a similar trend for low values of CT firing rate, but the opposite trend is observed when the CT firing rate increases. This further emphasizes the complex interactions that are present in muscle activation during phonation.

Chapter 5

CONCLUSIONS

5.1 Final Remarks

The present study introduces a neurophysiological muscle activation scheme for intrinsic laryngeal muscles. It is designed to capture the essential characteristics of force control, providing an activation signal for use in numerical models of the vocal folds. The resulting activation is controlled by the neural firing rate of the different motor units, therefore establishing a link between the nervous system and laryngeal muscle control. Synaptic stochasticity present in the neuronal input of the motor unit arises from the temporal and spatial summations that govern superposition of muscle twitches and motor unit recruitment, respectively. As a result, the muscle response has frequency content centered around both the firing rate and its harmonics, as well as a low-frequency DC component. These components influence both the fine structure variability of the signal, as well as the ability to achieve a target mean activation value for pitch control. These components can have a significant impact on vocal fold parameters and associated outputs of voice production models.

The proposed scheme is integrated into a body cover model of the vocal folds to assess the impact of muscle activation variability on overall laryngeal control. Neural firing rate is introduced as the fundamental control parameter for constructing the body cover model via the muscle activation rules; this novel control parameter offers a natural, physiologically-based, framework that governs vocal fold properties. Fluctuations arise in model parameters, such as vocal fold mass and stiffness, which in turn result in measurable perturbations in the model output. Variability in these parameters is not a simple function of one muscle, rather exhibiting complex dependence on multiple laryngeal muscle behaviors. In spite of this complexity, the fluctuations arising from inherent stochasticity in neural behavior do not introduce abnormal vocal fold vibratory patterns. The results of this study are in general agreement with

previous studies relating neural response of a single muscle to fundamental frequency [4]. However, our results highlight the importance of multi-variate laryngeal interactions, that yield more complex behaviors than previously reported.

The jitter map previously presented shows that although there is a tendency to follow this idea, the exact behavior is more complicated. The interactions with the BCM show that perturbations in f_0 depend on which muscle is involved and how these variations occur. The differences in these results may also be explained by the inclusion of a physiological behavior for muscular recruitment. On the other hand, we observe that part of neurological jitter is likely to be a significant portion of the total jitter in normal vocal fold vibration. The investigation carried out by Titze [4] predicts that perturbations in f_0 decrease with increasing mean firing rate.

5.2 Future Work

The proposed model includes several assumptions regarding muscle morphology and functionality, including a linear summation of muscle twitches, the Rule of Five for motor unit recruitment, and the collection of motor units into groups that are simultaneously recruited. Although linear twitch summation is well established in the literature [4,6], a non-linear summation model could be explored. We further note that other muscle recruitment models exist [20] and may warrant future examination. In this regard, the proposed morphology for the GMUs could be further revised. We acknowledge that the selection of the number GMU can have an effect on partial frequency components of the muscle activation signals, although the resulting VF model kinematics remain mostly unaffected. It would be of value to further validate the proposed model with measurements of intrinsic laryngeal muscle activity of human subjects during phonation, though this is quite invasive and requires particular expertise.

Future efforts will be devoted to exploring the neural effects of antagonistic muscles and extending the rules for controlling a triangular body-cover model [95]. A long-term goal of this work is to replicate the neural variability of common muscle-related pathologies like Parkinson's disease. In the case of Parkinson's disease, neurons exhibit an intricate pattern of inhibition and excitation, which leads to altered firing rate patterns [96]. The proposed model could potentially replicate this behavior and therefore serve as a starting

point to construct a physiologically-relevant model of Parkinson voices, which is currently lacking. There are also other applications of the proposed muscle activation scheme that do not involve a model of vocal folds. For instance, a 3D model of the vocal tract, in which the geometry depends on normalized activations [97], or lung muscle activation can benefit from the proposed scheme to incorporate inherent neural fluctuations in the simulations.

REFERENCES

[1] "National cancer institute website: Larynx coronal cut," https://training.seer.cancer.gov/head-neck/anatomy/larynx.html.

[2] "University of liverpool website: Anatomy of larynx," https://www.liverpool.ac.uk/ trh/local$_h$*tml/testsites/larynxcarcinoma/anatomy.htm*.

[3] A. Mårtensson and C. R. Skoglund, "Contraction properties of intrinsic laryngeal muscles," *Acta physiol. scand.*, vol. 60, pp. 318–336, 1964.

[4] I. R. Titze, "A model for neurologic sources of aperiodicity in vocal fold vibration," *J Speech Hear Res*, vol. 34, pp. 460–472, 1991.

[5] A. Roth and M. C. W. van Rossum, "Modeling synapses," *The MIT Press.*, p. 139–160, 2010.

[6] Y. C. Fung, *Biomechanics: mechanical properties of living tissues*. New York: Springer-Verlag, 1993.

[7] B. D. Erath, M. Zañartu, K. C. Stewart, M. W. Plesniak, D. E. Sommer, and S. D. Peterson, "A review of lumped-element models of voiced speech," *Speech Comm.*, vol. 55, pp. 667–690, 2013.

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

[9] M. Zañartu, L. Mongeau, and G. R. Wodicka, "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.*, vol. 121, pp. 1119—-1129, 2007.

[10] I. R. Titze and B. Story, "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.*, vol. 77, no. 2, pp. 257–286, 1995.

[11] P. Birkholz, B. J. Kröger, and C. Neuschaefer-Rube, "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*, 2011.

[12] M. E. Smith, G. S. Berke, B. R. Gerratt, and J. Kreiman, "Laryngeal paralyses: theoretical considerations and effects on laryngeal vibration." *J. Speech Hear. Res.*, vol. 35, pp. 545–554, 1992.

[13] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal-fold model," *J. Acoust. Soc. Am.*, vol. 93, pp. 1874—-1884, 1995.

[14] M. Zañartu, G. E. Galindo, B. D. Erath, S. D. Peterson, G. R. Wodicka, and R. E. Hillman, "Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunction." *J. Acoust. Soc. Am.*, vol. 136, pp. 3262–3271, 2014.

[15] I. R. Titze and B. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation," *J. Acoust. Soc. Am.*, vol. 112, pp. 1064–1076, 2002.

[16] C. T. Moritz, B. K. Barry, M. A. Pascoe, and R. M. Enoka, "Discharge rate variability influences the variation in force fluctuations across the working range of a hand muscle," *J. Neurophysiol.*, vol. 93, pp. 2449–2459, 2005.

[17] E. Henneman, "Functional organization of motoneuron pools: The size principle." *Proc. Int. Union Physiol. Sci*, vol. 12, p. 50, 1977.

[18] L. M. Mendell, "The size principle: a rule describing the recruitment of motoneurons." *J. Neurophysiol.*, vol. 93, pp. 3024–3026, 2005.

[19] E. Henneman, G. Somjen, and D. O. Carpenter, "Functional significance of cell size in spinal motoneurons." *J. Neurophysiol.*, vol. 28, pp. 560–580, 1965.

[20] A. J. Fuglevand, D. A. Winter, and A. E. Patla, "A model for neurologic sources of aperiodicity in vocal fold vibration," *J Neurophysiol.*, vol. 70(6), pp. 2470–2488., 1993.

[21] M. M. Lowery and Z. Erim, "A simulation study to examine the effect of common motoneuron inputs on correlated patterns of motor unit discharge," *J Comput Neurosci.*, vol. 19(2), pp. 107–124, 2005.

[22] K. G. Keenan and F. J. Valero-Cuevas, "Experimentally valid predictions of muscle force and emg in models of motor-unit function are most sensitive to neural properties," *J Neurophysiol.*, vol. 98(3), pp. 1581–1590, 2007.

[23] H. Cao, S. Boudaoud, F. Marin, and C. .Marque, "Surface emg-force modelling for the biceps brachii and its experimental evaluation during isometric isotonic contractions," *Comput Methods Biomech Biomed Engin*, vol. 18(9), pp. 1014–1023, 2014.

[24] T. J. M. Dick, A. A. Biewener, and J. M. Wakeling, "Comparison of human gastrocnemius forces predicted by hill-type muscle models and estimated from ultrasound images," *J. Exp. Biol*, vol. 220, pp. 1643–1653, 2017.

[25] I. R. Titze, *Principles of Voice Production.* Iowa City: IA: The National Center for Voice and Speech, 2000.

[26] I. R. Titze, E. Hunter, , and J. G. Svec, *The Myoelastic Aerodynamic Theory of Phonation.* Iowa City: IA: The National Center for Voice and Speech, 2006.

[27] F. Alipour and M. Karnell, "Aerodynamic and acoustic effects of ventricular gap," *J Voice.*, vol. 28(2), pp. 154–160, 2014.

[28] J. McGlashan, M. A. Thuesen, and C. Sadolin, "Overdrive and edge as refiners of "belting"?: An empirical study qualifying and categorizing "belting" based on audio perception, laryngostroboscopic imaging, acoustics, ltas, and egg." *J Voice.*, vol. 31(3), pp. 385.e11–385.e22., 2017.

[29] C. Eckers, D. Hutz, M. Kob, P. Murphy, D. Houben, and B. Lehnert, "Voice production in death metal singers." in *Proceedings of the International Conference on Acoustics*, 2009.

[30] M. Hirano, S. Kurita, and T. Nakashima, "The structure of the vocal folds." in *Vocal fold physiology.* Tokyo: University of Tokyo Press, 1981.

[31] M. Hirano, "Structure and vibratory behavior of the vocal folds." in *Dynamic aspects of speech production.* Tokyo: University of Tokyo Press, 1977.

[32] I. R. Titze, *Principles of voice production.* Iowa City: IA: National Center for Voice and Speech, 2000.

[33] A. Yang, M. Stingl, D. A. Berry, J. Lohscheller, D. Voigt, U. Eysholdt, and M. Döllinger, "Computation of physiological human vocal fold parameters by mathematical optimization of a biomechanical model," *J Acoust Soc Am.*, vol. 130(2), pp. 948–964, 2011.

[34] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.*, vol. 97, pp. 1249—-1260, 1995.

[35] T. Koizumi, S. Taniguchi, and S. Hiromitsu, "Two-mass models of the vocal cords for natural sounding voice synthesis," *J. Acoust. Soc. Am.*, vol. 82, pp. 1179—-1192, 1987.

[36] B. H. Story, "An overview of the physiology , physics and modeling of the sound source for vowels," *Acoust. Sci. & Tech.*, vol. 23, pp. 195–206, 2002.

[37] P. Birkholz, B. J. Kroger, and C. Neuscheafer-Rube, "Model-based reproduction of articulatory trajectories for consonant–vowel sequences." *IEEE Trans. Audio Speech Lang. Process.*, vol. 19(5), pp. 1422–1433, 2011.

[38] ——, "Synthesis of breathy, normal, and pressed phonation using a two-mass model with a triangular glottis." in *Proc. Interspeech 2011*, 2011, pp. 2681–2684.

[39] M. Hirano, "Morphological structure of the vocal cord as a vibrator and its variations." *Folia Phoniatr.*, vol. 26, pp. 89–94, 1992.

[40] I. R. Titze, "Parameterization of the glottal area, glottal flow, and vocal fold contact area." *J. Acoust. Soc. Am.*, vol. 75(2), pp. 570–580, 1984.

[41] R. Chan and I. R. Titze, "Dynamic shear modulus of vocal fold tissues and phonosurgical biomaterials," *J. Acoust. Soc. Am.*, vol. 101, p. 3179, 1997.

[42] E. Marieb and K. Hoehn, *Human Anatomy Physiology 10th Edition.* San Francisco: Pearson, 2016.

[43] G. E. Loeb and C. Ghez, "The motor unit and muscle action," in *Principles of neural science.* McGraw-Hill Companies, Inc., 2000.

[44] R. E. Burke and P. Tsairis, "Anatomy and innervation ratios in motor units of cat gastrocnemius," *J. Physiol. Lond.*, vol. 234, pp. 749–765, 1973.

[45] R. F. Mayer and J. L. Young, "Physiological properties of muscle units in humans." *Trans. Am. Neurol. Assoc.*, vol. 104, pp. 193–196, 1973.

[46] H. S. Milner-Brown, R. B. Stein, and R. Yemm, "Two-mass models of the vocal cords for natural sounding voice synthesis," *Journal of Physiology*, vol. 230, pp. 371—-390, 1973.

[47] H. Hatze, "The relative contribution of motor unit recruitment and rate coding to the production of static isometric muscle force," *Biological Cybernetics*, vol. 27, pp. 21–25, 1987.

[48] H. P. Clamann, "Motor unit recruitment and the gradation of muscle force," *Physical Terapy*, vol. 73, pp. 830—-843, 1993.

[49] ——, "Statistical analysis of motor unit firing patterns in a human skeletal muscle." *Biophys J*, vol. 9, pp. 1233–1251, 1969.

[50] P. B. Matthews, "Relationship of firing intervals of human motor units to the trajectory of post-spike after-hyperpolarization and synaptic noise." *J Physiol*, vol. 492, pp. 597–628, 1996.

[51] F. Negro, A. Holobar, and D. Farina, "Fluctuations in isometric muscle force can be described by one linear projection of low-frequency components of motor unit discharge rates." *J Physiol*, vol. 587, pp. 5925–5938, 2009.

[52] J. L. D. ans F. Negro, R. M. Enoka, and D. Farina, "Motor unit recruitment strategies and muscle properties determine the influence of synaptic noise on force steadiness." *J Neurophysiol*, vol. 107, pp. 3357–3369, 2012.

[53] H. S. Rhee, C. A. Lucas, and J. F. Y. Hoh, "Fiber types in rat laryngeal muscles and their transformations after denervation and reinnervation," *J Histochem Cytochem*, vol. 52(5), pp. 581–590, 2004.

[54] I. R. Titze, J. Jiang, and D. G. Drucker, "Preliminaries to the body-cover theory of pitch control," *Journal of Voice*, vol. 1, pp. 314–319, 1988.

[55] M. H. Hast, "Mechanical properties of the cricothyroid muscle." *Laryngoscope*, vol. 78, pp. 537–548, 1966.

[56] D. S. Cooper, M. Shindo, U. Sinha, M. H. Hast, and D. H. Rice, "Dynamic properties of the posterior cricoarytenoid muscle." *Ann Otol Rhinol Laryngol*, vol. 103, pp. 937–944, 1994.

[57] A. L. Perlman and F. Alipour-Haghighi, "Comparative study of the physiological properties of the vocalis and cricothyroid muscles." *Acta Otolaryngol*, vol. 105, pp. 372–378, 1988.

[58] M. H. Hast, "Mechanical properties of the vocal fold muscle." *Pract Otorhinolaryngol*, vol. 29, pp. 53–56, 1967.

[59] H. Hirose, T. Ushijima, T. Kobayashi, and M. Sawashima, "An experimental study of the contraction properties of the laryngeal muscles in the cat." *Ann Otol Rhinol Laryngol.*, vol. 78(2), pp. 297–306, 1969.

[60] C. Hinrichsen and A. Dulhunty, "The contractile properties, histochemistry, ultrastructure and electrophysiology of the cricothyroid and posterior cricoarytenoid muscles in the rat." *J Muscle Res Cell Motil*, vol. 3, pp. 169–190, 1982.

[61] I. A. Mardini, R. J. McCarter, G. D. Neal, M. L. Wiederhold, and C. E. Compton, "Contractile properties of laryngeal muscles in young and old baboons." *Am J Otolaryngol*, vol. 8, pp. 85–90, 1987.

[62] M. H. Hast, "The primate larynx. a comparative physiological study of intrinsic muscles." *Acta Otolaryngol*, vol. 67, pp. 84–92, 1969.

[63] E. C. B. Hall-Crags, "The contraction times and enzyme activity of two rabbit laryngeal muscles." *J Anat.*, vol. 102, pp. 241–255, 1968.

[64] E. Teig, H. A. Dahl, and H. Thorkelsen, "Actomyosin atpase activity of human laryngeal muscles," *Acta. Oto-Laryngologica*, vol. 85, pp. 272–281, 1978.

[65] W. Happak, M. Zrunek, U. Pechmann, and W. Streinzer, "Comparative histochemistry of human and sheep laryngeal muscles," *Acta. Otolaryngol. (Stockh)*, vol. 107, pp. 283–288, 1989.

[66] J. F. Y. Hoh, "Laryngeal muscle fibre types (review)," *Acta Physiol Scand*, vol. 183, pp. 133–149, 2005.

[67] G. Asmussen and F. Wohlrab, "Enzymhistochemische untersuchungen an der kehlkopfmuskulatur erwachsener kaninchen." *Z Mikrosk Anat Forsch*, vol. 86, pp. 383–393, 1972.

[68] L. Edstrom, C. Lindquist, and A. Martensson, "Histochemical properties of intrinsic laryngeal muscles." in *B. Wyke (ed.) Ventilatory and Phonatory Control Systems*, 1974, pp. 392–407.

[69] F. Mascarello and A. Veggetti, "A comparative study of intrinsic laryngeal muscles of ungulates and carnivores." *Basic Appl Histochem*, vol. 23, pp. 103–125, 1979.

[70] T. Yokoyama, S. Nonaka, and S. Mori, "Histochemical properties of intrinsic laryngeal muscles in cats." *J Auton Nerv Syst*, vol. 56, pp. 50–60, 1995.

[71] K. G. Braund, J. E. Streiss, A. E. Marshall, J. R. Mehta, and K. A. Amling, "Morphologic and morphometric studies of the intrinsic laryngeal muscles in clinically normal adult dogs." *Am J Vet Res*, vol. 49, pp. 2105–2110, 1988.

[72] I. Sanders, I. Jacobs, B. L. Wu, and H. F. Biller, "The three bellies of the canine posterior cricoarytenoid muscle: implications for understanding laryngeal function." *Laryngoscope*, vol. 103, pp. 171–177, 1993.

[73] K. M. Pell and J. W. Stanfield, "Mechanical model of skeletal muscle." *Am. J. Phys. Med.*, vol. 51, pp. 23–28, 1972.

[74] A. B. S. J. A. Faulkner and V. A. Kadhiresan., "A simple hill element—nonlinear spring model of muscle contraction biomechanics." *J. Appl. Physiol.*, vol. 70, pp. 803–812, 1991.

[75] L. M. Studer, D. G. Ruegg, and J. P. Gabriel., "A model for steady isometric muscle activation." *Biol. Cybern.*, vol. 80, pp. 339–355, 1999.

[76] A. S. Wexler, J. Ding, and S. A. Binder-Macleod., "A mathematical model that predicts skeletal muscle force." *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 337–348, 1997.

[77] R. D. Woittiez, P. A. Huijing, H. B. K. Boom, and R. H. Rozendal., "A three-dimensional muscle model: a quantified relation between form and function of skeletal muscles." *J. Morphol.*, vol. 182, pp. 95–113, 1999.

[78] R. Raikova, M. Pogrzebna, H. Drzymała, J. Celichowski, and H. Aladjov, "Variability of successive contractions subtracted from unfused tetanus of fast and slow motor units." *J Electromyogr Kinesiol.*, vol. 18(5), pp. 741–751, 2008.

[79] D. Song, G. Raphael, N. Lan, and G. E. Loeb, "Computationally efcient models of neuromuscular recruitment and mechanics," *J Neural Eng*, vol. 5(2), pp. 175–184, 2008.

[80] R. Raikova, H. Aladjov, J. Celichowski, and P. Krutki, "An approach for simulation of the muscle force modeling it by summation of motor unit contraction forces," *Comput Math Methods Med*, 2013, article ID 625427.

[81] J. Z. Liu, R. W. Brown, and G. H. Yue, "A dynamical model of muscle activation, fatigue, and recovery." *Biophys J.*, vol. 82(5), pp. 2344–2359., 2002.

[82] J. R. Potvin and A. J. Fuglevand, "A motor unit-based model of muscle fatigue," *PLoS Comput Biol*, vol. 13(6), 2017.

[83] K. Meijer, H. Grootenboer, H. Koopman, B. van der Linden, and P. Huijing, "A hill type model of rat medial gastrocnemius muscle that accounts for shortening history effects," *J Biomech*, vol. 31, pp. 555–563, 1998.

[84] B. K. Barry, M. A. Pascoe, M. Jesunathadas, and R. M. Enoka, "Rate coding is compressed but variability is unaltered for motor units in a hand muscle of old adults." *J Neurophysiol*, vol. 97, pp. 3206–3218, 2007.

[85] J. Yin and Z. Zhang, "The influence of thyroarytenoid and cricothyroid muscle activation on vocal fold on vocal fold stiffness and eigenfrequencies." *J. Acoust. Soc. Am.*, vol. 133, pp. 2972–2983, 2013.

[86] S. L. Smith and E. J. Hunter, "A viscoelastic laryngeal muscle model with active components," *J. Acoust. Soc. Am.*, vol. 135(4), pp. 2041–2051, 2014.

[87] J. Yin and Z. Zhang, "Laryngeal muscular control of vocal fold posturing: Numerical modeling and experimental validation." *J. Acoust. Soc. Am.*, vol. 140(3), pp. EL280–EL284, 2016.

[88] J. P. Teixeira, C. Oliveira, and C. Lopes, "Vocal acoustic analysis - jitter, shimmer and hnr parameters," *Procedia Technology*, vol. 9, pp. 1112–1122, 2013.

[89] D. Purves, G. J. Augustine, D. Fitzpatrick, and et al, *Neuroscience. 2nd edition.* Sunderland (MA): Sinauer Associates, 2001.

[90] E. P. Widmaier, H. Raff, and K. T. Strang, *Vander, Sherman, Luciano's Human Physiology: The Mechanisms of Body Function 18a ed.* Boston: McGraw-Hill Higher Education, 2004.

[91] I. R. Titze and B. Story, "Regulating glottal airflow in phonation : Application of the maximum power transfer theorem to a low dimensional phonation model," *J. Acoust. Soc. Am.*, vol. 111, p. 67–76., 2002.

[92] T. Tetzlaff, S. Rotter, E. Stark, M. Abeles, A. Aertsen, and M. Diesmann, "Dependence of neuronal correlations on filter characteristics and marginal spike train statistics." *Neural Comput.*, vol. 20(9), pp. 2133–2184., 2008.

[93] D. Talkin, "A robust algorithm for pitch tracking (rapt)," in *Speech Coding & Synthesis*, W. B. Kleijn and K. K. Paliwa, Eds. Elsevier Science B.V., 1995, pp. 497–518.

[94] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [Computer Program], Version 6.0.41," http://www.praat.org/, 2018.

[95] G. E. Galindo, S. D. Peterson, B. D. Erath, C. Castro, R. E. Hillman, and M. Zañartu, "Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 9, p. 2452–2471, 2017.

[96] R. Levy, W. D. Hutchison, A. M. Lozano, and J. O. Dostrovsky, "High-frequency synchronization of neuronal activity in the subthalamic nucleus of parkinsonian patients with limb tremor," *J Neurosci.*, vol. 20, p. 7766–7775, 2000.

[97] S. Dabbaghchian, M. Arnela, O. Engwall, and O. Guasch, "Synthesis of vv utterances from muscle activation to sound with a 3d model," in *Proc. Interspeech 2017*, 2017, pp. 3497–3501.