UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

Department of Electronic Engineering

High-density surface electromyography as a method for estimating intrinsic laryngeal muscle activity

Thesis proposal presented by
Josué Martínez Hernández

As a partial requirement to qualify for the degree of
Ph.D. in Electronic Engineering

Thesis Supervisor
Matías Zañartu

Valparaíso, Chile
Jan 2026

CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): □ Memoria o trabajo de título **X** Tesis de Postgrado

Título del trabajo: High-density surface electromyography as a method for estimating intrinsic laryngeal muscle

Nombre del candidato(a): Josué Daniel Martinez Hernández

Carrera / Grado: Doctorado en Ingeniería Electrónica

Campus: Sede Casa Central **Departamento:** Electrónica

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, **Matías Zañartu**, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

X - El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

□ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por **(marcar una opción)**:

□ 6 meses □ 12 meses □ 2 años □ 3 años □ 5 años □ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

_____

_____

_____

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: _____08-04-2026_____ **Firma:** _____

Estudiante o Candidato(a):

Fecha: _____08-04-2026_____ **Firma:** _____

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

"High-density surface electromyography as a method for estimating intrinsic laryngeal muscle activity," a dissertation prepared by Josué Daniel Martínez Hernández in partial fulfillment of the requirements for the degree, Doctor of Philosophy, has been approved and accepted by the following:

Matías Zañartu
Thesis Advisor

Alejandro Weisten
Chair of the Examining Committee

Date

Thesis Committee:

Dr. Matías Zañartu

Dr. Alejandro Weinstein, Chair

Dr. Pavel Prado

Dr. James Heaton

DEDICATION

I dedicate this work to my partner and family, whose unconditional support and love made it possible for me to achieve this goal. Thank you for always believing that I could.

ACKNOWLEDGMENTS

I would like to express my sincere and deep gratitude to my advisor, Professor Matías Zañartu, whose expertise in acoustic modeling and signal processing was essential during my doctoral studies. His continuous guidance, valuable advice and corrections, and unconditional support were key to the successful completion of this thesis. Beyond his academic contributions, his mentorship significantly influenced my doctoral trajectory at UTFSM, making it intellectually enriching and personally meaningful, helping me grow as a researcher. I am especially grateful for the trust and the opportunity he gave me to continue this research.

I thank the members of my thesis committee: Professor Alejandro Weinstein and Professor James Heaton. I thank Dr. James Heaton for his availability to work on my scientific article and for accepting to be part of this committee, whose clinical knowledge will be highly beneficial for the objectives of this thesis.

I thank my colleagues from the Voice Production Laboratory (VPLAB), Dr. Jesús Parra, Dr. Juan P. Cortés, Dr. Emiro Ibarra, Nicolás Quinteros, and Christian Castro, for all the moments together during the doctorate, the laughter and conversations, and above all for their friendship. I also thank the corrections and discussions we had, which helped to share information and to work together as a VPLAB group.

I also deeply thank the financial support provided by the Doctoral Scholar-

ship granted by UTFSM, the DPP PIIC 036/2023 programs, Basal CIA250006, Regular FONDECYT 1230828 and the P50DC015446 grant, which was supported by the National Institutes of Health (NIH), National Institute on Deafness and Other Communication Disorders.

I also thank the Center for Laryngeal Surgery and Voice Rehabilitation, at the Massachusetts General Hospital (MGH), for their participation in my scientific article and for sharing their data. I especially thank James Heaton, Daryush Mehta, and Robert E. Hillman for their availability, corrections, and guidance in my work, which were enriching in my training.

I would also like to give a deep gratitude to my partner, Andres Valencia, whose support, patience, and love throughout all these years were fundamental in the development of my doctorate. Finally, I thank my family for their constant support and prayers so that I could complete my process, whose company is missed from a distance.

It has been a great experience to have completed the doctorate, so I am completely grateful to everyone.

ABSTRACT

During speech, the activity and coordination of the intrinsic laryngeal muscles are essential for modulation and control, with the cricothyroid (CT) muscle being primarily responsible for regulating the fundamental frequency during phonation. Currently, methods exist for estimating the activity of these muscles; however, their applicability is limited by the invasive nature of intramuscular electromyography (iEMG) and by the spatial superposition inherent to conventional surface recordings. This thesis proposes and evaluates a non-invasive framework for estimating CT muscle activation using high-density surface electromyography (HDsEMG), supported by synchronized intramuscular recordings as a physiological reference. First, an initial experiment was conducted in which a dataset of simultaneous iEMG and acoustic recordings was established in healthy subjects at the Massachusetts General Hospital (MGH). In this experiment, an acoustic-based normalization strategy anchored to the fundamental frequency (f_0) and sound pressure level (SPL) was developed to improve inter-session and inter-task consistency of EMG measurements. These data allowed the creation of a database that served as the gold standard for validating the subsequent results. Subsequently, in a second experiment, 13 patients were recorded at the VPLab laboratory of AC3E, acquiring simultaneous iEMG, HDsEMG, and acoustic signals. The HDsEMG and iEMG signals were decomposed into motor unit discharge trains and validated using a two-source approach with concurrent intramuscular decomposition, enabling the identification of motor units detected by both techniques. The spatial characterization of motor units was performed by estimating the center of gravity and inferring depth based on Gaussian surface fitting, which allowed the separation of superficial and deep activation regions within the CT muscle.For each spatial group, cumulative spike trains (CSTs) were computed and compared with intramuscular EMG envelopes using correlation, coherence, and error metrics. Although the CSTs reliably reproduced temporal trends and shared synaptic input with the iEMG activity, amplitude mismatches were observed, particularly during dynamic phonatory tasks. To address this limitation, a subject-specific neural network was implemented to estimate normalized CT activation by combining CST-derived features with acoustic correlates (f_0 and SPL). This data-driven approach improved amplitude estimation while preserving physiologically meaningful temporal dynamics. Overall, the results demonstrate that HDsEMG, when combined with spatially informed processing, acoustic normalization, and machine learning, enables the non-invasive estimation of an intrinsic laryngeal muscle activation, providing a foundation for future studies in voice production modeling and potential clinical assessment.

Contents

1 Introduction

1.1 Motivation

It has been estimated that up to 15% of the general population has a voice disorder [1], with a higher number for teachers and high voice users. Current methods for monitoring laryngeal muscle function include techniques such as intramuscular electromyography (iEMG), external laryngeal palpation and laryngeal endoscopy [2, 3]. Although a laryngoscopy evaluation allows an identification of pathologies of the larynx, it is not easily found in many institutions or laboratories, forcing clinicians to resort to acoustic and aerodynamic metrics in vocal production to determine vocal disorders. However, the ability of these metrics to distinguish it from other types of voice disorders has not been well examined. Recent efforts have leveraged the physiological relevance of the low-order voice production model to develop neural networks that enable subject-specific models based on muscle activations [4]. However, these results could not be validated.

Electromyography (EMG) is the recording of muscle electrical activity and can be acquired using two recording modalities: iEMG and surface electromyography (sEMG). iEMG is a well-established means to evaluate the neuromuscular activity of both intrinsic and extrinsic laryngeal musculature.[5]. iEMG signal recorded by means of a needle or a fine-wire electrode has been used to investigate the behavior

of muscle activity and is considered the "gold standard" for the evaluation of the electrical activity of a muscle [6].

iEMG is an invasive procedure, uncomfortable for patients, and requires technical skill for both electrode placement and the interpretation of recorded signals. iEMG target confirmation using anticipated activity patterns is not reliable, given the similarity of activity among intrinsic muscles and the substantial intra- and inter-individual variability in how these muscles are used during phonation [7]. In most cases, the duration of the investigation is also limited due to the need for needle or hooked electrodes in clinical or laboratory settings to collect electrophysiological signals [5, 6]. Nevertheless, iEMG is an excellent technique for identifying the morphological characteristics of motor unit signals and has applicability in numerous clinical scenarios [8, 9]. The placement of needle electrodes directly into the target muscles by experienced operators provides confidence that the recorded signal reflects the activity of the muscle under examination. Although highly informative, its widespread use is impractical due to several limitations, such as its limited spatial resolution and the difficulty of normalizing these signals. Currently, EMG normalization methods based on maximum or submaximal values have been used across different muscles of the body [10, 11]; however, the larynx presents a considerable challenge due to the small size of its muscles and the difficulty of isolating their individual activation during phonatory tasks.Therefore, sEMG emerges as a viable non-invasive alternative.

Conventional sEMG offers a non-invasive means of assessing muscle activity with minimal discomfort for the subject. Conventional sEMG within the field of laryngology has been used for evaluation of laryngeal hyperfunction and swallow with inconsistent results[12, 13]. sEMG has a number of limitations that impair signal detection: (1) impedance of the skin-electrode interface, (2) distance between the myoelectric source and surface electrode, (3) lack of specificity due to the superposition of motor unit action potentials, resulting in cross-talk [14].

In addition to intrinsic laryngeal muscles, several superficial and extrinsic neck muscles may contribute to ventral neck sEMG recordings. In particular, the platysma and infrahyoid (strap) muscles, such as the sternohyoid, sternothyroid, and omohyoid, are located superficial or adjacent to the cricothyroid (CT) muscle and can therefore contribute to signal superposition. These muscles are known to influence laryngeal position and phonation biomechanics, and their activation may contaminate surface recordings, complicating the interpretation of intrinsic muscle activity.

Due to these limitations, High-density surface electromyography (HDsEMG) emerges as a promising alternative. HDsEMG uses multiple electrodes distributed over the skin surface to obtain a more detailed representation of sEMG activity [15]. By increasing the number of recording channels and reducing electrode size, this technique improves spatial resolution and enhances the separation of myoelectric sources, enabling the identification of contributions from different muscle

groups and positioning HDsEMG as a potentially valuable methodological tool for future investigations of laryngeal musculature.

1.2 Goals

1.2.1 General Aim

To develop and validate consistent and noninvasive HDsEMG–based methods for characterizing and estimating intrinsic laryngeal muscle activation, at least for the CT muscle, during human phonation and its relationship with acoustic features.

1.2.2 Specific Aims

1. Specific Objective 1 (SA1): To achieve consistent characterization of the relationship between intrinsic laryngeal muscle activation and acoustic features during sustained phonation. This objective involves the recording of iEMG signals from different intrinsic laryngeal muscles of a single patient across multiple sessions, obtained from a pilot database, as well as the simultaneous recording of acoustic signals. Based on these signals, a normalization framework will be developed to generate standardized data that will be used to validate the proposed method.

2. Specific Aim 2 (SA2): To identify intrinsic laryngeal muscle related components in HDsEMG recordings. This aim focuses on applying EMG decomposition methods to extract surface EMG components, both temporal and spatial, that correspond to intrinsic laryngeal muscle activity in subjects without voice disorders, using simultaneous iEMG recordings and the normalization framework from Specific Aim 1 to establish component correspondence and physiological relevance.

3. Specific Aim 3 (SA3): To estimate intrinsic laryngeal muscle activation from HDsEMG recordings and acoustic signals through neural network–based modeling. This aim develops and evaluates predictive models that infer intrinsic laryngeal muscle activation using surface EMG and acoustic features,including fundamental frequency (f_0) and sound pressure level (SPL), without reliance on iEMG, thereby advancing a fully noninvasive approach applicable to broader experimental and clinical contexts.

1.3 Hypotheses

- **Hypothesis 1:** If an acoustic-based normalization framework anchored to f_o and SPL is applied to iEMG recordings across multiple sessions and phonatory tasks, then the resulting normalized muscle activation features will statistically exhibit significantly lower inter-session variability and higher

inter-session consistency than traditional peak-based normalization approaches. This improvement will be reflected by reduced mean absolute error (MAE), root mean squared error (RMSE), and coefficient of variation (CV) in muscle–acoustic mappings, enabling more reliable comparisons of intrinsic laryngeal muscle activity across phonatory conditions.

- **Hypothesis 2**: If HDsEMG signals recorded from the anterior neck surface are decomposed using a method associated with intrinsic laryngeal muscles, specifically for the cricothyroid (CT) muscle, and validated through a two-source approach with simultaneous iEMG, then a rate of agreement (ROA) greater than 25% will be observed between motor unit discharge patterns detected by both techniques, across phonatory tasks and for each subject. Meeting this ROA criterion will indicate that a measurable portion of intrinsic laryngeal muscle activity is effectively captured at the surface, supporting the physiological validity of HDsEMG-based decomposition for assessing intrinsic laryngeal muscle activation in non-invasive recordings.

- **Hypothesis 3**: If HDsEMG signals recorded from the anterior neck surface are processed using an EMG decomposition method associated with intrinsic laryngeal muscles, then the resulting activation estimates, particularly for the cricothyroid (CT) muscle, will reproduce the temporal components observed in iEMG during phonatory tasks for each subject, with correla-

tion and accuracy comparable to those observed in iEMG signals. The performance of this approach will be quantified using RMSE and R values between the HDsEMG decomposition-based estimates and iEMG envelopes. This correspondence will demonstrate that HDsEMG-based decomposition can provide a non-invasive approximation of intrinsic laryngeal muscle activation during voice production.

- **Hypothesis 4**: If subject-specific neural network models are trained to estimate intrinsic laryngeal muscle activation using HDsEMG recordings and acoustic features (f_0 and SPL), then the predicted muscle activation, particularly of the cricothyroid muscle, will reproduce the temporal dynamics and amplitude trends observed in normalized iEMG signals across phonatory tasks for each subject, with a prediction accuracy comparable to that reported by recent subject-specific neural network approaches that rely on a larger set of input parameters. This performance will be quantified using RMSE and MAE values and will demonstrate that accurate estimation of intrinsic laryngeal muscle activation can be achieved using a reduced and non-invasive set of surface-based and acoustic inputs.

1.4 Overview of the proposed methods

The main objective of this thesis was to establish a framework capable of estimating the activity of intrinsic laryngeal muscles from non-invasive surface recordings, supported by synchronized intramuscular data as the physiological reference. The first stage focused on developing an acoustic-based normalization framework using simultaneous iEMG and acoustic recordings. This framework was designed to minimize session-to-session variability and allow intermuscular comparisons across tasks and recording conditions. The normalized signals served as a gold-standard dataset that links EMG activation patterns with voice-related acoustic parameters such as f_0 and SPL.

The second stage involved applying and validating HDsEMG decomposition algorithms to extract the firing activity of motor units from surface recordings. The decomposition performance was assessed using multiple metrics, including the Rate of Agreement and Decomposition Index which quantify accuracy and reliability under the two-source validation method. These analyses demonstrated the capacity of surface EMG to recover physiologically meaningful information that reflects the activation of intrinsic laryngeal muscles.

The third stage expanded the spatial interpretation of decomposition results by estimating the centers of gravity and activation depths of motor units. To achieve this, a Gaussian surface fitting procedure was applied to the HDsEMG

activation maps, from which the depth of each motor unit was estimated based on the principle of potential decay as the distance from the surface increases. This approach provided a physiologically grounded estimation of motor unit depth within the muscle volume. Once spatial coordinates were obtained, 3D clustering was performed to identify regions of activation that could correspond to specific intrinsic muscles. For each identified cluster, a CST was generated to represent the aggregate motor unit activity within that region. The two-source validation method was subsequently used to identify CSTs associated with the CT muscle, resulting in a composite CST signal that approximated CT activation from surface recordings. These findings are particularly relevant in light of previous reports suggesting that intrinsic laryngeal muscles, do not meaningfully contribute to ventral neck sEMG recordings [7]. In this context, the present results support the notion that, at least for the CT muscle, physiologically meaningful activity can be detected and characterized from HDsEMG signals recorded at the anterior neck surface.

Finally, the last stage explored predictive modeling approaches to estimate the activity of the CT muscle from surface EMG data. For this purpose, neural network models were implemented, trained with synchronized HDsEMG, acoustic (f_0, SPL), and intramuscular data. These models represent a transition from descriptive physiological mapping to computational estimation of muscle activation. The proposed methods provide a foundation for future work integrating vocal fold

models and muscle activity data, contributing to the development of non-invasive tools for clinical and experimental voice assessment.

1.5 Scientific contributions

This work represents the first attempt to estimate the intrinsic muscular activation of the larynx using HDsEMG. The integration of HDsEMG decomposition techniques with acoustic-based normalization provides a methodological framework capable of describing the physiological processes underlying laryngeal muscle activity during phonation.

First, this thesis establishes a reference dataset consisting of simultaneous iEMG and acoustic recordings obtained from the intrinsic laryngeal muscles, including the CT,TA, and lateral cricoarytenoid (LCA) muscles. These recordings were used to develop and validate a normalization framework that improves intersession and intermuscular consistency, enabling reliable comparisons across experimental sessions. This dataset also serves as a benchmark for validating future HDsEMG decomposition and estimation methods within the field of laryngeal EMG.

Second, the proposed framework introduces the use of HDsEMG decomposition as a non-invasive alternative to estimate the intrinsic activation of laryngeal muscles. By implementing the *FastICA* algorithm and the two-source validation

method, this work demonstrates the feasibility of extracting motor unit activation behavior and spatial recruitment patterns from the neck surface, as well as estimating the depth of active motor units. This approach enables the visualization of distinct activation clusters within different muscle regions, representing a novel contribution to the characterization of laryngeal muscle activity and bridging the gap between invasive intramuscular methods and clinically applicable surface techniques.

Finally, this thesis lays the groundwork for extending EMG-based estimation through neural network models capable of inferring intrinsic laryngeal activity from surface signals. The implementation of neural networks trained with synchronized acoustic and EMG data demonstrates the potential to estimate muscle activation, particularly that of the CT, from non-invasive recordings. This contribution represents a step toward physiologically meaningful models that integrate electrophysiological, acoustic, and computational domains.

The main findings of this thesis have been disseminated through peer-reviewed journal publications and international conferences, including studies on acoustic-based normalization of laryngeal EMG and HDsEMG decomposition for estimating muscle activation.

1.6 Publications

This thesis is supported by a set of publications in which the candidate is the main author. Related publications made during the research period are also included

1.6.1 Journals Papers

1. **Martínez, J.**, Ibarra, E. J., Parra, J. A., Mehta, D. D., Heaton, J. T., Hillman, R. E., ... & Zañartu, M. (2025). Toward Acoustic-Based Normalization of Laryngeal EMG for Improved Interspeaker Consistency in Muscle-to-Acoustic Mapping. Journal of Voice.

1.6.2 Conference Papers

1. **Martínez, J.**, Valencia, O., & Zañartu, M. (2023, March). High-density electromyography as a method for estimating laryngeal muscle activity: a preliminary study. In 18th International Symposium on Medical Information Processing and Analysis (Vol. 12567, pp. 373-379). SPIE.

1.7 Conferences

1. **J. Martinez** and M. Zañartu, *"Comparative Analysis of Intrinsic Muscle Activation of the Larynx Using the EMG Decomposition Method in Different Phonatory Tasks. A Preliminary Study"*, Proceedings of the Annual Voice Foundation Symposium 2023; Philadelphia, PA

2. **J. Martinez**, L. Rojas and M. Zañartu, *"Investigating the correlation between diverse vocal tasks and intermuscular coherence: a pilot study using high-density electromyography"*, Proceedings of 13th International Conference on Voice Physiology and Biomechanics (ICVPB 2024), Erlangen, Germany, 2024.

3. **J. Martinez**, O. Valencia, and M. Zañartu, *"High-density electromyography as a method for estimating laryngeal muscle activity: a preliminary study."*. Proceedings of 18th International Symposium on Medical Information Processing and Analysis, November, 9-11, 2022. Valparaiso, Chile.

4. **J. Martinez** and M. Zañartu *"Electromiografía de alta densidad como método para estimar la activación muscular de la laringe"*. 78° Congreso Chileno de Otorrinolaringología, November, 9-11, 2022. Viña del Mar, Chile.

5. **J.Martinez** and M. Zañartu, *"Uso de la electromiografía de alta densi-*

dad como método de estimación de la actividad muscular intrínseca de la laringe". I Congreso de la Sociedad Chilena de Fonoaudiología and IV Congreso de Fonoaudiólogos Investigadores, October, 21-22, 2022. Valparaiso, Chile.

2 Background

This research focuses on the use of HDsEMG signals and their decomposition to estimate the activation of intrinsic laryngeal muscles, particularly the CT muscle, whose measurement is especially difficult in clinical settings due to the invasive nature of traditional techniques. In this context, this chapter presents a review of methods that have sought to characterize laryngeal muscle function from EMG recordings, highlighting recent advances in iEMG and HDsEMG, as well as the limitations that remain. This overview aims to justify the selection of the proposed methodology, which is based on experimental recordings and signal decomposition. Finally, the theoretical foundations that support the estimation of muscle activity from EMG signals are described, along with their relationship to the validation of voice production models using machine learning techniques.

2.1 Model and Physiology of Muscle Activation

The generation of EMG signals is intrinsically linked to the physiology of motor unit (MU), which represent the functional building blocks of neuromuscular activation. The central nervous system (CNS) regulates muscle activity through the successive recruitment of MUs in skeletal muscles, including the intrinsic muscles of the larynx. Each MU is composed of a single α-motoneuron and the muscle fibers it innervates, and its electrical manifestation is observed as a motor unit

action potential (MUAP). A MUAP corresponds to the algebraic sum of the action potentials generated in all fibers belonging to a given MU, and its occurrence leads to muscle contraction, the magnitude of which depends on the MU firing rate. From this perspective, the activation of intrinsic laryngeal muscles during phonation can be described as the coordinated recruitment and modulation of multiple MUs. To interpret this process, several models of muscle activation have been developed, ranging from phenomenological descriptions of muscle force production [16, 17, 18, 19], neurophysiological approaches [20] and recent extensions applied to voice production [21]. These models provide the theoretical basis for linking neural drive, motor unit activity, and measurable EMG signals with muscle force and, ultimately, with vocal function.

2.1.1 MU twitch modeling

The work of Burke et al. [22] provided one of the earliest classifications of MU based on their contractile properties. In this study, the MU population was divided into three distinct physiological types: slow-twitch (S), fast-twitch resistant to fatigue (FR), and fast-twitch fatigable (FF). Each of these categories was associated with specific functional roles, such as sustained contractions, locomotor activity, and rapid force generation, respectively. This classification became a cornerstone for later physiological and computational descriptions of muscle behavior.

Early attempts to characterize the mechanical response of individual MUs were carried out by Thomas et al.[23] and later extended by Negro and Orizio[24], who used spike-triggered averaging techniques to describe the twitch response evoked by single motoneuron discharges [8, 24]. These studies showed that each twitch exhibits contraction and relaxation phases whose duration varies with MU type (slow, fast-resistant, or fast-fatigable). Building upon these early contributions, subsequent modeling efforts incorporated these physiological observations to reproduce tetanic contractions through the temporal summation of twitches driven by motoneuron firing patterns.

Initial mathematical descriptions represented the twitch as the impulse response of a critically damped second-order system. Milner-Brown et al. [25] proposed one of the first analytical formulations, which was later incorporated into the widely known Fuglevand model of a MU pool. More flexible approaches, including four- and six-parameter models (Raikova et al. [17, 26]), were developed to better fit experimental twitch shapes across different MU types. These models highlighted the continuous variability of MU properties, rather than a strict categorical separation.

Figure 2.1 illustrates the main parameters typically included in twitch modeling, such as contraction time, half-relaxation time, and peak force. The mechanical event of a muscle contraction,$f(t)$, for a unit motor i, was expressed

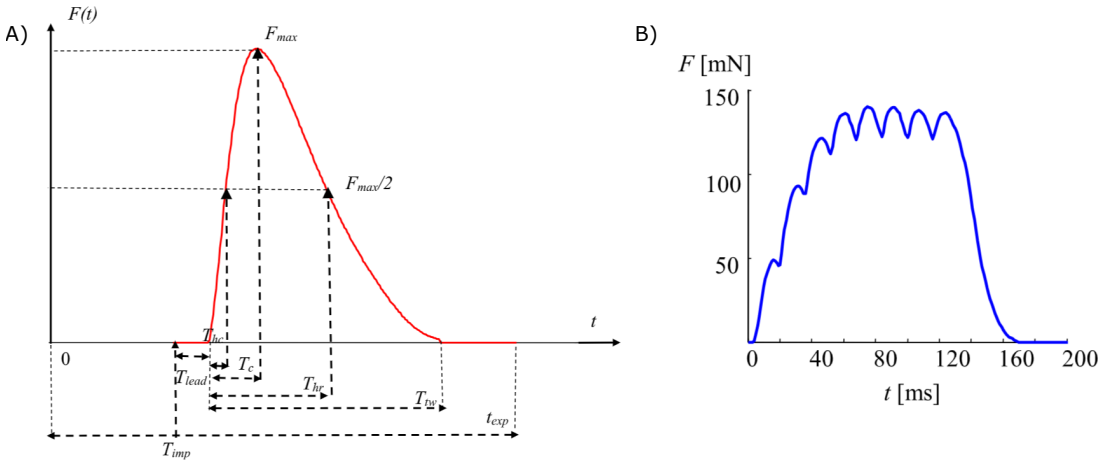


Figure 2.1: (A) Schematic representation of the twitch model parameters as a function of motor unit force over time, $F(t)$. $F_{\max}$ denotes the maximum force; T_{imp} corresponds to the timing of the action potential; T_{lead} represents the lead time; T_c is the contraction time; and T_{hc} indicates the half-contraction time, defined as the interval between the onset of contraction and the point at which the force reaches half of its maximum value. (B) Example of an unfused tetanic contraction produced by repetitive stimulation. Figure from [27].

as:

$$f_i(t) = \frac{P_i \cdot t}{T_i} \cdot e^{1-(t/T_i)} \tag{2.1}$$

where T_i is the rise time (contraction time) to peak force of a muscle contraction and the P_i is related to the peak amplitude of twitch force (Figure 2.1.A), which was expressed as:

$$P(i) = e^{b \cdot i} \tag{2.2}$$

where coefficient b was set as $b = (\ln RP)/n$, RP was range of twitch forces and n

was the number of units in the pool. In any motor unit i and for a train containing k discharges, the force produced in the unit, $F_i(t)$, was equivalent to the sum of individual muscle contractions.

$$F_i(t) = \sum_{j=1}^{k} f_{i,j}\,(t - t_{i,j}) \quad t - t_{i,j} \geq 0 \tag{2.3}$$

$$F_M(t) = \sum_{i=1}^{n} F_i(t) \tag{2.4}$$

Milner-Brown et al. [25] and De Luca and Erim [28] were among the first to describe how the intensity of the EMG signal, as well as muscle force, depends on the number of active motor units and their firing rates. Their studies revealed that the relationship between EMG amplitude and isometric force can take different forms depending on the muscle examined. In small muscles, such as those controlling finger movements, this relationship tends to be linear [25, 28]. Figure 2.1.B illustrates this relationship between excitatory drive and muscle force.

2.1.2 Models of muscle based on MU recruitment

The first attempts to model the recruitment and firing dynamics of motor units were introduced by Fuglevand et al. [16]. Their work established a physiologically grounded computational framework that described how muscle force emerges from the coordinated activation of a motor unit pool organized according to the size principle. In this model, a set of 120 MUs was defined, each with exponentially scaled twitch properties to capture the broad distribution of force capacities across

the pool. Fuglevand also formalized the relationship between motoneuron firing rate and excitatory drive through a simple but powerful expression:

$$FR_i(t) = [E(t) - RTE_i] + MFR, \quad E(t) \geq RTE_i \tag{2.5}$$

where $E(t)$ represents the global excitatory drive reflecting the net synaptic input during voluntary contraction, RTE_i is the recruitment threshold excitation—the minimum level of drive required to initiate repetitive discharge in the i-th motor unit—and MFR denotes the minimum firing rate common to all motoneurons. This formulation allowed the model to reproduce the gradual recruitment of motor units and the corresponding increase in firing rates as excitation intensified. Equation (1) was later reformulated by expressing the instantaneous firing rate as the inverse of the interspike interval (ISI), providing a direct link between simulated discharge times and neural activity patterns:

$$FR_i(t) = \frac{1}{(t_{ij} - t_{ij-1})} \tag{2.6}$$

$$ISI = t_{ij} - t_{ij-1} \tag{2.7}$$

Building on this foundation, subsequent studies expanded the Fuglevand model to include phenomena such as fatigue. Potvin and Fuglevand [20] simulated reductions in twitch amplitude and prolonged contraction times, highlighting how fatigue alters MU firing rates and overall force output. Similarly, Contessa and

De Luca [29] developed a model for specific muscles such as the vastus lateralis and first dorsal interosseous, integrating the "onion skin" firing rate organization. Their approach adopted a four-parameter twitch model and accounted for force potentiation and fatigue effects, providing more physiological realism compared to the original formulation.

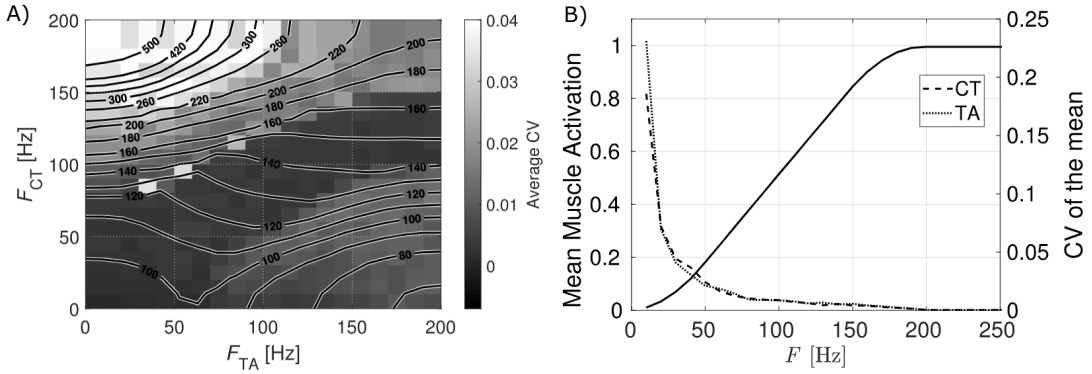


Figure 2.2: (A) Mean muscle activation (left axis, solid lines for both muscles) and coefficient of variation of the mean (right axis, dashed lines; see legend) as functions of firing rate for the CT and TA muscles. (B) Mean fundamental frequency (f_0) represented by iso-lines and average CV shown as a filled contour map for the BCM, expressed as functions of TA and CT firing rates. Figure from [21].

In a voice-specific application, Manríquez et al. [21] adapted the Fuglevand framework to the intrinsic muscles of the larynx. Their model preserved the recruitment and firing properties of motor units while introducing stochastic variability into discharge patterns to mimic physiological noise. This stochastic drive was then used to simulate the activation of two key intrinsic muscles, the CT and

TA, whose activity was mapped into a triangular body–cover vocal fold model. The simulations produced motor activation plots (MAPs) that allowed direct comparison of how different combinations of CT and TA activity affect vocal fold dynamics (Figure 2.2).

A significant advance was introduced by Raikova and colleagues [19], who proposed a model grounded in experimental data from the rat medial gastrocnemius. Unlike earlier models that relied on fixed scaling laws, this framework incorporated measured twitch parameters from three MU types (S, FR, FF) using a six-parameter twitch function. The model also included nonlinear twitch summation and recruitment thresholds derived from rheobase currents. This approach provided a more realistic representation of MU heterogeneity and revealed the stabilizing role of slow MUs in smoothing force output during contractions.

2.2 Laryngeal muscle function from iEMG

iEMG has been a fundamental tool for investigating the physiology of the intrinsic laryngeal muscles (ILMs). By inserting fine-wire or needle electrodes directly into the target muscles, iEMG enables direct observation of motor unit activity that cannot be accessed with surface recordings. Over the years, this technique has been employed in both animal models and human subjects, providing insight into the mechanisms of motor unit recruitment, firing behavior,

and muscle coordination during phonation. Despite its unique advantages, iEMG also faces significant methodological challenges, including invasiveness, variability in electrode placement, and difficulties in normalization across subjects. These factors have motivated a range of studies that attempt to characterize ILM activation in controlled conditions, while also exploring the clinical and research applications of iEMG in voice science. The following subsections review key contributions in animal studies, human recordings, normalization strategies, and the broader implications and limitations of iEMG for understanding laryngeal muscle function.

2.2.1 Animal studies of ILM activation

Initial investigations into ILM function were performed in canines, focusing on the role of specific muscles during phonation. Choi et al. [30] used an in vivo dog model to characterize the posterior cricoarytenoid (PCA), traditionally described as the sole vocal fold abductor. Their findings indicated that the PCA also contributes to phonation by modulating glottal width and balancing the opposing forces of adductors and the CT, rather than acting exclusively as a respiratory muscle [30]. Building on this line of work, Cox, Alipour, and Titze [31] provided a detailed geometric characterization of canine and human CT and TA muscles, quantifying fiber orientation, length, and cross-sectional area [31]. This comparative approach highlighted both similarities and species-specific differences, offering

a structural basis for interpreting functional studies in animal models.

Later studies refined the understanding of ILM activation dynamics. Vahabzadeh-Hagh et al. [32] performed in vivo canine experiments to measure onset and offset times of ILM contraction during neuromuscular stimulation, reporting that TA exhibited the fastest response (34 ms), whereas CT activation was considerably slower (100 ms) . These temporal differences underscored the differential roles of ILMs in fine-tuning phonation and vocal fold posture. Their results showed consistent synergy between these muscles even in the absence of external innervation, suggesting intrinsic or network-level coordination in laryngeal motor control .

Chhetri et al. [33] developed an in vivo canine model to examine the independent and combined effects of ILM activation. By applying graded stimulation to the CT, TA, and LCA/IA complexes, they characterized how each muscle modulates vocal fold posture, f_o, and phonation onset pressure. This study showed the strong f_o dependence on CT activity, the dual role of TA in raising or lowering f_o depending on the activation level, and the necessity of LCA/IA input for achieving posterior glottic closure (Figure 2.3). These findings provided a systematic framework linking ILM activation to phonatory outcomes.

In subsequent work, Chhetri and colleagues [34] refined the functional understanding of adductor muscles. They demonstrated that selective stimulation of TA versus LCA produced distinct glottal configurations, with neither muscle alone achieving complete closure. Later, Chhetri and Neubauer [35] reported that

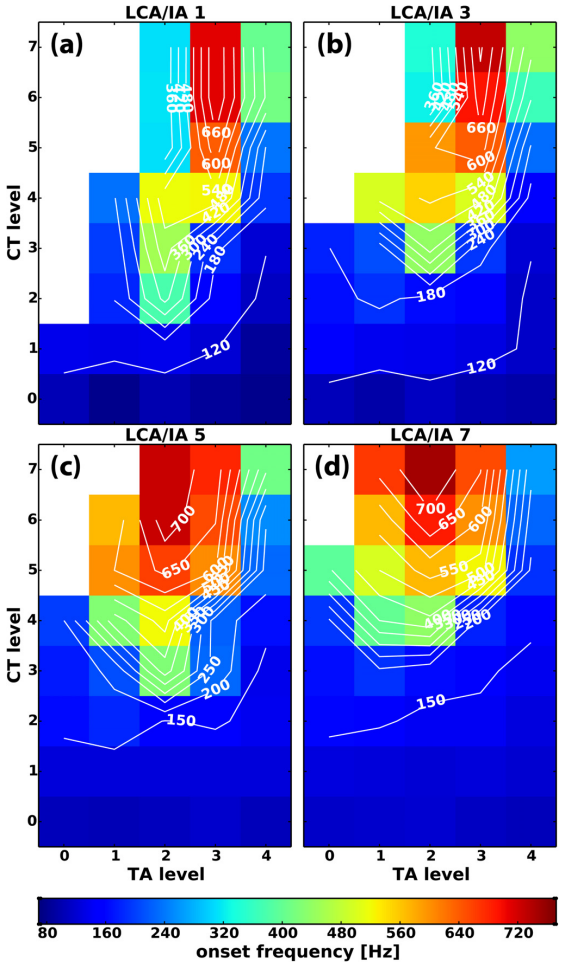


Figure 2.3: Representative maps of muscle activation at phonation onset, illustrating f_0 as a function of graded CT and TA activation for different LCA/IA activation levels: (a) level 1, (b) level 3, (c) level 5, and (d) level 7. Blank regions indicate activation conditions for which phonation onset was not achieved. Figure from [33].

TA activation was associated with abrupt, periodic vibration at phonation onset, whereas LCA stimulation led to irregular vibratory patterns and incomplete closure. These results indicated that TA and LCA have complementary, rather than

redundant, roles in glottal closure and vibratory stability.

Chhetri et al. [36] integrated muscle activation patterns with subglottal pressure to investigate their combined effects on voice production. They showed that different combinations of CT, TA, and adductors could produce similar values of f_0 and SPL, demonstrating the principle of laryngeal motor equivalence. Using MAPs, they identified regions of synergy and redundancy among ILMs, offering a more comprehensive view of neuromuscular control of phonation.

2.2.2 Human intramuscular recordings

iEMG in humans has provided the most direct insights into ILM activation, despite its invasive and technically demanding nature. Early fine-wire recordings by Hirano [37] and Gay et al. [8] established baseline observations on TA and CT activity during pitch changes and glottal tasks. Subsequent refinements demonstrated how differential recruitment of TA, LCA, and CT underpins both phonatory and non-phonatory gestures.

A landmark contribution came from Titze, Luschei, and Hirano [9], who systematically analyzed TA and CT activation in relation to f_o using MAPs. They showed that TA and CT act in a coupled yet context-dependent manner: at low frequencies and intensities, TA activity correlated positively with f_o, whereas at higher frequencies and falsetto-like phonation, increased TA activity could actually lower f_o [9]. This dual role was explained by the body-cover model, where TA

stiffens the vocal fold body but may slacken the cover, producing both positive and negative effects on vibration. Importantly, CT behaves tonically and tracks f_o, while the strap muscles exhibit phasic, range-dependent support—active at very low and very high pitches—with a transition region where CT dominates control.

Hillel [38] extended the field by performing simultaneous eight-channel fine-wire recordings in healthy subjects and patients with laryngeal dystonia. His data provided the first comprehensive normative values for latencies and amplitudes across ILMs during phonation, connected speech, Valsalva, cough, and swallowing (Figure 2.4). A characteristic "set pattern" was described: TA and LCA exhibit a burst at phonation onset followed by reduced activity, while the interarytenoid (IA) sustains adduction. In contrast, CT activation scaled with pitch glides, and posterior PCA recruitment was specific to inspiration. These patterns served as a normative baseline for distinguishing pathological activation in different dystonia subtypes([38]).

Poletto et al. [39] conducted a systematic investigation of the correspondence between intramuscular activity and vocal fold motion in humans. Using hooked-wire EMG from the PCA, TA, LCA, and CT muscles, along with synchronized nasolaryngoscopic imaging, they quantified correlations between EMG amplitude and ipsilateral vocal fold angle across tasks such as sniff, cough, and rapid syllable repetition. Their results demonstrated task- and muscle-specific patterns:

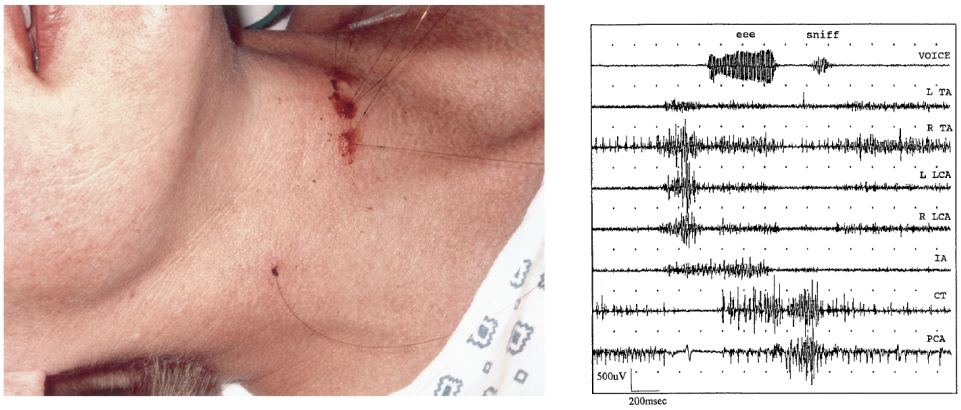


Figure 2.4: Subject instrumented with fine-wire electrodes. The wire exiting laterally from the neck corresponds to the PCA electrode. An eight-channel recording of the "eee/sniff" task is shown, where Channel 1 contains the acoustic signal, and Channels 2 to 8 correspond to EMG recordings from the labeled laryngeal muscles. Figure from [38].

PCA activation correlated positively with vocal fold abduction in all trials, with peak correlations during sniff (typical r $\approx$ 0.74). In contrast, TA and LCA activity showed predominantly negative correlations with abduction, particularly during cough, where they drove forceful adduction (10/14 significant TA trials and 10/12 LCA trials yielded negative r values). These data provided the first quantitative evidence that laryngeal motor control is not based on strict reciprocal activation of agonist–antagonist pairs, but instead often involves task-dependent co-contraction, particularly during speech (Figure 2.5). The findings underscored that CT tends to behave tonically, tracking f_o, while strap muscles provide pha-

sic, range-dependent support—active at very low and very high pitches—with a transition zone where CT dominates.

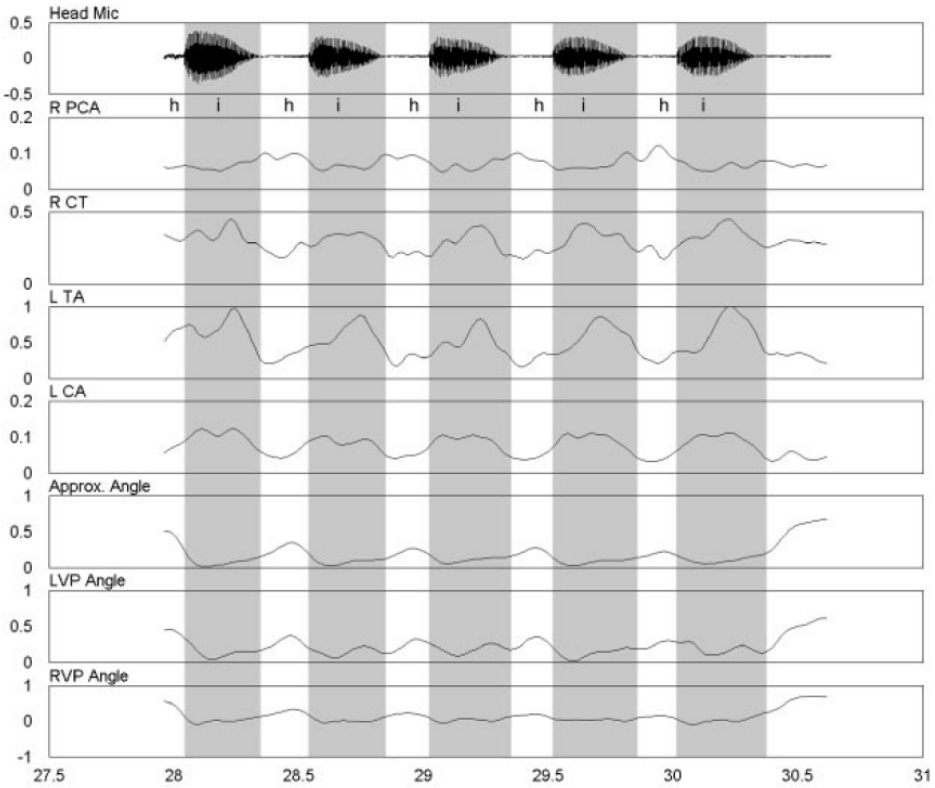


Figure 2.5: Multiple repetitions of the syllable /hi/ are shown. The top panel presents the acoustic waveform together with the corresponding phonetic transcription. Rectified and smoothed EMG recordings are displayed in the subsequent four panels. A clear reciprocal activation pattern between the adductor muscles (LCA and TA) and the abductor muscle (PCA) is observed across all repetitions of /hi/. Figure from [39].

Complementary studies further elaborated on these roles. Roubeau [40] highlighted task-dependent CT activity in high-frequency phonation; Kochis-Jennings

[41] examined LCA activation in dystonia; and Finnegan [42] addressed the interplay of TA and CT during register transitions. However, the invasiveness and variability of electrode placement remain major limitations, restricting widespread clinical use. These methodological challenges motivate the search for non-invasive alternatives such as surface EMG, which aim to capture similar physiological information without the constraints of intramuscular techniques.

Others intramuscular EMG studies of the laryngeal muscles provided foundational insights into the fine motor control of phonation. Seminal works by Hirano and colleagues demonstrated distinct activation patterns of the CT and TA muscles during pitch modulation and register changes, highlighting their complementary roles in vocal fold tension and stiffness control [43, 8]. Subsequent investigations using hooked-wire electrodes further characterized task-specific recruitment and timing differences among intrinsic laryngeal muscles during speech, singing, and voicing transitions [44, 42]. Despite their high temporal and spatial specificity, intramuscular recordings are inherently limited by small sample sizes, inter-subject variability, and sensitivity to electrode placement depth and orientation. These constraints hinder reproducibility across sessions and complicate longitudinal or large-scale clinical studies, reinforcing the need for alternative, less invasive approaches capable of capturing comparable neural drive and activation dynamics.

2.3 Laryngeal muscle function from sEMG

2.3.1 sEMG in the assessment of vocal hyperfunction and MTD

Disorders of the glottis (the area of the vocal folds) are often caused by or accompanied by maladaptive behaviors referred to collectively as vocal hyperfunction. Vocal hyperfunction has been defined as "conditions of abuse and misuse of the vocal mechanism due to excessive and 'imbalanced' muscular forces" [45], characterized by excessive laryngeal and paralaryngeal tension. Despite the widespread use of the vocal hyperfunction designation, diagnosis and assessment in current clinical practice is dependent upon subjective interpretation of patient history and physical examination. Commonly associated symptoms of vocal hyperfunction are not limited to the larynx. Many muscles in the neck that attach to the larynx and/or hyoid bone have voice and speech-related contractions.

Previous studies have attempted to use sEMG to objectively quantify neck muscle tension. Redenbaugh and Reich [46] measured mean neck sEMG of a single anterior neck electrode in seven individuals with healthy normal voice and seven "hyper-functional" individuals, finding that the individuals with disordered voice had significantly greater mean normalized neck sEMG during phonation than individuals with healthy normal voice. Hocevar-Boltezar et al. [47] recorded

sEMG from 18 pairs of differential electrodes on the face and anterior neck in 11 women with disorders associated with vocal hyperfunction (nodules, muscle tension dysphonia) with respect to five women with healthy normal voice. Although this study found significant differences between the mean sEMG, sEMG signals were not normalized.

Stepp [12] explored intermuscular coherence in the beta band as a possible indicator of vocal hyperfunction, a common condition associated with many voice disorders. sEMG was measured from two electrodes on the anterior neck surface of 18 individuals with vocal nodules and 18 individuals with healthy normal voice. Coherence was calculated from sEMG activity gathered while participants produced both read and spontaneous speech. There was no significant effect of speech type on average coherence (Figure 2.6). Their results suggest that bilateral EMG–EMG beta coherence in neck strap muscle during speech production shows promise as an indicator of vocal hyperfunction.

Muscle tension dysphonia (MTD) has been defined as a functional dysphonia in which there is an excessive tension in the (para)laryngeal musculatures [48]. Although different dimensions of MTD have been understood through comprehensive evaluations [49], the diagnostic entity of MTD has not been fully understood. Excessive tension of the (para)laryngeal muscles has been considered as a clinical hallmark of MTD [49]. The assessment of laryngeal muscles to explore muscle tension symptoms is an essential component in the clinical assessment and

diagnosis of MTD [2].

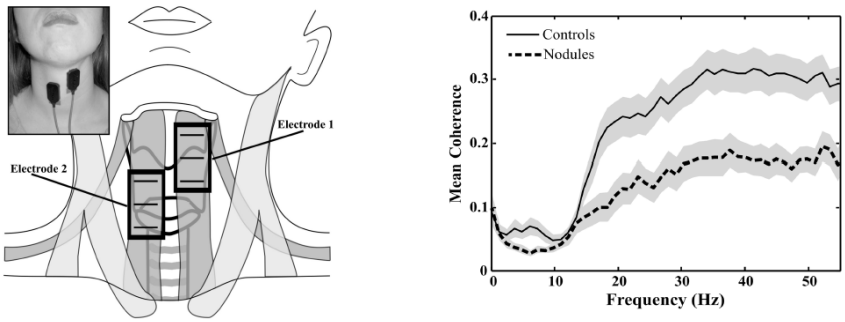


Figure 2.6: Left: schematic representation of the anterior neck showing the placement of double differential sEMG electrodes over the sternohyoid muscle. Right: mean coherence spectra for the two participant groups. The solid black line corresponds to healthy control subjects, whereas the dashed black line represents individuals with vocal nodules. Figure from [12].

Previous investigations have used sEMG to investigate the possible value of sEMG in the assessment of patients with MTD. Redenbaugh and Reich [46] and Hocevar-Boltezar et al, [47] are the first to study (para)laryngeal muscle tension in MTD using sEMG. The authors claimed that the EMG levels were higher in MTD and that sEMG can differentiate patients with MTD from the healthy ones. However, investigations such as the one by Stepp et al, [50] and Van Houtte et al, [13] did not provide enough evidence on the discriminative validity of sEMG in patients with MTD. Unlike previous investigations, a recent study by Balata et al,[51] demonstrated that the electrical activity of the suprahyoid and infrahyoid

muscles was lower in dysphonic subjects compared with non-dysphonic subjects.

2.3.2 Advances in HDsEMG for laryngeal assessment

sEMG within the field of laryngology has been used for evaluation of laryngeal hyperfunction with inconsistent results [13, 52]. sEMG has a number of limitations that impair signal detection such as impedance of the skin-electrode interface and distance between the myoelectric source and surface electrode. HDsEMG has the potential to compensate for the previously described spatial selectivity limitations by application of a large number of electrodes within defined area. A multielectrode array that spans the anterior neck can ensure a high number of electrodes concurrently capture the signal of interest during phonation. Additionally, gross visualization of activity can be intuitively displayed as power-density energy maps [53, 54]. Task-specific muscle activation is used to highlight differences between adjacent musculature.

Bracken et al, [55] proposed to differentiate the activity of the CT muscle at rest, and to identify different patterns of muscle activity between high and low pitch phonation.Ten adults performed phonatory tasks while a 20-channel array recorded spatio-temporal data of the anterior neck. EMG maps were created to evaluate use in gross identification of muscle location (Figure 2.7). Comparison of each electrode to itself across phonatory tasks yielded differences in all subjects during rest versus low versus high. Review of EMG maps allowed for gross

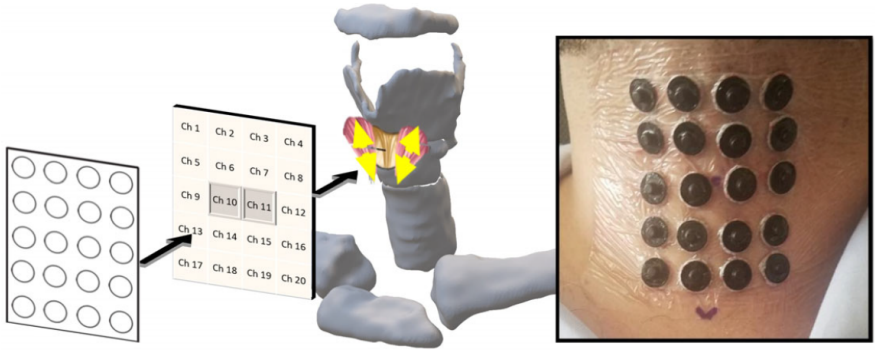


Figure 2.7: High-density sEMG electrode array placed on the anterior neck surface. The central marker located between electrodes 10 and 11 denotes the cricothyroid space, while the inverted V marker below the array indicates the sternal notch. Figure from [55].

identification of CT muscle amidst anterior strap musculature. Zhu et al, [56] recorded 80 channels of HDsEMG from four healthy subjects when they uttered the vowel /a/ at different pitches. They used the root mean square (RMS) of the HDsEMG signals to construct dynamic EMG maps. Their results showed that muscle activities increased steadily with increasing tone levels. And the RMS maps represented the energy distribution of the front neck muscles, which would provide both the temporal and spatial information in accordance with the physiological and biomechanical principles of phonation.

Later, in [57] they calculated Normalized RMS (NRMS) to quantitatively display the correlation of the neck muscles between the left and right sides. The results of sEMG waveforms and NRMS values showed obvious symmetry between

the left and right sides of the front neck muscles. Their research suggests that HD sEMG could be a potential tool to visualize the distribution of muscle activities and observe the coordination of muscle contractions during phonation.

2.4 Methods for Estimating Intramuscular Activation

Efforts to estimate intramuscular activation have led to a variety of methodological approaches that extend beyond classical EMG decomposition. These approaches differ in their conceptual basis, measurement modality, and computational framework, but all share the common objective of linking surface or multisite recordings with MU behavior and internal muscle activity. Below, we summarize representative developments that illustrate this evolution.

One early line of work focused on the spatial distribution of MUPs. Stålberg et al, [58] proposed a model in which unipolar and bipolar MUPs were recorded at multiple radial distances from superficial and deep motor units. Their observations demonstrated that both electrode configuration and motor unit depth strongly shape the recorded potentials, thus providing a foundation for reconstructing intramuscular activity from surface recordings.

Building on this physiological basis, finite element (FE) frameworks have been employed to simulate the conduction of motor unit action potentials through

realistic tissue geometries. Nakajima et al. [59] developed a muscle model in which virtual fibers and elements are embedded within a discretized muscle cross-section, while surface electrodes capture the propagated signals. This approach highlights how geometric complexity and conduction paths influence the recorded EMG, and it offers a computational platform to test hypotheses about intramuscular activation. A schematic of this FE approach is shown in Figure 2.8.

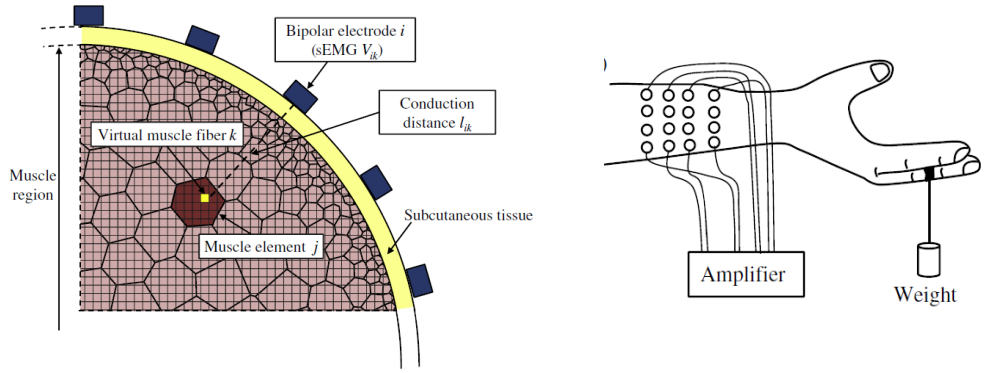


Figure 2.8: Schematic representation of the EMG conduction model used for signal estimation. The virtual muscle fiber k belongs to muscle element j. The parameter l_{ik} denotes the distance between muscle fiber k and bipolar electrode i, while V_{ik} represents the sEMG contribution from fiber k recorded by the bipolar electrode. Figure from [59].

Alternative strategies have combined EMG with imaging modalities. Piovanelli et al. [60] proposed a graph-based analysis for deep muscle activation, where nodes represent recording sites and edges encode functional connectivity derived from EMG features. This method emphasizes network-level coordination and allows identification of active regions even in muscles that are difficult to

access with invasive recordings. Such approaches bridge electrophysiological and systems-level perspectives on muscle control.

In a similar effort to improve spatial resolution, recent methods have integrated HDsEMG decomposition with ultrafast ultrasound imaging [61]. By convolving spike trains of motor unit firing patterns with local tissue velocity waveforms, researchers have been able to localize regions of contraction and track MU-specific activation patterns in real time.

Finally, methods for real-time estimation of active regions have been proposed, combining decomposition outputs with adaptive filtering and spatio-temporal analysis [62]. These systems aim to provide online identification of MU activity, supporting both clinical applications and closed-loop control systems. Collectively, these approaches highlight the progressive shift from simplified models of MUP generation toward multimodal, computationally enriched frameworks that better approximate intramuscular activation.

2.5 Voice production model

Numerical models of voice production have been developed with varying levels of physiological detail, mathematical complexity, and computational cost [63, 64, 65]. Among these, lumped-element approaches have proven particularly useful for simulating vocal fold dynamics, synthesizing natural-sounding voices, and

representing pathological conditions [66, 67, 68, 69, 70]. Typically formulated as non-linear mass–spring–damper systems, these models incorporate tissue biomechanics, mucosal wave propagation, aeroacoustic interactions, and laryngeal muscle control [71, 72, 73, 74]. In this thesis, two such models were adopted: a symmetric three-mass body–cover model with posterior glottal opening [70], chosen for its simplicity and compatibility with Bayesian inference, and the Triangular Body–Cover Model (TBCM), which integrates coordinated activation of the five intrinsic laryngeal muscles and offers a more physiologically representative framework [75].

2.5.1 Active stress of muscles

The generation of force in intrinsic laryngeal muscles arises from the active stress developed during contraction of individual fibers. This macroscopic effect has been quantified experimentally, for instance in canine thyroarytenoid muscles, where measurements distinguish between passive stress (no fibers contracted) and maximal active stress (all fibers contracted) (Alipour-Haghighi & Titze, 1985)[76]. Passive stress reflects the elastic response of the tissue, while active stress captures the additional load supported by contractile elements. Notably, the maximum active stress is approximately 100 kPa, with little dependence on strain within the physiological range of up to 40%.

A key observation is that the total stress in the tissue combines both passive

and active components. While the passive contribution becomes evident even at low strain levels, the active component dominates under voluntary contraction. Nevertheless, the condition of simultaneous recruitment of all fibers, necessary to achieve maximum active stress, rarely occurs during phonation. Instead, natural vocal tasks rely on partial and graded activation patterns, underscoring the role of motor unit recruitment strategies in balancing efficiency and control.

Mathematically, the stress σ can be described as the sum of its passive and active contributions:

$$\sigma = \sigma_{\text{passive}} + \sigma_{\text{active}}. \tag{2.8}$$

Here, σ_{passive} represents the elastic response of connective tissue and non-contractile fibers, while σ_{active} arises from cross-bridge cycling within muscle fibers. To normalize across different experimental conditions, Titze [74] introduced a formulation of active stress that accounts for the level of muscle activation a and the maximum active stress σ_0:

$$\sigma_{\text{active}} = a \cdot \sigma_0. \tag{2.9}$$

This formulation highlights that active stress scales with neural input, ranging from subthreshold activation to near-maximal contraction.

Taken together, these insights emphasize that while active stress values measured in isolated muscle preparations are rarely achieved in vivo, they provide critical boundaries for modeling laryngeal muscle function. Including both pas-

sive and active components in vocal fold models ensures that biomechanical simulations capture the interplay between tissue elasticity and contractile capacity, thereby improving the accuracy of numerical frameworks used to study phonation dynamics.

Accurate passive and active stress values are considered according to the different mechanical characteristics and functional roles of glottic and paraglottic tissues (e.g., muscles, fascia, ligaments, glands, fat, ciliated epithelial mucosa, and non-ciliated mucosa). In addition, the presence of different muscle fiber heavy-chain isoforms is taken into account, as these determine contraction speed and relaxation times, thereby allowing the generation of muscle-specific active stress.

2.5.2 Vocal fold model

The development of lumped-element models of the vocal folds has evolved from the classic two-mass representation by Ishizaka and Flanagan [66] toward more physiologically detailed structures. Story and Titze later introduced the body–cover model (BCM) [67], which added a third mass to represent the layered body–cover configuration of the folds. In this scheme, two cover masses are coupled laterally to a central body mass through nonlinear spring and damping elements, while the body is also attached to a rigid boundary using similar connections. This formulation allowed the body mass, representing the muscular layer, to interact dynamically with the pliable cover layers.

Further refinement came from Zañartu and colleagues [70], who expanded the BCM by incorporating a posterior glottal opening (PGO). This modification enabled systematic analysis of how incomplete posterior closure influences tissue vibration, energy transfer, aerodynamics, and acoustic loading. By accounting for posterior leakage, the model improved physiological realism and provided a better platform for studying pathological and asymmetric phonation patterns.

Building on this foundation, Galindo et al. [73] introduced the TBCM. Unlike the linear configuration of the BCM, the TBCM organizes three interacting mass–spring–damper subsystems in a triangular arrangement that reflects glottal geometry [77]. A key innovation of this model is the inclusion of a zipper-like, progressive closure mechanism, allowing for partial and time-varying collision between the folds. This feature provides a more accurate description of vocal fold impact pressures and glottal closure dynamics during phonation.

The governing equations for both BCM and TBCM are derived from Newtonian mechanics, where each mass is influenced by coupling forces between adjacent elements and external aerodynamic pressures. In the generic formulation:

$$F_u = m_u\,\ddot{x}_u = F_{k,\,u} + F_{d,\,u} - F_{kc} + F_{e,\,u} + F_{Col,\,u}, \tag{2.10a}$$

$$F_l = m_l\,\ddot{x}_l = F_{k,\,l} + F_{d,\,l} + F_{kc} + F_{e,\,l} + F_{Col,\,l}, \tag{2.10b}$$

$$F_b = m_b\,\ddot{x}_b = F_{k,\,b} + F_{d,\,b} - \left[F_{k,\,u} + F_{d,\,u} + F_{k,\,l} + F_{d,\,l}\right], \tag{2.10c}$$

m denotes the mass of the upper (u), lower (l), and body (b) elements, x represents

their medial–lateral displacement, and F denotes the set of forces acting on each block. These include spring (F_k), damping (F_d), aerodynamic (F_e), coupling (F_{kc}), and collision (F_{Col}) components. While the specific definitions differ between models, they follow the analytical formulations detailed in Story and Titze [67] for the BCM and in Galindo et al. [73] for the TBCM.

Finally, both models can be parametrically controlled by intrinsic laryngeal muscle activations following the empirical rules proposed by Titze and Story [71]. Normalized activation levels of the CT (a_{CT}), TA (a_{TA}), and LCA (a_{LCA}) modulate mechanical properties such as stiffness distribution, effective mass, glottal convergence, and fold thickness. This muscle-driven control enables the BCM and TBCM to serve as physiologically interpretable platforms for investigating muscle-to-acoustic mappings.

2.5.3 Triangular body-cover model: TBCM

Alzamendi et al. [75] introduced a multi-physics extension of the TBCM that allows coordinated activation of all five intrinsic laryngeal muscles (CT, TA, LCA, IA, and PCA). This framework builds upon earlier formulations for low-order vocal fold control [71], vocal fold posturing [74], and triangular body–cover dynamics [73]. A schematic of the model is shown in Figure 2.9. By incorporating antagonistic muscle interactions, the model enables simulation of both normal and dysregulated activation patterns, relevant for conditions such as non-phonatory

voice disorders [78].

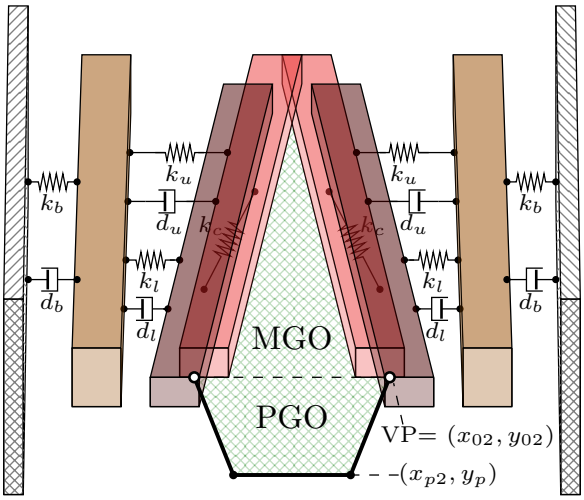


Figure 2.9: Schematic of the triangular body-cover model of the vocal folds. Figure from [75].

In this model, intrinsic muscle forces and passive elements of the ligament (LIG) and mucosa (MUC) generate the internal stresses that deform the vocal folds [79, 34]. Normalized muscle activations a_i ($0 \leq a_i \leq 1$) modulate translational and rotational dynamics around the cricoarytenoid joint center (x_{CAJ}, y_{CAJ}). The displacement (ξ, ψ) and rotation θ of the arytenoid evolve according to:

$$\ddot{\xi} = \frac{1}{M_a}\sum_{i\in\mathcal{I}} \alpha_i F_i - k_x\xi - d_x\dot{\xi} \tag{2.11a}$$

$$\ddot{\psi} = \frac{1}{M_a}\sum_{i\in\mathcal{I}} \beta_i F_i - k_y\psi - d_y\dot{\psi} \tag{2.11b}$$

$$\ddot{\theta} = \frac{1}{I_a}\sum_{i\in\mathcal{I}} \gamma_i F_i - \kappa\theta - \delta\dot{\theta} \tag{2.11c}$$

where F_i represent the forces, and α_i, β_i, and γ_i are the directional cosines for the Cartesian displacements and the directional moment arm, applicable to laryngeal tissue $\mathcal{I}$ = {LCA, IA, PCA, CT, TA, LIG, MUC}. Moreover, the parameters M_a and I_a denote the mass and moment of inertia of the arytenoid cartilage, respectively; k_y, k_x, and κ are the translational and rotational stiffnesses, while d_x, d_y, and δ are the translational and rotational damping coefficients, respectively.

As the vocal process is structurally linked to the arytenoid cartilage, the Cartesian coordinates of the vocal process position (x_{02}, y_{02}) are computed as follows [74]:

$$x_{02} = x_{CAJ} - (x_{CAJ} - \bar{x}_0)\cos\theta + y_{CAJ}\sin\theta + \xi, \tag{2.12a}$$

$$y_{02} = y_{CAJ}(1-\cos\theta) - (x_{CAJ} - \bar{x}_0)\sin\theta + \psi + L_0(\epsilon_r + \epsilon_t). \tag{2.12b}$$

where x_0 represents the cadaveric horizontal position of the vocal process, L_0 denotes the rest length, ϵ_r and ϵ_t correspond to the rotational and translational movements around the cricothyroid joint, respectively.

Based on VF adduction and the symmetry along the midsagittal plane, the

glottal area for the upper (A_u) and lower (A_l) cover masses in this model are calculated as follows:

$$A_u = 2(1-\alpha_u)L_g\big(\tilde{x}_u + 0.5(1+\alpha_u)x_{01}\big), \tag{2.13a}$$

$$A_l = 2(1-\alpha_l)L_g\big(\tilde{x}_l + 0.5(1+\alpha_l)x_{02}\big), \tag{2.13b}$$

where $\tilde{x}_u = x_u - x_{u,0}$ and $\tilde{x}_l = x_l - x_{l,0}$ are block displacements relative to the rest positions, and L_g is VF length. Additionally, α_u and α_l are the proportions of mass length for the upper and lower blocks undergoing collision at the solution time, where $0 \leq \alpha_u, \alpha_l \leq 1$. As a result, the area for the membranous glottal opening is $A_{\text{MGO}} = \min\{A_u,\, A_l\}$. Also, the model simulates the effects of laryngeal posture on the posterior cartilaginous portion of the glottis, assuming that the area of the posterior glottal opening has a trapezoid shape [79]. Thus:

$$A_{\text{PGO}} = \max\{0,\, \min\{(x_{p1} + x_{01}),\, (x_{p2} + x_{02})\}(y_{02} - y_p)\}, \tag{2.14}$$

where x_{p1} is the posterior wall half-width at the bottom, x_{p2} is the posterior wall half-width at the top, and y_p is the posterior wall position along the longitudinal axis. The total glottal area comprises both the membranous and the cartilaginous parts:

$$A_g = A_{\text{MGO}} + A_{\text{PGO}}. \tag{2.15}$$

In addition to controlling intrinsic muscle activation, the model facilitates the adjustment of aerodynamic lung pressure, denoted as P_L. The aerodynamic forces

exerted on the vocal fold cover layer are computed using the resulting subglottal pressure P_s, and supraglottal pressure, P_e, according to [80].

The concept of MAPs was first introduced as a way to visualize how combinations of intrinsic laryngeal muscle activity shape the f_0 during phonation. In particular, Titze and Story [71] developed a low-dimensional vocal fold model in which normalized activation levels of the CT (a_{CT}) and TA (a_{TA}) muscles are systematically varied. This framework provided a mapping between muscle activity space and f_0, represented as isofrequency contour lines. The MAPs highlight regions of redundancy, where different combinations of a_{CT} and a_{TA} can produce similar f_0 outcomes, revealing the flexibility and complexity of neuromuscular control strategies in the larynx.

More recently, Gabriel et al. [75] expanded the use of MAPs by embedding them within the TBCM. This allowed the integration of vocal tract interactions and the generation of MAPs under different subglottal pressure conditions and vowel configurations. The resulting plots (see Figure 2.10) demonstrated how sustained phonatory gestures can be systematically described as trajectories within the muscle activation space, where iso-f_0 contours shift depending on vocal tract shape and aerodynamic load. These findings reinforce the value of MAPs as a unifying framework, linking biomechanical modeling to observed phonatory behaviors, and providing a predictive tool for understanding both normal and disordered voice production.

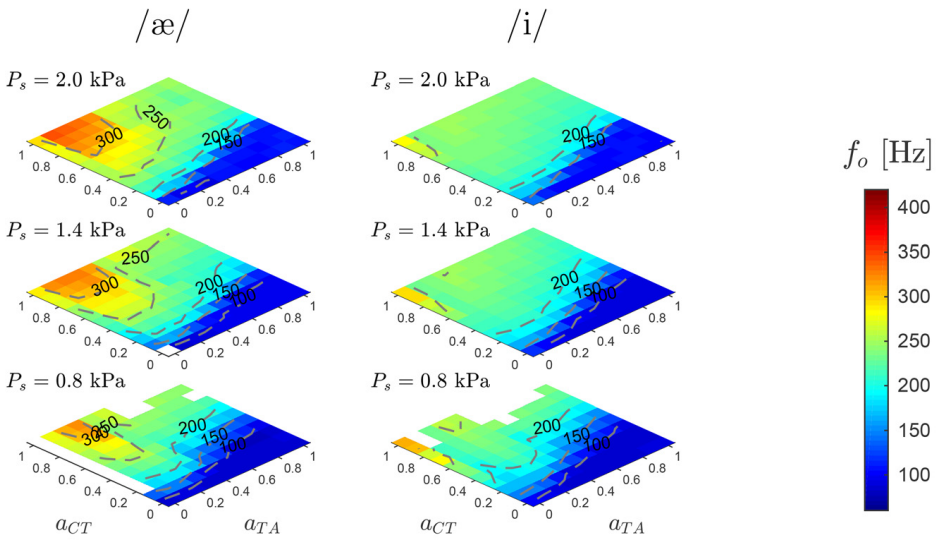


Figure 2.10: Example of MAPs generated with the TBCM, illustrating the relationship between a_{CT}, a_{TA}, and f_0 across different subglottal pressures and vowel configurations. Figure from [75].

An important aspect of the MAPs proposed by Gabriel is the visualization of systematic trends across phonatory conditions. For instance, f_0 contours tend to rise steeply with increasing a_{CT}, particularly at moderate to high levels of subglottal pressure, confirming the dominant role of CT in pitch elevation. In contrast, TA activation shows a nonlinear pattern: at low f_0, greater a_{TA} can increase pitch by stiffening the vocal folds, whereas at higher frequencies, excessive a_{TA} reduces f_0 due to increased vocal fold mass participation. The MAPs also reveal regions of balanced CT–TA coordination where stable f_0 is maintained despite changes in subglottal pressure or vocal tract configuration. These trends illustrate how MAPs can serve as a predictive atlas of neuromuscular strategies

for controlling pitch in both physiological and pathological conditions.

2.6 Chapter Conclusion

This chapter presented a comprehensive review of the experimental and theoretical approaches that have been developed to study laryngeal muscle activation. First, the physiological basis of motor units and their recruitment was discussed, highlighting how classical models, such as those by Fuglevand and later extensions, provide the link between neural drive, firing rate, and force generation. These frameworks are essential for interpreting EMG signals and form the foundation for integrating experimental measurements with computational descriptions of muscle activation. Special attention was given to the CT and TA muscles, given their central role in controlling vocal fold length and tension.

The analysis of laryngeal muscle function using iEMG has been fundamental for understanding motor unit behavior and the coordination of intrinsic muscles during phonation. Studies in both animals and humans have clarified recruitment patterns, firing rates, and their relationship to vocal fold kinematics. Despite its value as a gold standard, iEMG remains invasive, requires specialized expertise, and is limited in duration and subject comfort. These constraints motivated the development of less invasive techniques, while iEMG continues to provide the reference necessary for validating new approaches.

sEMG has offered a more accessible method for evaluating laryngeal muscle activity, particularly in the context of hyperfunction and clinical monitoring. However, traditional electrodes present significant challenges, including cross-talk, low spatial specificity, and filtering effects due to electrode size and placement. To overcome these limitations, HDsEMG emerged as a promising alternative. By using grids of small electrodes, HDsEMG improves spatial resolution and allows decomposition of motor unit activity. Initial applications in the larynx demonstrated its feasibility, although decomposition remains challenging compared to larger limb muscles. Nonetheless, HDsEMG provides a pathway toward non-invasive estimation of motor unit firing patterns.

In parallel, several computational methods have been proposed for estimating intramuscular activity from non-invasive signals. These include inverse modeling frameworks, such as finite element representations of motor unit potentials, spatial decomposition methods, graph-based approaches, and multimodal strategies that incorporate ultrasound. Although heterogeneous, these approaches share the objective of capturing the spatial distribution and recruitment of motor units with improved accuracy. Their application to laryngeal physiology is still emerging, but they illustrate the broader scientific effort to bridge invasive and non-invasive measurements of muscle activation.

Finally, low-order voice production models were reviewed, including the BCM and the TBCM. These models link muscle activation to vocal fold posturing and

dynamics, enabling the generation of MAPs that describe the interplay between CT and TA activity in controlling f_o and glottal closure. More recent developments extended these models to incorporate all five intrinsic laryngeal muscles and posterior glottal openings, enhancing physiological realism. When combined with experimental EMG data, such models provide a powerful framework for interpreting how muscle activity translates into acoustic output. Overall, this chapter establishes the rationale for the use of HDsEMG decomposition and voice models as complementary tools in estimating intrinsic laryngeal activation, which constitutes the central aim of this thesis.

3 Physiological of Muscle Activation and recording intramuscular EMG

This chapter presents an experimental framework that establishes a reference dataset of laryngeal electromyographic activity, serving as the physiological foundation for the subsequent analyses of this thesis. Specifically, it introduces an acoustic-based normalization strategy that provides the gold-standard data required for validating computational models and data-driven methods aimed at estimating ILM activation from surface EMG decomposition. The study focuses on synchronized recordings of iEMG and acoustic signals during controlled phonatory tasks that vary f_o and SPL. These experimental data allow a detailed characterization of the coupling between laryngeal muscle activation and acoustic output, highlighting how ILMs such as the CT, TA, and LCA coordinate to control pitch and intensity. However, the inherent sensitivity of EMG amplitude to electrode placement, impedance, and intersubject differences restricts cross-session or cross-speaker comparisons. To address this limitation, the normalization method calibrates EMG activity relative to acoustic target defined by fixed combinations of f_o and SPL, generating MAPs that enhance the interpretability and consistency of

EMG measurements. The work described in this section was published in Journal of Voice under the title "Toward Acoustic-Based Normalization of Laryngeal EMG for Improved Interspeaker Consistency in Muscle-to-Acoustic Mapping" (Martínez et al., 2025).

3.1 Laboratory recordings and task protocol

The intramuscular recordings were conducted with a 71-year-old male participant who was healthy and had no history of dysphonia or laryngeal abnormalities. The participant, a professional voice user, provided informed consent prior to the study. All experimental procedures were reviewed and approved by the Massachusetts General Hospital Institutional Review Board (Protocol 2008P000652). Before electrode insertion, I dedicated a training session to guide him through the phonatory tasks included in the protocol, ensuring consistent performance during data collection.

The experimental design consisted of four recording sessions plus a prior pilot session, each including three sets of phonatory tasks, as well as non-phonatory and combined tasks. The tasks were selected based on previous literature for their ability to elicit specific activation patterns of the ILM. The first set consisted of sustained vowel productions (/a/ and /i/) of approximately five seconds under different pitch and loudness conditions. These tasks were used to evaluate

Table 3.1: Experimental protocol. The phonatory tasks are categorized into three sets: sustained vowels, pitch glides, and repeated syllable tasks, with variations in pitch and loudness conditions indicated. Non-phonatory and combined tasks are also listed.

Category	**Task**	**Pitch**	**Loudness**
Phonatory Set 1	Sustained vowels /a/, /i/	Low, Habitual, High	Comfortable
		Habitual	Soft, Comfortable, Loud
Phonatory Set 2	Pitch glides on /a/, /i/	Ascending, Descending	Comfortable
Phonatory Set 3	Repeated syllables /pae/, /pi/	Low, Habitual, High	Loud to Soft
Non-Phonatory	Cough, Swallow, Throat clear, Chin tuck		
Combined	Vowel + Valsalva maneuver	Habitual	Very Loud

steady-state activation across various acoustic conditions [37]. The second set included ascending and descending pitch glides on the same vowels at a comfortable loudness level, with the aim of eliciting a broad range of CT muscle activation [40, 41, 81, 8]. The third set involved repeated production of syllables (/pae/, /pi/) at low, habitual, and high pitch levels, performed either at a constant volume or with a gradual decrease from loud to soft, which elicited activation of the TA and LCA muscles [82]. In addition, non-phonatory gestures such as coughing, swallowing, throat clearing, and chin retraction were recorded to capture maximal muscular activation. A combined task was also included, consisting of vowel production during a Valsalva maneuver at habitual pitch and very high loudness, with the purpose of eliciting near-maximal ILM activity for reference normalization [7].

A summary of all task conditions is presented in Table 3.1.

3.1.1 EMG Data Acquisition Protocol

Recordings were obtained from three ILMs: the CT, TA, and LCA muscles. Bipolar hooked-wire electrode pairs were inserted following standard anatomical guidelines [83]. In some sessions, two pairs were inserted into the same muscle to evaluate spatial differences in activation, labeled as TA1–TA2 or LCA1–LCA2. Sterile, disposable hooked wires (44-gauge, enamel-insulated stainless steel with 2 mm of exposed tips) were inserted using a 27-gauge, 30 mm carrier needle through the cricothyroid gap. For the CT, the needle was introduced approximately 5 mm lateral to the midline and angled 50° from the sagittal plane to a depth of 15 mm. For the TA and LCA, insertions were performed at the midline, angled 30° from the sagittal plane, passing laterally beneath the cricothyroid membrane to depths of approximately 20 and 26 mm, respectively [39, 6]. These trajectories allowed for accurate targeting of the muscles while avoiding airway penetration.

Verification gestures were used to confirm electrode placement based on characteristic activation patterns. The CT showed increased activity during high-pitch phonation and minimal activity during low pitch or swallowing, confirming the absence of strap muscle contamination. The TA and LCA demonstrated strong burst activity at phonation onset [38], with the TA exhibiting sustained activation during glottal closure. These signal-based criteria were used to validate

co-localization and signal quality.

Table 3.2: Recorded ILM by session, grouped by VF laterality (left and right). For sessions where the same muscle was recorded with two bipolar electrode pairs, the placements are differentiated by numbers (e.g., TA1 and TA2).

Session	**Left Side Muscles**	**Right Side Muscles**
1	CT	CT, TA, LCA
2	CT, TA1, TA2, LCA	CT, TA1, TA2
3	CT, TA1, TA2, LCA1, LCA2	CT, TA, LCA
4	CT1, CT2, TA, LCA	CT1, CT2, TA

Note: CT = cricothyroid; TA = thyroarytenoid; LCA = lateral cricoarytenoid.

Muscle activity was recorded using an eight-channel EMG system (Bagnoli, Delsys Inc., Natick, MA) equipped with miniature expansion springs to support the fine wires. A 2.5-inch ground electrode was placed on the dorsal neck near the C6–C7 vertebrae. The system band-pass filtered the differential EMG signals between 20–450 Hz, applied a 1000× amplification, and digitized the signals at 20 kHz (Digidata® 1440, Molecular Devices, San Jose, CA). Simultaneous acoustic recordings were obtained using an omnidirectional microphone (Sennheiser MKE104) positioned 15 cm from the lips of the participant. Both EMG and acoustic signals were synchronized and calibrated to physical units: sound pressure in pascals (Pa) for the microphone and voltage (V) for the EMG signals.

3.2 Signal Processing for Acoustic and EMG Features

3.2.1 Acoustic-based Features

Both SPL and f_o were computed from 50 ms analysis windows of the calibrated acoustic signal, with a 25 ms overlap between consecutive frames. SPL was derived from the root mean square value of each windowed segment of the microphone signal, recorded at a distance of 15 cm from the lips [84]. The f_0 was extracted using the `pitch` function in MATLAB, applying the Log-Harmonic Summation (LHS) method [85].

3.2.2 EMG Signal Processing and Electrode Identification

The raw EMG signals were first high-pass filtered using a fourth-order Butterworth filter (10 Hz) [14]. The filtered signals were then full-wave rectified and subsequently low-pass filtered at 8 Hz to obtain the amplitude envelope, as described in [39].The resulting envelope was segmented into 50 ms windows with 25 ms overlap, and the mean EMG amplitude was extracted from each segment. These segments were then synchronized with the corresponding acoustic windows, producing time-aligned EMG–acoustic segments.

To identify electrode pairs likely inserted into the same muscle, cross-correlation was applied to the raw EMG signals under the assumption that high temporal similarity indicates spatial proximity and potential crosstalk between nearby channels [86, 87]. A cross-correlation threshold of 0.5 was used to infer co-localization within the same muscle. Additionally, Pearson correlation coefficients (r) were calculated between all pairs of EMG envelopes. A strong negative correlation with CT activation ($r < –0.5$) was used to help differentiate TA muscle, based on their known antagonistic roles during pitch modulation [81, 43]. For LCA identification, envelope-based correlation was insufficient. Consequently, qualitative inspection of EMG envelopes was performed during tasks with stable adductory demand—specifically, /pae/ descending and /ifi/. Channels exhibiting sustained envelope amplitude with limited modulation across these tasks were considered representative of LCA activity.

3.3 Normalization procedure

3.3.1 Correlation-Based Channel Selection

Channel selection was guided by both cross-correlation of raw EMG signals and Pearson correlation coefficients of EMG envelopes, in accordance with the thresholds defined in the EMG signal processing methodology (Section 3.2.2). The resulting correlation heatmaps are presented in Figure 3.1.In the top row, cross-

correlation matrices computed from raw EMG signals revealed spatial patterns that were consistent across all tasks within each session. Most channel pairs exhibited low correlation ($r > 0.1$), indicating substantial functional independence. However, electrode pairs with correlation values exceeding 0.5 were interpreted as co-localized within the same muscle, reflecting spatial proximity and potential crosstalk [87, 86].

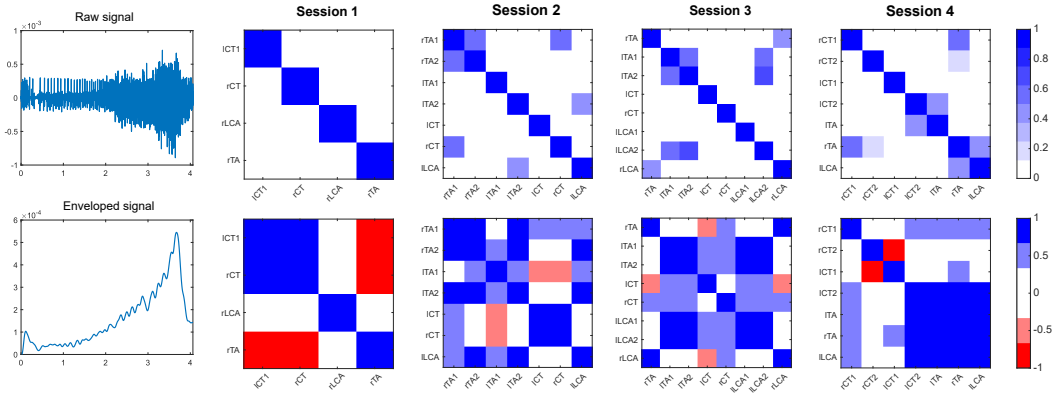


Figure 3.1: Comparative bar plot of mean maximal CT, TA and LCA muscle activation values from EMG signal envelope data. The results of each session per phonatory and non-phonatory task (Valsalva and swallow) are shown.

The bottom row displays r-value matrices derived from EMG envelopes recorded during pitch glide descending tasks across all sessions, capturing co-activation patterns. A consistent strong negative correlation ($r < -0.5$) was observed between CT and TA in some electrode pairs during descending glides, reflecting their well-established biomechanical antagonism in pitch modulation. This behavior was used to distinguish TA from LCA electrodes.

Based on spatial and functional correlation patterns, one representative CT, TA, and LCA channel triplet was selected per session: rCT-rTA-rLCA (Session 1), lCT-lTA1-lLCA (Session 2), lCT-rTA-lLCA2 (Session 3), and lCT1-rCT2-lTA (Session 4).

The normalization procedure used in this study was grounded in the principles of submaximal task-based normalization [88, 89, 10], based on acoustic targets (f_o and SPL). This approach has been proposed as an alternative to MVIC techniques for dynamic and task-specific EMG applications, particularly to reduce intrasubject variability. Given that each acoustic target may be associated with different levels of EMG activation, due to varying combinations of subglottal pressure and co-activation of other ILM, normalization was performed by grouping EMG–acoustic segments into non-overlapping 1 Hz bins for f_o (CT) and 1 dB bins for SPL (TA and LCA). Within each bin, the maximum EMG amplitude was extracted to represent a pattern of peak EMG amplitudes across the acoustic range. The corresponding EMG normalization value was obtained by linearly interpolating between these peak values and the acoustic target. To this end, phonatory tasks that elicited the highest activation levels were used for each muscle.

For the CT muscle, maximum activation was associated with pitch glide tasks from Phonatory Set 2 [40, 81, 8, 43], which typically produce the highest f_o values. In contrast, for the TA and LCA muscles, both involved in glottal adduction and

P_s regulation, maximum activation was associated with high SPL values [37, 42], which are generally achieved during high-intensity productions such as increasing the loudness of /pae/ gestures performed in Phonatory Set 3 [82]. Only phonatory gestures involving vocal fold vibration were considered, as these tasks are directly related to voiced sound production and better reflect the functional role of ILM during speech.

For comparison purposes, a traditional normalization approach was also implemented. In this method, the EMG signal of each muscle was normalized by dividing it by the maximum amplitude observed within the session, regardless of the task. The peak value was selected from the highest activation measured across all available tasks, including both phonatory and non-phonatory conditions, such as swallowing and Valsalva maneuvers, performed during that session [39, 81, 41]. To ensure the reliability of the extracted measures, the raw EMG signals were manually reviewed to confirm that the selected maximum values were not affected by non-EMG events, such as motion artifacts. When such artifacts were identified within a trial, that specific trial was excluded from the normalization process, while retaining the remaining repetitions of the corresponding task. Both approaches were applied in parallel to compare proposed normalization outcomes in downstream analyses, as shown in the Results section.

3.3.2 Muscle–Acoustic Mapping Consistency Across Sessions

To assess the repeatability of muscle–acoustic relationships across sessions, MAPs were generated using normalized EMG activation values and interpolated acoustic features [90, 80]. Each MAP represents a two-dimensional grid of activation for two muscles being compared, with color-coded values indicating either f_o or SPL. Three configurations were analyzed: (1) CT vs. TA with f_o, (2) CT vs. TA with SPL, and (3) LCA vs. CT with f_o. MAPs were computed independently for each session and for both normalization strategies (traditional and proposed). To enable cross-session comparison, a shared activation region was identified within each configuration. Each session was treated as a repeated measure, yielding four values per grid point in the overlapping region. MAPs offer a structured and interpretable framework for evaluating whether normalization strategies preserve consistent and meaningful activation–acoustic relationships across sessions.

To evaluate intrasubject variability, pairwise comparisons were performed across all combinations of the four recording sessions (i.e., S1–S2, S1–S3, ..., S3–S4), resulting in six session pairs. For each pair, acoustic values (f_o or SPL) were extracted from the common MAP region and compared using Pearson correlation (r), MAE, and RMSE. In addition, global consistency across all sessions was as-

sessed using intraclass correlation coefficient (ICC, model 2,1) and coefficient of variation (CoV).

3.4 Results

3.4.1 Acoustic-Based EMG Normalization

Figure 3.2 shows the average peak EMG amplitudes of the CT, TA, and LCA muscles across four recording sessions during a set of phonatory and non-phonatory tasks. Under the conventional normalization scheme, the task eliciting the maximum EMG amplitude was selected as the reference for each muscle. For the CT muscle, this reference corresponded to the pitch glide task (Phonatory Set 2), which consistently produced strong activation across sessions. In contrast, the highest EMG amplitudes for the TA and LCA muscles were observed during non-phonatory gestures, mainly coughing and swallowing.

In contrast, the normalization strategy proposed in this thesis was based exclusively on phonatory tasks. The pitch glide task was selected as the reference condition for the CT muscle, while descending loudness productions of the /pae/ syllable at low and high pitch (Phonatory Set 3) were used as reference tasks for the TA and LCA muscles, respectively. These reference conditions are illustrated in Figure 3.3. The upper panels show scatterplots describing the relationship between EMG activation and acoustic parameters, with CT activation plotted

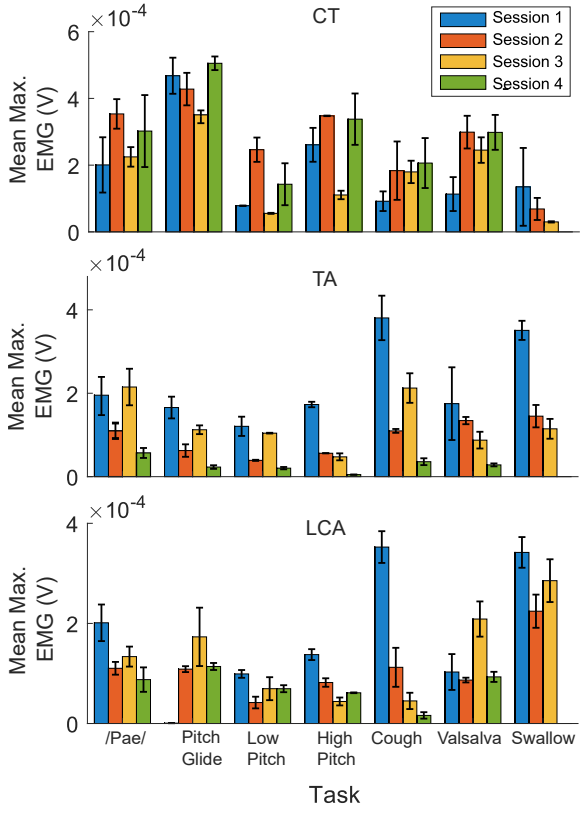


Figure 3.2: Comparative bar plot of mean maximal CT, TA and LCA muscle activation values from EMG signal envelope data. The results of each session per phonatory and non-phonatory task (Valsalva and swallow) are shown.

against f_0 in the first column and TA and LCA activation plotted against SPL in the second and third columns, respectively. The lower panels present discrete peak EMG profiles obtained by grouping EMG–acoustic data into 1 Hz bins for the CT muscle and 1 dB bins for the TA and LCA muscles, as described in the Methods section.

For the CT muscle, EMG amplitude increased systematically with f_0 across

all sessions, following an approximately linear relationship up to 200 Hz. Beyond this range, a transient reduction in activation was observed, potentially associated with a register transition, followed by a renewed increase at higher f_0 values. This behavior suggests a non-linear activation pattern with progressive recruitment at higher frequencies. For the TA muscle, EMG amplitude increased moderately but consistently with SPL across sessions, indicating a stable association between vocal intensity and TA activation. In contrast, the LCA muscle did not show a consistent trend with SPL; instead, intermittent activation peaks appeared at both low and high SPL values, with variable magnitude and occurrence across sessions.

Based on these results, interpolated EMG amplitudes corresponding to subject-specific acoustic reference points were identified for normalization across the four sessions. As indicated by the dashed horizontal lines in the lower panels of Figure 3.3, the acoustic reference targets were set to 350 Hz for the CT muscle and 90 dB SPL for the TA and LCA muscles. These values were selected because they corresponded to submaximal conditions consistently achieved across all sessions and are commonly used as reference maxima in computational voice models. A linear interpolation approach was adopted as a practical approximation, assuming smooth variations in EMG amplitude over small acoustic ranges. Although non-linear relationships may be present, the linear approach was chosen for its simplicity and reproducibility.

Although acoustic targets such as $f_0 = 350$ Hz and SPL = 90 dB may vary across individuals, the task-specific relationships underlying muscle activation are expected to remain comparable due to shared biomechanical constraints. Therefore, submaximal acoustic targets can be adapted to different subject populations while maintaining the robustness of the normalization procedure.

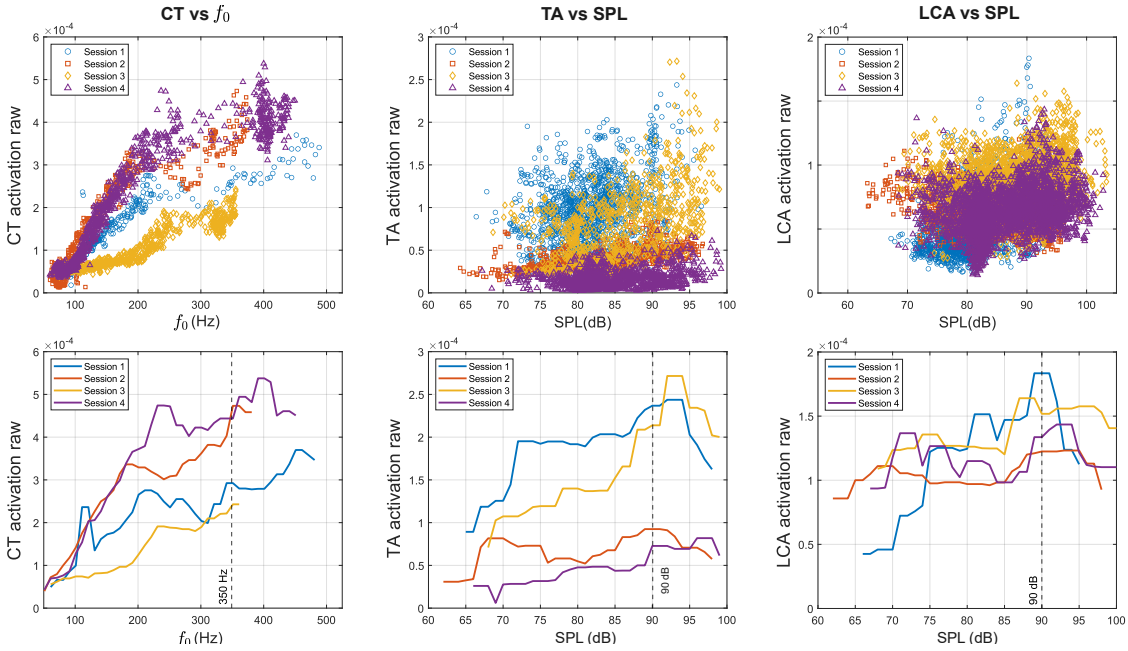


Figure 3.3: (Top row) Scatterplots showing CT activation as a function of f_o, and TA and LCA activation as a function of SPL, obtained from pitch and loudness variation tasks across the four recording sessions. (Bottom row) Maximum activation values per 1 Hz (for f_o) and 1 dB (for SPL) bin. Horizontal lines indicate the interpolated normalization values used for each session: 350 Hz for CT, and 90 dB SPL for TA and LCA.

3.4.2 Normalization Effects on CT and TA with Acoustic Mapping

Figure 3.4 presents MAPs of f_o as a function of normalized CT and TA activation across four recording sessions. The data were derived from tasks involving the vowel /a/—both sustained productions and pitch glides from Phonatory Sets 1 and 2—as well as /pae/ productions with descending loudness from Phonatory Set 3. The MAPs on the left were generated using the traditional normalization approach, whereas those on the right were obtained using the proposed method.

As shown in Figure 3.4b, higher f_o values (250–350 Hz) were primarily distributed in the upper-left quadrant, corresponding to strong CT activation levels (> 0.6) and moderate-to-low TA activation (< 0.5). These results align with previous *in vivo* observations in the canine larynx [35, 33]. Conversely, lower f_o values (50–150 Hz) were concentrated in the lower-left and lower-right quadrants, while the upper-right quadrant—where both CT and TA activation were high—was associated with intermediate f_o values (150–250 Hz). These outcomes are consistent with predictions from computational models of phonation [91, 92, 90]. The upper-right region was reached during conditions demanding higher vocal intensity, particularly during /pae/ descending tasks, where simultaneous activation of CT and TA was observed. In contrast, in the lower portion of the MAPs, increasing TA activation under low CT conditions did not substantially affect f_o.

Despite minor inter-session variations, such as slightly increased TA activation at higher f_o in Session 2, the overall spatial organization of MAPs remained consistent across sessions when using the proposed normalization procedure. In contrast, MAPs generated with the traditional normalization (Figure 3.4a) exhibited greater variability, with inconsistent quadrant-level distributions and reduced agreement across sessions.

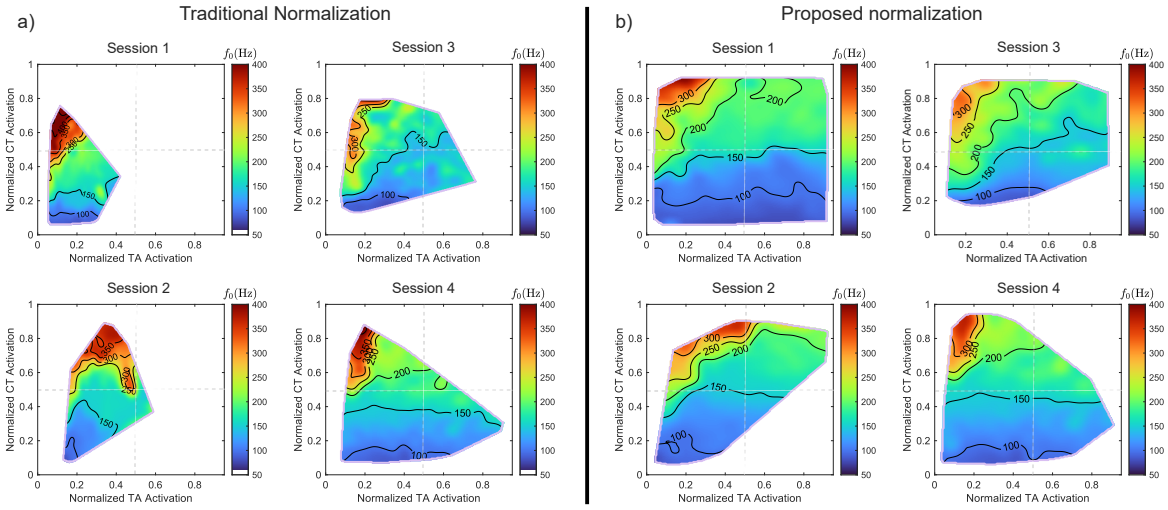


Figure 3.4: MAPs with f_o showing the relationship between normalized CT and TA muscle activation for each of the four sessions. Data points were collected from all tasks involving the vowel /a/, including sustained vowels and pitch glides, as well as the /pae/ task. The x-axis represents TA activation, while the y-axis shows CT activation, with colors indicating f_o in Hz (iso-contours indicated for f_o).

These observations were supported by quantitative results summarized in Table 3.3, showing a clear improvement in inter-session consistency. The mean

Table 3.3: Comparison between the traditional and the new proposed normalization in terms of r-value, MAE, RMSE, ICC and CV between f_o sessions.

Sessions		**Traditional Normalization**			**New Normalization**		
		(r)	MAE (Hz)	RMSE (Hz)	(r)	MAE (Hz)	RMSE (Hz)
S1	S2	0.87	59.21	75.13	0.88	17.07	25.06
S1	S3	0.77	94.04	131.84	0.85	21.86	27.19
S1	S4	0.89	61.81	88.21	0.91	17.20	22.33
S2	S3	0.49	111.24	156.77	0.83	23.63	33.13
S2	S4	0.82	44.31	67.11	0.88	20.44	26.58
S3	S4	0.53	53.02	74.22	0.84	19.18	25.58
Mean value		**0.73**	**70.60**	**98.88**	**0.87**	**19.23**	**26.98**
ICC(2,1)		**0.52**	–	–	**0.83**	–	–
CV medio (%)		**44.62**	–	–	**15.48**	–	–

Pearson correlation coefficient (r) increased from 0.73 to 0.87 using the proposed normalization. The MAE decreased by approximately 73% (from 70.60 Hz to 19.23 Hz), and the RMSE dropped by nearly 72% (from 98.88 Hz to 26.98 Hz). Additionally, the coefficient of variation (CV) was reduced from 44.62% to 15.48%, while the intraclass correlation coefficient [ICC(2,1)] improved from 0.52 to 0.83, indicating a substantial increase in both accuracy and reliability of the normalized data across sessions.

Figure 3.5 displays the corresponding MAPs of SPL as a function of normalized CT and TA activation across the four sessions, based on the same data presented in Figure 3.4. The left plots correspond to the traditional normalization, whereas

the right plots illustrate the proposed approach. Although the proposed normalization reduced intrasubject variability, the SPL-based maps exhibited less defined spatial patterns compared to the CT–TA–f_o maps. SPL values above 90 dB were generally located in regions where both CT and TA activation reached moderate to high levels, while values below 85 dB consistently appeared in the lower-left quadrant, corresponding to low activation of both muscles. In contrast, the traditional normalization displayed greater variability and less stable contour distributions, particularly in regions characterized by elevated TA activation.

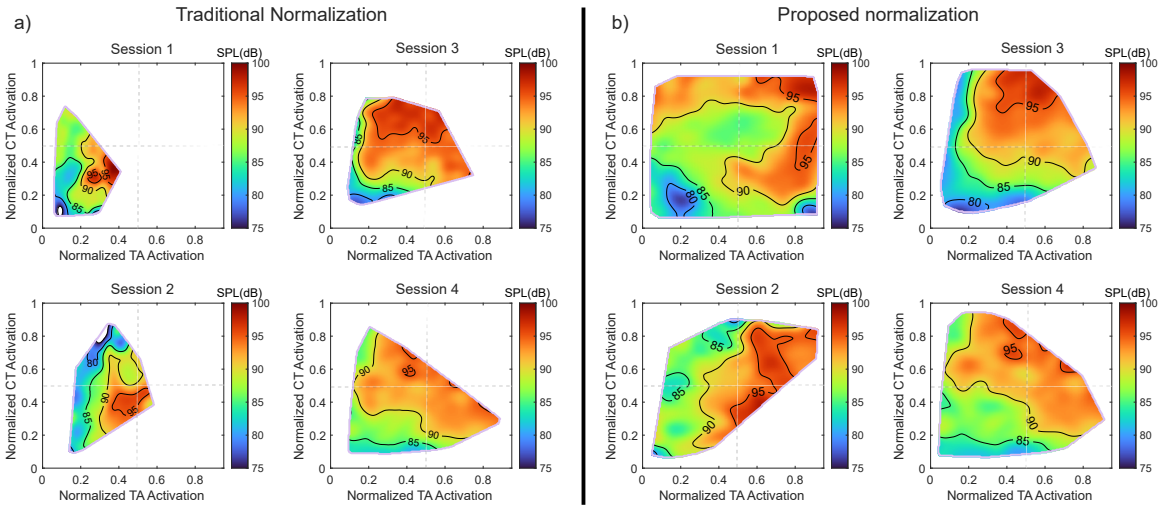


Figure 3.5: MAPs with SPL showing the relationship between normalized CT and TA muscle activation for each of the four sessions. Data points were collected from all tasks involving the vowel /a/, including sustained vowels and pitch glides, as well as the /pae/ task. The x-axis represents TA activation, while the y-axis shows CT activation, with colors indicating SPL in dB (iso-contours indicated for SPL).

As shown in Table 3.4 the mean r increased from 0.52 to 0.73. Mean MAE was reduced by 55% (from 8.20 dB to 3.66 dB), and RMSE decreased by over 58% (from 10.68 dB to 4.47 dB). The coefficient of variation dropped from 41.45% to 19.62%, representing a 52.7% reduction, and ICC(2,1) improved from 0.51 to 0.76.

Table 3.4: Comparison between the traditional and the proposed normalization in terms of r-value, MAE, RMSE, ICC, and CV for SPL across sessions.

Sessions		**Traditional Normalization**			**Proposed Normalization**		
		(r)	MAE (dB)	RMSE (dB)	(r)	MAE (dB)	RMSE (dB)
S1	S2	0.77	5.79	7.33	0.85	3.96	4.69
S1	S3	0.48	8.54	10.42	0.75	4.03	4.84
S1	S4	0.58	7.81	9.62	0.72	3.14	3.70
S2	S3	0.10	13.46	17.11	0.54	4.52	5.65
S2	S4	0.30	9.97	12.82	0.60	3.78	4.71
S3	S4	0.89	3.63	4.77	0.93	2.51	3.22
Mean		**0.52**	**8.20**	**10.68**	**0.73**	**3.66**	**4.47**
ICC(2,1)		**0.51**	–	–	**0.76**	–	–
CV (%)		**41.45**	–	–	**19.62**	–	–

Figure 3.6 shows the f_0 maps obtained after applying the same normalization method to three additional participants, including two males and one female, following the same recording protocol. All maps were computed from a single session per subject and include data from identical phonatory tasks. After normalization, a similar f_0 value of approximately 150 Hz is consistently observed around a normalized activation level of 0.4 across all participants, suggesting that

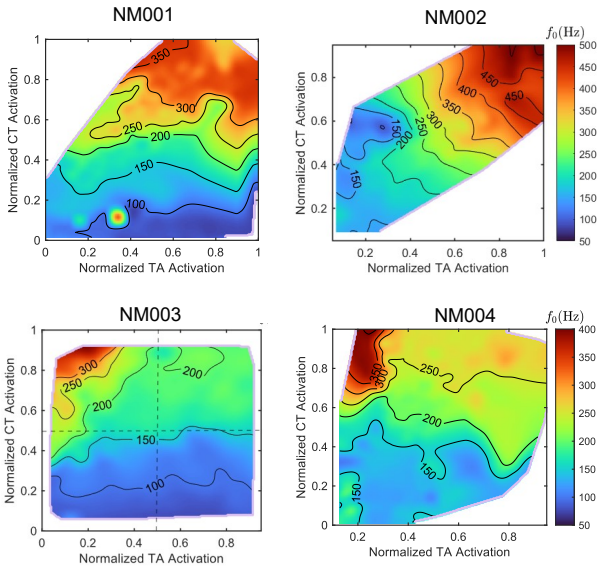


Figure 3.6: Maps of f_0 showing the relationship between normalized CT and TA muscle activation for each of the four participants. Each panel corresponds to a different subject (NM001–NM004), computed from a single recording session per subject using the proposed normalization framework. Data points were collected from all tasks involving the vowel /a/, including sustained vowels and pitch glides, as well as the /pae/ task. The x-axis represents TA activation, while the y-axis shows CT activation, with colors indicating f_0 in Hz (iso-f_0 contours are shown).

the proposed normalization framework yields comparable activation–acoustic relationships across subjects. However, the female participant exhibits a markedly different activation pattern, which may indicate a suboptimal placement of the intramuscular electrode, potentially positioned closer to the CT muscle rather than the TA, thus biasing the observed activation trend.

3.5 Chapter Conclusions

This chapter presented the development and validation of an acoustic-based normalization framework designed to reduce inter-session variability and improve intermuscular consistency in EMG recordings of intrinsic laryngeal muscles. The proposed method aligned EMG amplitudes with subject-specific acoustic references of f_0 and SPL, allowing a physiologically meaningful comparison across sessions and muscles.

The results demonstrated that normalization using acoustic targets significantly enhanced the stability and repeatability of MAPs, particularly for the CT and TA muscles, achieving higher correlation and lower error metrics compared with traditional normalization. The proposed framework also provided robust representations of f_0 and SPL distributions across muscle activations, consistent with prior physiological and computational findings.

The next chapter builds upon these results by applying decomposition algorithms to HDsEMG signals to extract motor unit activity and explore their spatial and temporal characteristics in the estimation of intrinsic laryngeal muscle activation.

4 HDsEMG signals Decomposition Method

This chapter presents the complete methodological framework developed in this thesis for estimating the intrinsic activation of the CT muscle from HDsEMG recordings. Building upon the iEMG analyses and physiological characterization introduced in the previous chapter, the focus is shifted toward non-invasive estimation strategies capable of capturing both the temporal and spatial organization of CT motor unit activity. To this end, a multi-stage processing pipeline is proposed, integrating HDsEMG signal preprocessing, blind source separation–based motor unit decomposition, and reconstruction of MUAP spatial distributions. Given the lack of a direct ground truth for surface-based decomposition, a two-source validation experiment is first introduced to assess the extent to which intramuscular CT activity can be detected and reliably represented at the surface level. Following this validation stage, a spatial processing framework is developed to characterize the distribution of identified motor units within the muscle, including the estimation of their center of gravity and relative depth beneath the electrode grid. These spatial descriptors are subsequently used to separate motor units into superficial and deep activation regions, enabling the estimation of cluster-specific CSTs as representations of collective neural drive. Finally, the resulting CSTs

are quantitatively compared with the iEMG envelope using correlation, error, and coherence metrics, allowing the assessment of temporal agreement, amplitude discrepancies, and shared synaptic input between invasive and non-invasive measurements. Together, the methods presented in this chapter establish a validated and physiologically informed framework for extracting CT muscle activation from HDsEMG signals, providing a critical foundation for the modeling and neural network–based estimation strategies explored in the following chapter.

4.1 Laboratory Recordings

The iEMG and HDsEMG values were recorded *in vivo* at the Voice Production Laboratory (VPLab) of the Advanced Center for Electrical and Electronic Engineering (AC3E) under controlled acoustic and electromagnetic conditions. Thirteen participants (ten males and three females; mean age 32 years, SD = 8.0 years) with no history of voice disorders took part in the study. All subjects were evaluated by a certified speech-language pathologist prior to data collection to confirm normal phonatory status and provided written informed consent according to the ethical protocols approved by the Universidad de Valparaíso (CEC-UV 273-23). Participants exhibited low levels of subcutaneous fat in the neck region, a selection criterion intended to minimize attenuation of intramuscular potentials from the CT muscle. Each participant underwent preliminary training sessions

to ensure consistent task performance during data acquisition. The experimental setup included simultaneous recordings of iEMG from the CT muscle, HDsEMG and a recording of acoustic pressure.

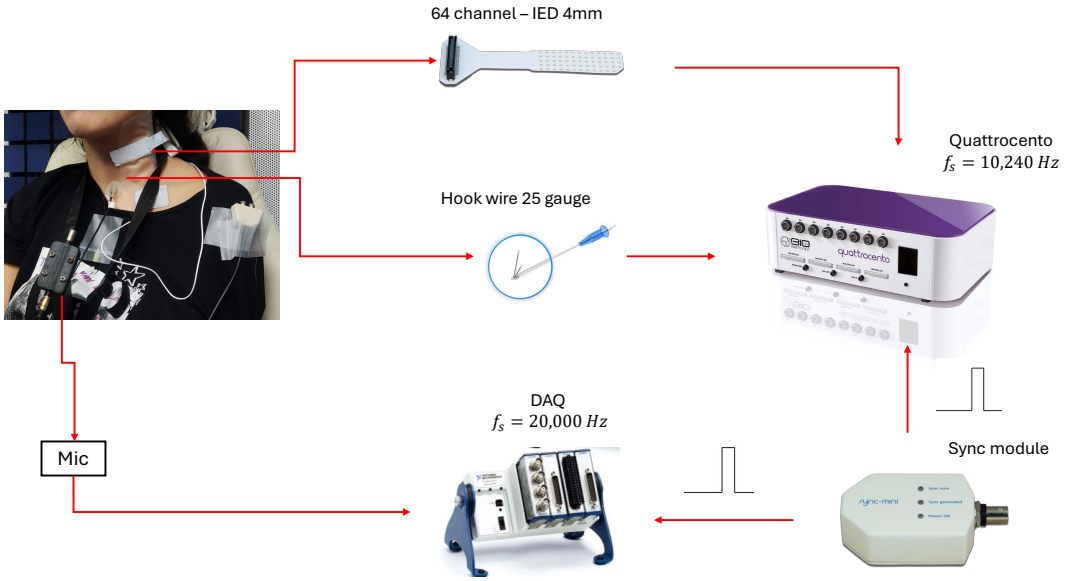


Figure 4.1: Experimental setup for simultaneous recordings of iEMG, HDsEMG, and acoustic signals. The hooked-wire electrode was inserted into the CT muscle and connected to the Quattrocento system (10,240 Hz), while the HDsEMG grid was placed on the laryngeal surface. A condenser microphone (20 kHz) recorded the acoustic signal. Both EMG and acoustic systems were synchronized using a SyncSE module and a National Instruments DAQ device.

iEMG and HDsEMG signals were acquired using a Quattrocento system (OT Bioelettronica, Italy). Pre-sterilized disposable hooked-wire electrode pairs were

inserted into the CT muscle by a neurologist specialized in laryngeal procedures, using a 27-gauge, 30 mm-long hypodermic carrier needle (221-28SS-730, Rhythmlink International LLC). The stainless-steel, enamel-insulated wires were 44-gauge (0.002" diameter) with approximately 2 mm of exposed metal at the bent-over (hooked) ends extending 2–6 mm from the carrier needle.

The neurologist inserted the carrier needle through the anterior neck skin into the cricothyroid space. Targeting the CT muscle, the needle was inserted approximately 5 mm lateral to the midline at an angle of about 50° from the sagittal plane, to a depth of roughly 10 mm [39, 6]. Prior to insertion, the skin surface was prepared with an abrasive solution followed by alcohol cleaning to reduce impedance. During insertion, the iEMG signal was reproduced in real time through a loudspeaker, allowing the neurologist to audibly verify whether the needle was positioned within muscle tissue, fatty tissue, or near the innervation zone. Additional validation was performed by monitoring the signal response during tasks that selectively engage the CT muscle (e.g., sustained vowels at high and low pitch).

After confirming correct iEMG placement, a 64-channel HDsEMG electrode grid with 4 mm inter-electrode distance (IED) was positioned over the laryngeal region as determined by the neurologist, ensuring alignment with the surface projection of the CT muscle and avoiding interference with the hooked-wire leads. The HDsEMG preamplifier and Quattrocento amplifier were connected to sepa-

rate ground references: one placed on the clavicle and another on the shoulder. Baseline noise levels were verified to remain within 10–20 µV RMS prior to recording.

Table 4.1: Experimental protocol. The phonatory tasks are categorized into three sets: sustained vowels, pitch glides, and repeated syllable tasks, with variations in pitch and loudness conditions indicated. Non-phonatory and combined tasks are also listed.

Category	**Task**	**Pitch**	**Loudness**
Phonatory Set 1	Sustained vowels /a/, /i/	Low, Habitual, High	Comfortable
		Habitual	Soft, Loud
Phonatory Set 2	Pitch glides on /a/, /i/	Ascending, Descending	Comfortable
Phonatory Set 3	Repeated /pae/	Low, Habitual, High	Loud to Soft
Non-Phonatory	Cough, Swallow, Throat clear		

In addition, the simultaneous recording of acoustic pressure was incorporated using a condenser microphone (MKE104, Sennheiser Electronic GmbH, Wedemark, Germany) placed 10 cm from the lips, with full bandwidth in the 0–6 kHz range and connected to a DAQ system (NI Compact DAQ, National Instrument). Both the acoustic and EMG systems were synchronized using a SyncSE device (OT Bioelettronica), which generated a 5 V, 10 ms square TTL pulse serving as a trigger to temporally align the onset of both acquisition systems (Figure 4.1).

Following instrumentation, participants performed a series of phonatory tasks, including sustained vowels (/a/ and /i/) at different pitch and loudness levels, pitch glides, and /pae/ sequences with descending intensity, as defined in Table 4.1. Each task was repeated across multiple trials to ensure signal stability and reproducibility. All signals were continuously monitored to confirm data quality and to minimize movement artifacts during acquisition.

4.2 Proposal scheme for activation estimation from HDsEMG

Figure 4.2 presents the general framework proposed to estimate the intrinsic activity of the CT muscle from HDsEMG signals. The framework integrates decomposition and spatial analysis methods to infer the activation and distribution of motor units, similar to the methodologies described in [93, 94, 95, 96, 97].

The first two blocks correspond to preprocessing and removal of electrocardiographic (EKG) artifacts, aimed at filtering and improving the signal-to-noise ratio. The filtered signals were extended and whitened to decorrelate the channels and enhance spatial independence between sources, a necessary step prior to blind source separation. The EMG decomposition stage was performed using the FastICA method, which relies on a fixed-point algorithm to estimate individual MU spike trains, as described by Farina and Negro [93]. After decomposition, a fil-

tering process was conducted in which each identified MU was evaluated through metrics such as the coefficient of variation (CoV) and the decomposition index (DI), retaining only physiologically reliable motor units [98].

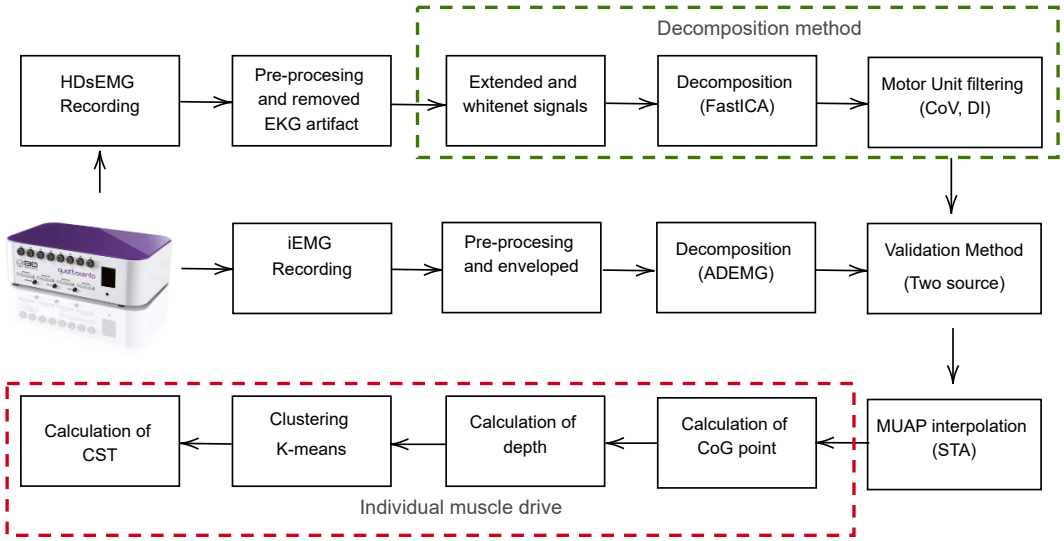


Figure 4.2: Proposed scheme and HDsEMG signal processing. The input of the method was HDsEMG and the output (activity and individual muscle impulse) was compared with the iEMG standard.

Following this filtering stage, the spike-triggered averaging (STA) method [99, 98, 100] was applied to reconstruct the spatial distribution of each MUAP across the electrode grid. From these maps, the center of gravity (CoG) of each MUAP was computed to determine its spatial position within the muscle area [101], while an exponential attenuation model was applied to infer its relative depth beneath the electrode matrix [102]. Once both spatial position and estimated depth were

obtained for each MU, three-dimensional plots were generated to visualize their distribution. A k-means clustering algorithm was then applied to group motor units with similar spatial characteristics, allowing separation between superficial and deep activation regions. Finally, for each group, a CST was computed to represent the overall neural activation of each MU cluster [103, 104, 95, 105].

Simultaneously, a two-source validation experiment was conducted to identify motor units detected both at the surface and intramuscularly, allowing verification that part of the surface information originated from the same intramuscular sources[93, 106, 107]. The cluster containing these common MUs was identified, and its corresponding CST was compared with the envelope signal obtained from the iEMG reference. Finally, the Pearson correlation coefficient, RMSE, and coherence metrics were computed to quantitatively compare the CT activation estimated from surface data with the iEMG gold-standard recordings.

4.2.1 Processing Signals

The acoustic signals were sampled at 20 kHz / 16 bits , low-pass filtered with an 8 kHz cutoff frequency, and calibrated in physical units [108]. Both SPL and f_o were computed from 50 ms analysis windows of the calibrated acoustic signal, with a 25 ms overlap between consecutive frames. SPL was derived from the RMS value of each segment of the microphone signal, recorded at a distance of 15 cm from the lips [58]. The f_o was extracted using the MATLAB `pitch` function,

implementing the harmonic logarithmic summation (LHS) method [85].

The raw HDsEMG and iEMG signals were sampled at 10 kHz / 24 bits, calibrated in physical units, and synchronized with the acoustic signal using a SyncSE device (OT Bioelettronica). The raw HDsEMG signal was first processed using a fourth-order band-pass filter (20 - 400 Hz) [14].

Subsequently, an EKG artifact removal procedure was applied to the HDsEMG signals. Due to the anatomical proximity between the laryngeal muscles and the carotid artery, ECG contamination was frequently observed, particularly in the low-frequency components of the EMG recordings. To mitigate this effect, an approach combining low-pass filtering and median filtering (LPF+MF) was employed, as described in [109]. Although high-pass filtering (HPF) is commonly used to remove the main ECG frequency components, which are typically confined to the 1–30 Hz range [15], this procedure also removes relevant low-frequency components of the EMG signal. Therefore, instead of blindly eliminating these frequency components, the ECG signal was first estimated and subsequently subtracted in the time domain. Specifically, the dominant ECG frequency components (1–30 Hz) were tracked using a fourth-order elliptic IIR low-pass filter with pass-band ripple and stop-band attenuation of 0.1 dB and –50 dB, respectively. Filtering was performed in a forward–backward zero-phase manner to avoid phase distortion associated with the non-linear phase response of IIR filters. Afterwards, a temporal median filter with a 40 ms window was applied to further refine the

estimation and extraction of the ECG patterns [109].

For the raw iEMG signals, full-wave rectification was applied and the signals were low-pass filtered at 8 Hz to obtain the amplitude envelope, as described in [39, 110]. The iEMG signals were normalized as described in Chapter 3.

4.2.2 EMG Decomposition Method

The EMG decomposition method consists of the identification of the firing patterns of the motor units whose action potentials contribute to the recorded signal from EMG and allows the study of the recruitment and firing trains of the motor units [107, 111, 15].

For the **decomposition of iEMG**, which is single-channel, method described by K. C. McGill [112, 113] is used. This approach combines template matching with an optimization-based algorithm for resolving superimposed motor unit action potentials. In this framework, the recorded EMG segment $\mathbf{y}$ is modeled as the linear superposition of time-shifted MUAP templates,

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{h}_k(\tau_k) + \mathbf{n}, \tag{4.1}$$

where $\mathbf{h}_k$ represents the MUAP template of the k-th motor unit, τ_k its discharge time, and $\mathbf{n}$ additive noise.

Template matching is first used to identify recurrent MUAP waveforms and their approximate occurrence times. When superpositions occur, the method

formulates their resolution as an optimization problem, seeking the set of discharge times that minimizes the squared reconstruction error,

$$E(\{\tau_k\}) = \|\mathbf{y} - \sum_{k=1}^{K} \mathbf{h}_k(\tau_k)\|^2. \tag{4.2}$$

The optimal alignment is obtained using a branch-and-bound search strategy combined with sub-sample temporal refinement through interpolation, allowing accurate separation of overlapping MUAPs from iEMG[112].

However, for **decomposition of HDsEMG** it has the limitation of not being able to separate the action potentials of different motor units that overlap in time. To overcome this, blind source separation (BSS) techniques have been developed [100], which estimate the underlying sources contributing to the observed EMG by inverting a linear mixture model. Among these, the Fast Independent Component Analysis (FastICA) with a fixed-point algorithm proposed by Negro et al. [93] has demonstrated high robustness and computational efficiency. This approach models the MU discharges as sparse, quasi-binary sources composed of Dirac delta impulses, allowing reliable separation of motor unit activity (Figure 4.3).

First,the HDsEMG signals can be modeled as a convolutive mixture of MUAPs:

$$x_i(k) = \sum_{l=0}^{L-1} \sum_{j=1}^{n} h_{ij}(l) s_j(k-l) + n_i(k), \tag{4.3}$$

where $x_i(k)$ is the signal recorded at the i-th channel, $h_{ij}(l)$ represents the

MUAP detected at the i-th electrode from the j-th motor unit, $s_j(k)$ is the spike train of that unit, and $n_i(k)$ denotes additive noise. The decomposition process aims to estimate both the source spike trains $\mathbf{s}(k)$ and their spatial filters $\mathbf{h}_{ij}(l)$ from the recorded signals $\mathbf{x}(k)$.

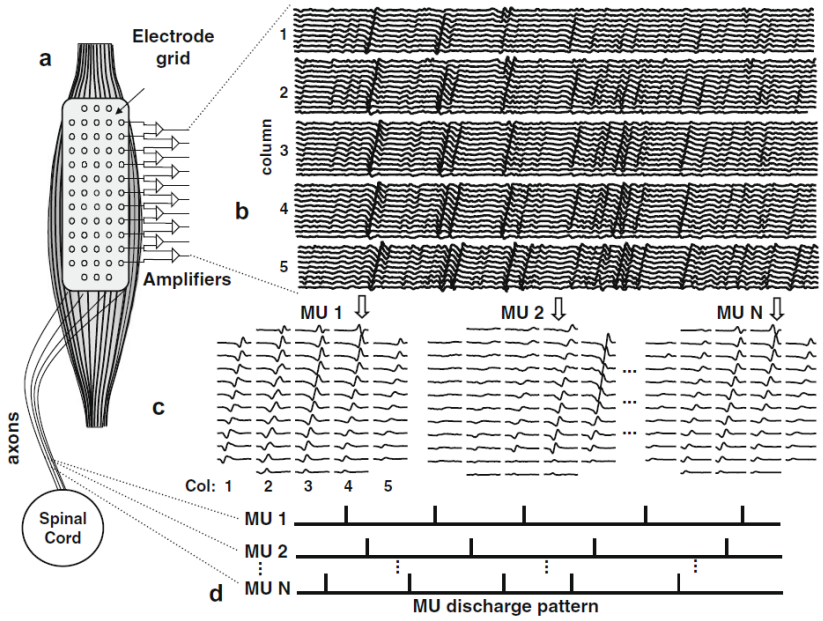


Figure 4.3: Illustration of the decomposition process of HDsEMG. (a) sEMG are recorded from the muscle with a multi-channel detection system. (b) Segment of 500ms duration of bipolar EMG detected. (c) spatio-temporal representation of three representative MUs. (d) Estimated discharge patterns for the three MUs. Figure extracted from [114]

Next, the EMG signals were temporarily extended to increase the number of effective channels. An extension factor of 1000 samples was used, following the

configuration proposed by Negro et al. [93]. The extended signal matrix $\tilde{\mathbf{x}}(k)$ was whitened to remove spatial correlations and equalize the variances across channels:

$$\mathbf{z}(k) = \mathbf{W}\tilde{\mathbf{x}}(k), \tag{4.4}$$

where the whitening matrix $\mathbf{W}$ is computed as

$$\mathbf{W} = \mathbf{U}\mathbf{D}_S^{\gamma}\mathbf{U}^T, \quad \text{with } \gamma = -1/2, \tag{4.5}$$

and $\mathbf{U}$ and $\mathbf{D}_S$ are obtained from the eigenvalue decomposition of the covariance matrix of the extended signals:

$$\mathbf{R}_{\tilde{x}\tilde{x}} = \mathbf{U}\mathbf{D}_S\mathbf{U}^T. \tag{4.6}$$

After whitening, the independent components were estimated using the FastICA method. Each separation vector $\mathbf{w}_i$ was updated iteratively by the fixed-point rule:

$$\mathbf{w}_i(n) = \mathbb{E}\{\mathbf{z}g[\mathbf{w}_i(n-1)^T\mathbf{z}]\} - \mathbb{E}\{g'[\mathbf{w}_i(n-1)^T\mathbf{z}]\}\mathbf{w}_i(n-1), \tag{4.7}$$

where $g(\cdot)$ is the contrast function, chosen here as logcosh$(\cdot)$, and $g'(\cdot)$ its derivative. Orthogonalization and normalization of $\mathbf{w}_i$ were performed after each iteration to ensure statistical independence among extracted sources. The algorithm converged after 150 iterations, yielding the estimated sources. Each source estimate was computed as:

$$\hat{s}_i(k) = \mathbf{w}_i^T\mathbf{z}(k). \tag{4.8}$$

After source estimation, the discharge times of each potential MU were identified by detecting high-amplitude peaks in the source signal. A K-means++ clustering procedure was then applied to separate high-amplitude discharges (corresponding to stronger MUs) from lower-amplitude components (corresponding to noise MUs). The separation vector $\mathbf{w}_i$ was subsequently updated to minimize the variability in the firing pattern of the detected unit. This iterative refinement continued until the coefficient of variation of the discharge times reached a minimum, indicating stable firing behavior. Once stability was achieved, the silhouette index (SIL) was computed for each candidate unit to quantify clustering quality and isolation. Units with SIL > 0.9 were retained as reliable sources and added to the matrix of reliable UMs.

After decomposition, MUAP shapes were reconstructed using the **STA method**, aligning signal segments within a 30 ms window ($\pm$15 ms around each discharge). The averaged waveforms yielded a representative MUAP template for each unit. For intramuscular EMG, a single template was obtained per unit, while in HD-sEMG one averaged template per channel was generated, resulting in 64 spatially distributed templates per identified motor unit. Following reconstruction, motor units were evaluated using two metrics to ensure physiological reliability. First, the CoV of inter-spike intervals was computed to quantify firing regularity:

$$\mathrm{CoV} = \frac{\mathrm{SD}_{\mathrm{ISI}}}{\mathrm{mean}_{\mathrm{ISI}}} \times 100\%. \tag{4.9}$$

Units with CoV values exceeding 50% were discarded as unstable. Second, the DI introduced by Florestal et al. [98] was used to estimate the difficulty of separating each MU, defined as:

$$\mathrm{DI}_{ki} = \frac{\min\left(\left\|m_{ki}\right\|,\ \left\|m_{ki} - m_{k*i}\right\|\right)}{v_i^{\mathrm{RMS}}} \tag{4.10}$$

where m_{ki} represents the MUAP template of the k-th MU in the i-th channel, $m_{k*k,i}$ the most similar MUAP or baseline noise, and V_i^{RMS} the RMS of the corresponding EMG channel. The DI expresses the morphological distinctiveness of a MUAP relative to the surrounding activity, normalized by background noise. Units presenting low DI values (< 3), indicative of poor separability, were discarded from subsequent analysis.

4.2.3 Two-Source Validation Method

A two-source validation experiment was conducted to assess whether part of the intramuscular activity of the CT muscle can be detected at the surface or is lost due to the superposition of neighboring muscles. This approach, widely adopted in EMG decomposition studies [107, 115, 116, 117], is considered the only sufficiently validated method to quantify decomposition accuracy in the absence of a gold standard.

The method relies on the simultaneous decomposition of two EMG signals, an iEMG and a HDsEMG, recorded from the same muscle region but using indepen-

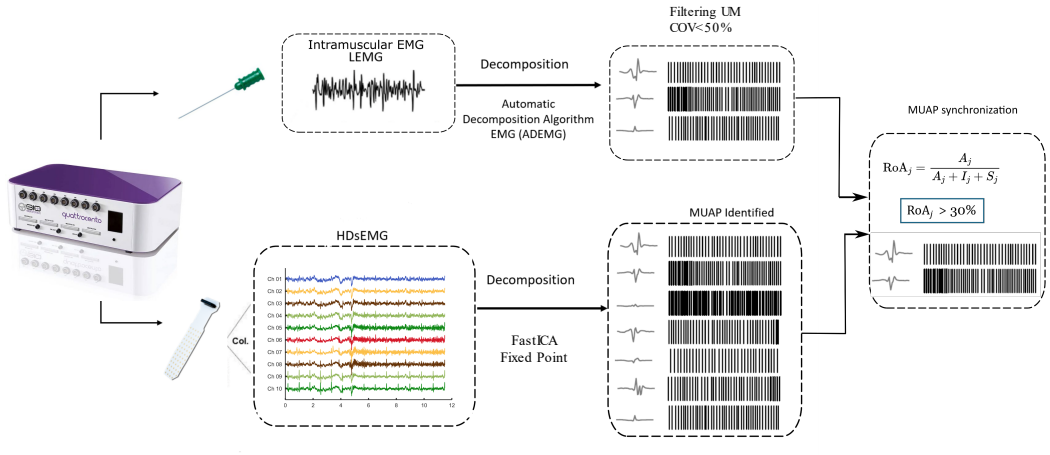


Figure 4.4: Representation of two-source decomposition method for validation

dent acquisition systems and decomposition algorithms. Agreement between the discharge timings of both decompositions is used as a conservative indicator of validity, since identical errors across independent recordings are highly improbable. The degree of correspondence between spike trains was quantified using the rate of agreement (RoA) metric [100], defined as:

$$\text{RoA} = \frac{c_j}{c_j + A_j + B_j} \times 100\%, \tag{4.11}$$

where c_j is the number of coincident discharges of the j-th motor unit detected by both decompositions within a tolerance of ± 1 ms, A_j the discharges identified only in the surface recording, and B_j those detected only intramuscularly. Motor units were considered common if at least 25% of their discharges overlapped, while units with inter-spike interval variability exceeding 65% were excluded. Figure 4.4 illustrates the two-source setup and decomposition framework, highlighting

the independent processing of intramuscular and surface signals through distinct algorithms, which strengthens the reliability of the cross-validation results.

4.2.4 Individual Muscle Drive

After validating the presence of simultaneous motor units using the two-source method, a spatial framework was implemented to estimate the individual activation of the CT muscle. Since only one or two simultaneous motor units do not provide a representative description of overall CT activation, it was necessary to determine the anatomical region from which these units originated, using them as spatial references. To this end, a spatial processing approach was employed to determine the *individual muscle drive* based on the distribution of MUAPs across the electrode grid. The first step involved computing the CoG of each MUAP, following the method proposed by Xia *et al.* [97]:

$$\bar{x} = \frac{1}{P}\sum_{j=1}^{20} x_j \sum_{i=1}^{13} p_{ij}, \qquad \bar{y} = \frac{1}{P}\sum_{i=1}^{13} y_i \sum_{j=1}^{20} p_{ij} \tag{4.12}$$

where x_j and y_i denote the horizontal and vertical positions of each electrode, p_{ij} is the normalized peak-to-peak value of the MUAP at electrode (i, j), and P is the total sum of all p_{ij}. Each CoG point thus represents the spatial location of an individual motor unit within the grid.

Motor unit depth estimation

The estimation of MU depth was performed by fitting a two-dimensional Gaussian surface to the peak-to-peak amplitudes of the MUAP distributions recorded with HDsEMG. This model assumes that the potential decays more rapidly perpendicular to the muscle fibers and more slowly along their direction, thus allowing inference of both fiber orientation and MU depth [102]. The surface distribution of MUAPs was modeled as:

$$f(x, y) = A \cdot \exp\left[-\left(a(x - x_0)^2 + 2b(x - x_0)(y - y_0) + c(y - y_0)^2\right)\right] \tag{4.13}$$

where A is the peak amplitude, and the coefficients a, b, and c were obtained by nonlinear least-squares fitting of the Gaussian function to the measured surface amplitudes, following the approach described by Lundsberg *et al.* [102].

To determine MU depth, the model assumes exponential signal attenuation with distance according to the Bouguer–Lambert–Beer law:

$$V = V_0 e^{-Qd} \tag{4.14}$$

where V_0 is the amplitude at the MU center, d is depth, and Q is the attenuation factor (fixed at $Q = 1$). The relationship between the Gaussian peak amplitude V_{max} and the half-maximum amplitude V_{FWHM} defines the effective

radial distance r_F, yielding:

$$\frac{V_{max}}{V_{FWHM}} = \frac{V_0 e^{-Qd}}{V_0 e^{-Qr_F}} \Rightarrow \ln 2 = Q(r_F - d) \tag{4.15}$$

$$d = r_F - \frac{\ln 2}{Q} \tag{4.16}$$

where r_F represents the distance from the surface position at half-maximum potential, and d is MU depth. This equation provides a direct estimation of MU depth based on the fitted Gaussian spread.

After estimating the depth of each motor unit, a k-means clustering algorithm was applied to group the CoG positions of all identified MUs into two spatial clusters, corresponding to superficial and deep activation regions. This classification allowed separating the contributions of MUs located near the surface from those situated deeper within the CT muscle. For each cluster, a CST was computed by summing the binary discharge trains of all MUs within that group and smoothing the result using a 30 ms Hanning window, as described by Farina *et al.* [103]. To enable comparison across tasks and sessions, the CST amplitude was normalized by the maximum value of a 500 ms activity index, defined as the number of active motor units within a non-overlapping window of that duration. The resulting normalized CST represents an estimate of the collective neural drive for each spatial cluster of MUs.

Finally, to compare the CST signals resulting from the decomposition with the iEMG envelopes (gold standard), three metrics were used: the correlation

coefficient (R), the RMSE, and the magnitude-squared coherence (C_{xy}). The correlation coefficient was used to assess the ability of the CST signal to reproduce the temporal trend of the iEMG envelope, while the RMSE quantified the deviation between their normalized amplitudes. The coherence was employed to evaluate whether two signals shared a common synaptic input, that is, whether they originated from the same neural source.

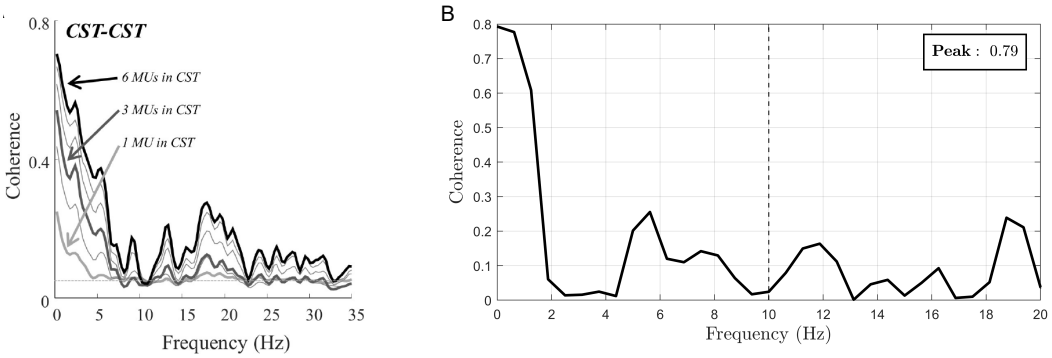


Figure 4.5: Coherence analysis of CSTs. (A) Intra-CST coherence as a function of the number of motor units included in the CST. An increase in the number of MUs enhances the reliability of the coherence estimation, with stable values achieved when at least six units are combined. (B) Coherence between intramuscular and HDsEMG CSTs from CM006 patient, showing a peak of 0.79 below 10 Hz, which indicates a strong shared synaptic input between both decomposition sources.

The coherence function between two signals $x(t)$ and $y(t)$ was defined as:

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)\,P_{yy}(f)}, \tag{4.17}$$

where $P_{xy}(f)$ represents the cross-power spectral density between the two signals,

and $P_{xx}(f)$ and $P_{yy}(f)$ denote their respective auto-power spectral densities [12].

Coherence is a frequency-domain measure widely used in EMG studies to evaluate signal crosstalk or to infer the presence of shared synaptic inputs between motor units [118, 119]. High coherence values at low frequencies (typically below 10 Hz) indicate a strong common synaptic drive between the signals being compared [118]. As shown in Figure 4.5, coherence peaks greater than 0.4 were observed at frequencies below 10 Hz, with the highest magnitudes typically occurring near 1 Hz. The left panel illustrates that as the number of motor units included in the CST increases, the resulting coherence becomes more stable and its amplitude rises, reflecting the improvement in the estimation of the common drive. A minimum of approximately six motor units per CST was required to obtain reliable coherence values. The right panel shows an example from a patient (CM006) of coherence between the HDsEMG-derived CST and the intramuscular CST, where a clear low-frequency peak ($C_{xy} = 0.79$) was observed, indicating a strong shared neural input to the CT muscle.

4.3 Results

Figure 4.6 shows an example of the results obtained from HDsEMG decomposition during a pitch-glide task. Panel (A) displays the spike occurrences over time for each identified motor unit, represented by a different color.

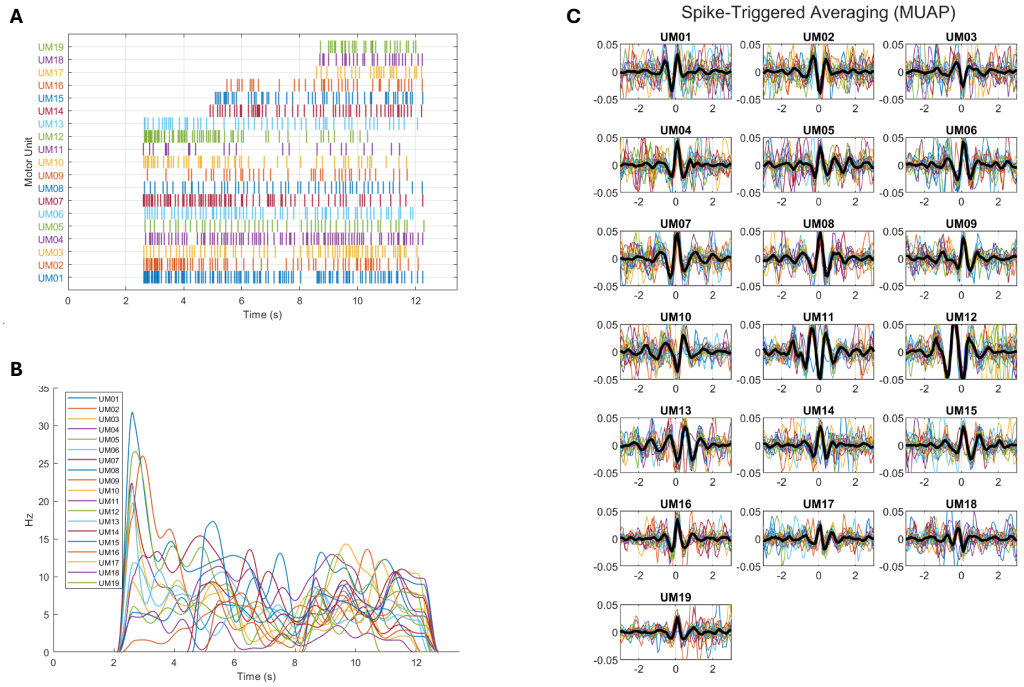


Figure 4.6: Results obtained from the HDsEMG decomposition during a pitch glide task. Panel (A) shows the spike trains of identified MUs. Panel (B) illustrates the instantaneous firing rates. Panel (C) displays the STA of the MUAPs

Panel (B) illustrates the corresponding instantaneous firing rates computed using Equation 2.4. Panel (C) shows the spike-triggered average (STA) of each MUAP, computed using segments generated by a 30-ms window aligned to each spike obtained from the decomposition. The smoothed averaged template of all segments is shown in black. Lower variability across segments indicates a more reliable motor unit, providing clearer information about its morphology..

A clear recruitment pattern can be observed, in which smaller motor units are

activated first, followed by progressively larger units, in agreement with the size principle described in the Fuglevand model. It can be observed that by second 30 all motor units have been recruited (Panel A), after which two additional units appear and rapidly increase their firing rates, while some of the initially recruited units begin to decrease their firing rates.

4.3.1 Decomposition and Two source method validation

The results of the two-source validation experiment revealed a variable number of MUs detected simultaneously across participants and phonatory tasks. As illustrated in Figure 4.7, simultaneous units were more frequently identified during sustained vowel tasks (Normal pitch and High pitch), where laryngeal muscle activation remains stable over time, resembling an isometric contraction. In contrast, the pitch-glide task showed the lowest number of concurrent MUs due to the continuous change in CT muscle length during frequency transitions, which complicates the temporal alignment of motor unit discharges.

The summary in Table 4.2 shows the RoA, CoV and DI results for each subject, reported as the mean ± standard deviation across the phonatory tasks. It can be observed that the female participants exhibited a lower or null number of simultaneous detections, likely due to greater anatomical differences in CT depth, as well as reduced muscle volume and, in some cases, higher levels of subcutaneous fat. For the remaining patients, simultaneous motor units were successfully iden-

tified, supporting the presence of CT muscle EMG components that are effectively captured by surface recordings.

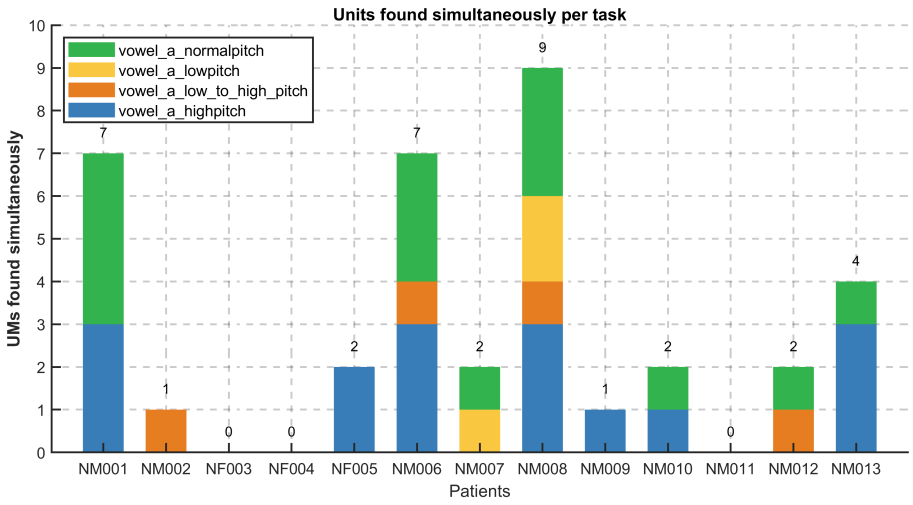


Figure 4.7: Number of MUs identified simultaneously across phonatory tasks for each subject. Bars indicate the total number of common units per patient, with colors denoting the contribution of each task (normal pitch, low pitch, low-to-high pitch, and high pitch).

Overall, acceptable CoV and DI values were obtained, remaining below 50% and above 5, respectively. Although the 25% threshold established for the other patients was exceeded, RoA values higher than 35% were not achieved, highlighting the difficulty of maintaining synchronization with deep CT motor units and the consequent loss of temporal spikes during the recording process.

The quantitative summary in Table 4.3 shows that the RoA between intramuscular and surface decompositions was higher in sustained vowel /a/ tasks

Table 4.2: Decomposition accuracy for the two-source validation experiment. Results reported per patient as mean ± standard deviation across the available phonatory tasks.

Patient	**MUs (Intra)**	**MUs (Surface)**	**MUs Common**	**RoA [%]**	**CoV [%]**	**DI**
CM001	20 ± 2	78 ± 40	3 ± 3	30.2 ± 0.9	46.7 ± 2.2	14.2 ± 5.2
CM002	25 ± 10	17 ± 10	0 ± 1	28.0 ± 0.0	34.7 ± 7.5	31.4 ± 35.4
CF003	16 ± 6	18 ± 11	0 ± 0	NaN ± NaN	47.0 ± 2.7	16.6 ± 13.2
CF004	5 ± 4	7 ± 5	0 ± 0	NaN ± NaN	53.4 ± 14.6	38.7 ± 70.1
CF005	6 ± 1	59 ± 45	1 ± 1	28.6 ± 0.1	46.5 ± 2.5	24.1 ± 11.1
CM006	14 ± 4	48 ± 48	2 ± 2	30.3 ± 0.4	45.3 ± 2.2	17.8 ± 5.9
CM007	10 ± 3	37 ± 35	1 ± 1	28.4 ± 0.5	41.6 ± 6.7	27.8 ± 20.7
CM008	16 ± 5	71 ± 40	5 ± 6	30.0 ± 1.2	42.3 ± 3.7	20.9 ± 9.4
CM009	7 ± 5	69 ± 58	0 ± 1	30.1 ± 0.0	38.8 ± 11.7	17.9 ± 3.3
CM010	12 ± 6	65 ± 52	1 ± 1	29.8 ± 0.4	46.3 ± 4.3	14.4 ± 6.3
CM011	21 ± 0	33 ± 0	0 ± 0	NaN ± NaN	49.9 ± 0.0	7.3 ± 0.0
CM012	15 ± 9	54 ± 36	1 ± 1	31.9 ± 2.6	40.9 ± 3.4	13.3 ± 6.4
CM013	13 ± 3	56 ± 41	3 ± 4	31.7 ± 3.4	40.8 ± 8.3	37.7 ± 40.1

(approximately 32–38 %), while lower RoA values (around 27 %) were observed during the pitch-glide condition. The CoV of inter-spike intervals followed a similar trend, showing greater stability (lower CoV) during steady phonations compared to pitch transitions. Overall, these results indicate that stable phonatory conditions favor the detection and synchronization of MUs between intramuscular and surface recordings.

Table 4.3: Decomposition accuracy for the two-source validation experiment. Comparison between intramuscular and surface decompositions across the phonatory tasks. Values are reported as mean ± standard deviation across the 13 patients.

Task	**MUs (Intra)**	**MUs (Surface)**	**MUs Common**	**RoA** [%]	**CoV** [%]	**DI**
Vowel /a/ highpitch	12 ± 8	73 ± 39	2 ± 2	28.6 ± 0.8	42.3 ± 7.8	28.8 ± 44.5
Vowel /a/ low_to_high_pitch	16 ± 7	12 ± 10	0 ± 0	29.7 ± 2.4	39.2 ± 9.2	33.3 ± 26.8
Vowel /a/ lowpitch	11 ± 3	53 ± 33	1 ± 1	28.1 ± 0.6	45.1 ± 1.6	17.6 ± 4.1
Vowel /a/ normalpitch	12 ± 7	62 ± 35	2 ± 4	29.5 ± 1.9	45.0 ± 2.8	18.0 ± 8.1

4.3.2 CT Estimation using CST and clusters

After the decomposition process, the STA technique was applied to reconstruct the MUAPs across the HDsEMG grid. This procedure allows visualizing the local potential fields generated by each identified MU and provides insight into the spatial propagation and morphology of the underlying electrical sources. Figure 4.8 summarizes the resulting spatial representations. Panel (A) shows the electrode placement above the CT muscle, aligned with the estimated fiber orientation. Panel (B) presents the reconstructed MUAP distributions obtained through STA, where each segment shows the averaged waveform of a single channel, corresponding to the mean MUAP template across the electrode grid.

Based on these spatial MUAP representations, the COG of each motor unit was computed using the amplitude-weighted centroid method, representing the

spatial location of the electrical activity within the muscle section. Panel (C) shows a two-dimensional heatmap of a MUAP, where the maximum amplitude corresponds to the center of electrical activity, and the red marker indicates the calculated COG. Panel (D) presents the spatial distribution of all identified COGs across the electrode array, showing the relative arrangement of motor units and allowing their grouping into regions activation.

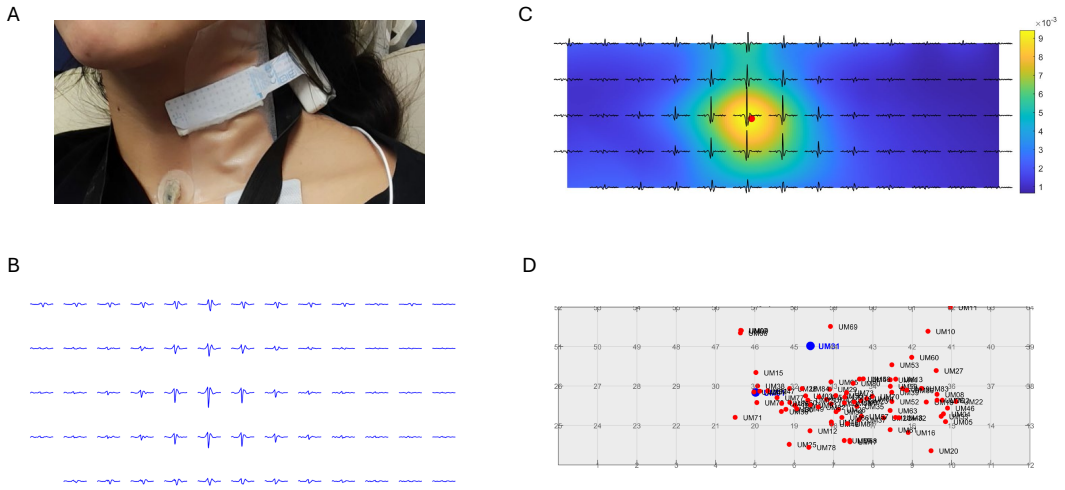


Figure 4.8: Spatial mapping of MUAPs obtained from HDsEMG decomposition. (A) Electrode placement above the CT muscle. (B) Spatial representation of MUAPs obtained by spike-triggered averaging. (C) Two-dimensional potential distribution and corresponding COG for a representative motor unit. (D) Distribution of COGs for all identified motor units across the electrode grid.

After obtaining the spatial COG coordinates from the HDsEMG signals, the next step was to estimate the MU depth to extend the representation into three

dimensions. While Zhou et al. [101] limited the COG-based approach to a two-dimensional surface representation, in this work the methodology was expanded by incorporating the cylindrical volume-conductor model previously described in Section 4.2.4 [102]. This model assumes that the amplitude of the MUAP decays exponentially with increasing distance from the recording surface, as expressed in Equation 4.12 and 4.14. This approach allowed the reconstruction of a volumetric distribution of motor units based on their estimated COG coordinates (x, y, z).

Figure 4.9 presents the three-dimensional spatial distribution of MUs identified across three patients during the vowel /a/ at low, normal, and high pitch conditions. The medio-lateral (CoG x) and antero-posterior (CoG y) coordinates correspond to the positions estimated from the HDsEMG grid, while the vertical axis represents the estimated depth (in mm) obtained from the exponential attenuation model. A hierarchical k-means clustering algorithm was applied to group the MUs according to their depth, following the same clustering methodology proposed by Jarque-Bou [97]. Two primary clusters were identified: superficial units (Cluster 1, blue) and deeper units (Cluster 2, red). Additionally, validation units identified from simultaneous intramuscular recordings were highlighted in yellow to verify consistency between methods.

Across patients, differences in MU spatial distribution were observed depending on the vowel condition. During the high-pitch production, MUs showed a more compact clustering pattern with increased depth values. In contrast, dur-

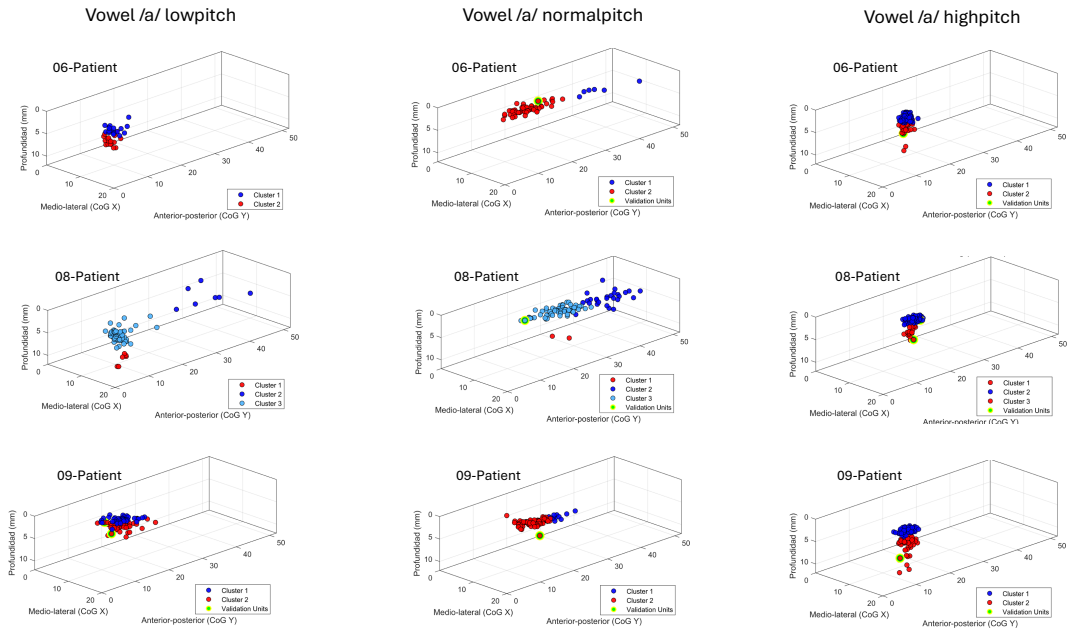


Figure 4.9: Three-dimensional spatial distribution of motor units (MUs) identified across three patients during the production of the vowel /a/ at low, normal, and high pitch conditions. The medio-lateral (CoG x) and antero-posterior (CoG y) coordinates represent the estimated CoG positions obtained from the HDsEMG grid, while the vertical axis corresponds to the estimated MU depth (mm) based on the exponential attenuation model. Two main clusters were identified using the k-means algorithm: superficial units (Cluster 1, blue) and deeper units (Cluster 2, red). Yellow markers denote validation units obtained from simultaneous intramuscular recordings.

ing the normal-pitch condition, MUs exhibited a broader spatial spread across the surface area, suggesting a wider activation region. A similar distribution was observed during the low-pitch condition, characterized by a larger lateral spread

and reduced depth across the identified units.

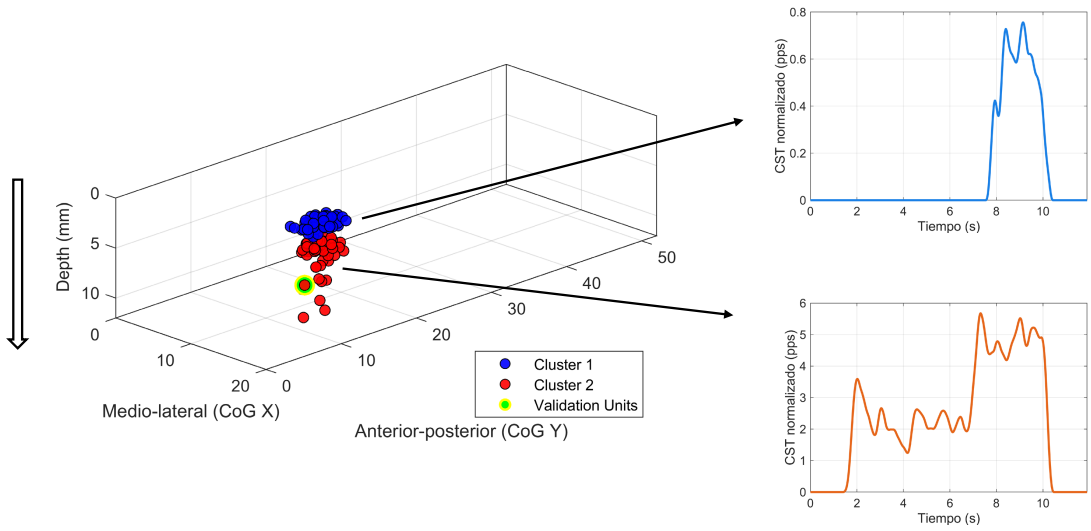


Figure 4.10: The 3D plot shows the spatial distribution of MUs based on their CoG in medio-lateral, anterior-posterior, and depth coordinates, grouped into two clusters: Cluster 1 (superficial, blue) and Cluster 2 (deep, red). The right panels depict the normalized CSTs for each cluster.

Finally, CSTs were computed separately for each cluster to assess differences in temporal activation patterns. The right panels of Figure 4.10 display the normalized CSTs for Cluster 1 and Cluster 2, respectively, showing distinct activation envelopes that correspond to different neural drive contributions across muscle depth.

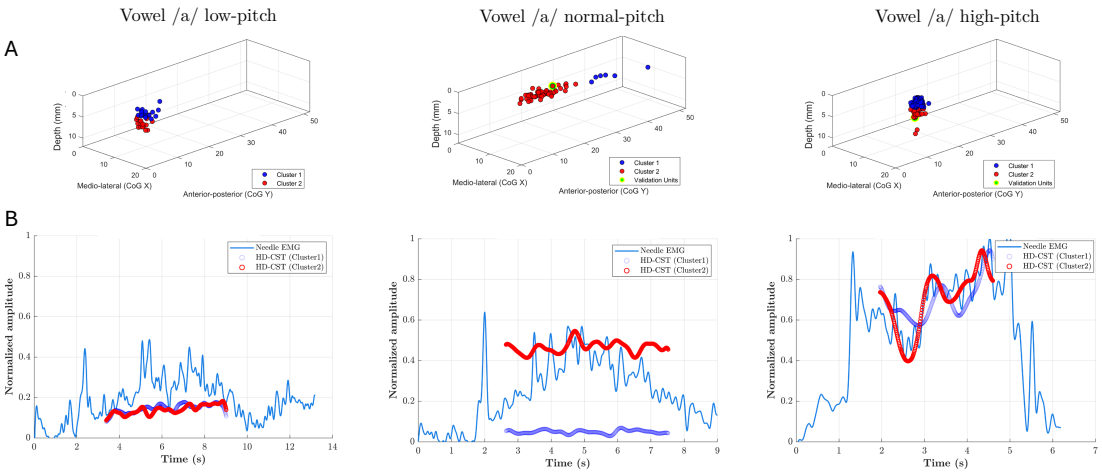
In addition to the CST computation, an activity index was generated using a 500 ms non-overlapping window, as described in [95]. This index quantified

the number of simultaneously active motor units within each time segment. The maximum value of this activity index across all tasks was used as a normalization factor for the CST amplitude, allowing consistent comparison of neural drive levels between clusters and tasks.

4.3.3 Comparison between iEMG and CST

After the two spatial clusters were defined according to MU depth, their ability to represent the activity captured by the iEMG signal was evaluated. For this purpose, two quantitative metrics were used: the R and the (C_{xy}). Both metrics were computed between the iEMG envelope and the CST derived from each cluster. The correlation and coherence analyses were performed as described in Section 4.2.4.

An example of these comparisons is presented in Figure 4.11. The three-dimensional spatial distribution of the MUs from a representative participant (CM006) is shown for three phonatory tasks of the vowel /a/ at different pitch levels (low, normal, and high). In Panel A, two clusters were identified: Cluster 1 (superficial, in blue) and Cluster 2 (deep, in red). In Panel B, the comparison between the iEMG envelope and the CST signals estimated for both clusters is displayed. In general, higher correlation and coherence values were observed for the deep cluster (Cluster 2). Although slightly lower correlation values were found for the superficial cluster (Cluster 1), its coherence values were considerably smaller.



EMG - CST		
	R	C_{xy}
Cluster-1	0.26	0.79
Cluster-2	0.31	0.77

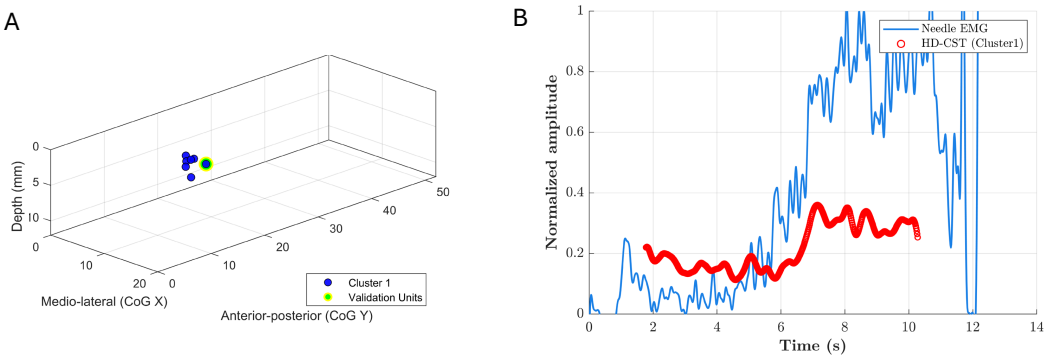
EMG-CST		
	R	C_{xy}
Cluster-1	0.166	0.548
Cluster-2	0.352	0.863

EMG - CST		
	R	C_{xy}
Cluster-1	0.63	0.84
Cluster-2	0.71	0.85

Figure 4.11: Comparison between iEMG envelopes and CST estimates obtained from two spatial clusters during three phonatory tasks of the vowel /a/ at different pitch levels. Panel A displays the three-dimensional spatial distribution of MUs for a representative participant (CM006), grouped into superficial (Cluster 1, blue) and deep (Cluster 2, red) regions, with simultaneously detected validation units highlighted in green(Seccion 4.3.1). Panel B shows the normalized iEMG envelope and CST signals for each cluster, along with their corresponding correlation (R) and (C_{xy}) values. In general, higher coherence and correlation were observed for the deep cluster, indicating a stronger shared synaptic input.

High correlation values with very similar amplitude in both cluster were observed in the high-pitch task, probably because most MUs were centered in the same region. In contrast, greater differences between clusters were found in the normal-

pitch task due to the wider spatial dispersion of MUs on the surface, where the largest differences in both correlation and coherence between clusters were observed. These results suggest that, while both clusters reflect a shared neural drive, the deeper units may exhibit a more synchronized and stable activation pattern, possibly indicating a common synaptic input to the CT muscle.



EMG-CST		
	R	C_{xy}
Cluster-1	0.90	0.81

Figure 4.12: Comparison between iEMG envelope and CST estimated for a dynamic phonatory condition (pitch-glide task). Panel A shows the spatial distribution of identified MUs and the simultaneously validated unit (green). Panel B presents the normalized iEMG envelope and the CST estimated from the superficial cluster (Cluster 1), along with the correlation (R) and coherence (C_{xy}) values.

On other hand, a dynamic phonatory condition was analyzed using the pitch-glide task. In this case, the decomposition was less efficient compared with sus-

tained vowels, as only a few motor units could be successfully identified. This limitation was mainly due to the continuous change in muscle length and the vertical displacement of the larynx during pitch elevation, which affects the spatial stability of the CT muscle signals. For the case illustrated in Figure 4.12, only ten units were decomposed. As a result, the CST amplitude differed considerably from the iEMG reference. Nevertheless, its overall trend was very similar. A high correlation value of $R = 0.90$ and a coherence of $C_{xy} = 0.81$ were obtained, indicating that although the amplitude pattern could not be fully reproduced, the trend of the neural drive was successfully captured.

Table 4.4: EMG–CST performance metrics across phonatory tasks and 3 patients.

EMG – CST	**R**	**RMSE**	$\boldsymbol{C_{xy}}$
Vowel /a/ low-pitch	0.31 ± 0.13	0.56 ± 0.29	0.61 ± 0.24
Vowel /a/ normal-pitch	0.48 ± 0.12	0.18 ± 0.05	0.79 ± 0.16
Vowel /a/ high-pitch	0.62 ± 0.19	0.15 ± 0.06	0.72 ± 0.23
Vowel /a/ pitch glide	0.84 ± 0.05	0.33 ± 0.12	0.71 ± 0.19

Table 4.4 summarizes the performance metrics obtained from three representative participants, applying the same procedure across all phonatory tasks. The table presents the mean and standard deviation values for the R, the C_{xy}, and the RMSE, allowing the comparison of the temporal similarity, shared synaptic input, and normalized amplitude error between the iEMG and the estimated CST

signals for each condition. The lowest values in each metric are highlighted in red.

Overall, the results indicate that the correspondence between the CST and iEMG signals improved in the sustained vowel tasks, particularly for the normal- and high-pitch conditions, which exhibited higher correlation and coherence values along with lower RMSE. In contrast, the low-pitch task showed weaker correlations and greater amplitude differences, likely due to lower activation levels of the CT muscle. The pitch-glide task, characterized by dynamic changes in muscle length, presented the highest correlation value ($R = 0.84$) but also a relatively large RMSE, suggesting that while the CST captured the temporal trend of the intramuscular activation, the amplitude variations were less accurately reproduced due to the non-stationary nature of the movement.

4.4 Chapter conclusions

This chapter presented the validation of a dual-source decomposition framework that combined intramuscular and high-density surface EMG recordings of the CT muscle. By estimating the spatial position of each motor unit and classifying them into superficial and deep clusters, the proposed method allowed the separate analysis of neural drive contributions from distinct muscle regions. The CST was computed for each cluster, normalized, and compared to the standard iEMG envelope to assess the feasibility of reconstructing motor-unit activity from

surface recordings.

The validation results demonstrated that the number of simultaneously identified motor units varied across tasks and subjects, with the highest synchronization observed in sustained phonations. The normal- and high-pitch vowels exhibited higher rates of agreement and lower variability compared to the pitch-glide task, which involved continuous changes in CT muscle length. Female participants generally presented fewer simultaneous units, likely due to anatomical factors such as reduced muscle volume and greater fiber depth. These findings confirmed that stable phonatory conditions favor cross-modality synchronization and enhance decomposition reliability.

Although the percentage of common motor units identified between HDsEMG and iEMG was relatively low (approximately 2% on average, as shown in Table 4.2), this does not necessarily imply that HDsEMG recordings lack meaningful CT muscle activity. A key factor underlying this result is the limited spatial sampling of the intramuscular electrode. The hooked-wire electrode records from a highly localized region of the muscle, typically capturing a small subset of motor units relative to the total population within the cricothyroid muscle. In contrast, the HDsEMG electrode array covers a broader surface area, potentially detecting motor unit activity originating from different regions along the muscle. Therefore, motor units identified exclusively in the surface recordings may still correspond to CT activity but remain unobserved in the intramuscular recordings due to

spatial mismatch between recording modalities. This spatial limitation should be considered when interpreting the rate of agreement and supports the notion that HDsEMG may capture a wider representation of muscle activation than that directly validated by the two-source method.

The comparison between iEMG envelopes and CST estimates revealed moderate-to-high correlation and coherence values, particularly in the deep cluster group, indicating that the reconstructed CST signals captured the temporal behavior and shared synaptic input of the CT muscle. The RMSE analysis showed that, although the amplitude reproduction was limited in dynamic tasks, the general trends were preserved. High coherence values below 10 Hz further supported the presence of a common neural drive between both recording modalities, validating the physiological consistency of the proposed decomposition approach.

Overall, the results of this chapter support the feasibility of estimating the collective neural drive of the CT muscle from high-density surface EMG. The methodology provides a bridge between invasive and non-invasive measurements, enabling the investigation of motor-unit synchrony and depth-dependent activation. These findings form the basis for future integration of surface-based CST estimation into biomechanical and acoustic voice models, paving the way for improved non-invasive assessment of laryngeal muscle control during phonation.

5 Muscle activation estimation using networks and HDsEMG

5.1 Introduction

Building upon the findings presented in the previous chapter, where correlation and coherence analyses revealed that surface-derived neural activity (CST) could successfully reproduce the temporal evolution of iEMG but not its amplitude, this chapter explores the use of nonlinear regression techniques to enhance the estimation of normalized laryngeal muscle activation. The previous framework demonstrated that although the temporal correlation between CST and iEMG was strong, the amplitude mismatch, reflected in relatively high RMSE values, suggested that additional acoustic correlates could be incorporated to improve amplitude prediction. In this context, since both the f_0 and the SPL are acoustical parameters directly related to the activation of the CT muscle, they were integrated as complementary inputs to a neural network model designed to better reproduce the amplitude of normalized CT activation.

Accordingly, this chapter introduces a machine-learning framework based on a multilayer perceptron (MLP) architecture trained to map surface EMG-derived and acoustic features to intramuscular activation of the CT muscle. In this initial

implementation, a *subject-specific* approach was adopted, where data from each participant were trained and evaluated independently.

The remainder of this chapter is structured as follows. First, the general framework and configuration of the neural network are presented, describing the data organization, input features, and training procedure. Next, the performance of the proposed model is evaluated through regression metrics and plot to assess its ability to predict normalized CT activation from surface and acoustic inputs. Finally, the main outcomes of the study are discussed, emphasizing their implications for future extensions toward generalized estimation frameworks.

5.1.1 Neural network architecture and training

Figure 5.1 illustrates the structure of the neural network framework implemented in this study. The network was designed to model the relationship between surface-derived neural activity and acoustic parameters to estimate the normalized activation of the CT muscle. Data from three participants were used, each performing four phonatory tasks involving the vowel /a/ at different pitch conditions: low-pitch, normal-pitch, high-pitch, and pitch-glide. For each participant and task, synchronized recordings of HDsEMG, iEMG, and microphone signals were obtained and preprocessed as described in previous sections. This first implementation was performed using a *subject-specific* approach, in which data from each participant were trained and evaluated independently.

Three input features were used: the normalized CST estimated from HD-sEMG decomposition, the f_0, and the SPL. These acoustic inputs were selected due to their high correlation with CT muscle activation. As shown in our results (Figure 3.3, Chapter 3), the fundamental frequency (f_0) increases with an almost linear trend with respect to CT activation, a relationship that is also supported by previous iEMG studies. Sound pressure level (SPL) likewise exhibits a correlation with CT activity, particularly at high loudness levels, as observed in Figure 3.5 for Sessions 1 and 3, where high CT activation was required together with TA activation. Including these features allows the model to capture variability in muscle activation due to different pitch and pressure levels during network training.

The target output corresponded to the normalized iEMG envelope of the CT muscle (a'_{CT}). The acoustic features f_0 and SPL were preprocessed following the procedure described in Section 3.2.1, while the intramuscular envelope was obtained as detailed in Section 3.2.2. Once the four signals were synchronized, they were segmented into 50 ms windows with a 50% overlap. For each window, the mean value was extracted to generate the muscle activation tokens used for model training.

Prior to training, all input variables were standardized using z-score normalization, while the target output was normalized to the range [0 - 1] [120]. The dataset from each subject was randomly partitioned into training (70 %), validation (15 %), and testing (15 %) subsets to ensure representative coverage of all

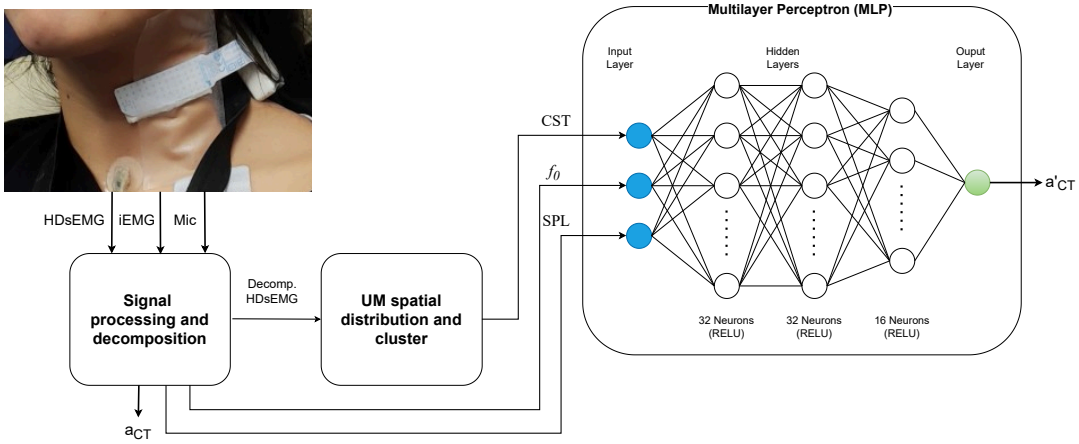


Figure 5.1: Scheme of the Neural network framework procedure

phonatory tasks.

A MLP architecture was implemented [120] with three hidden layers containing 32, 32, and 16 neurons, respectively (see Figure 5.1). The $ReLU$ activation function was applied to all hidden layers, while the output layer employed a linear activation to reproduce continuous amplitude values [121]. The network parameters were optimized using the Adam algorithm [122] with a learning rate of 0.001 and a mini-batch size of 128 samples. Training was carried out for 60 epochs, using the MSE as the loss function. Early stopping criteria based on validation loss were used to prevent overfitting. The trained models were evaluated using the value R, MAE, and the RMSE between predicted and true normalized activations.

5.2 Results

5.2.1 Results

Figure 5.2 shows the evolution of the MSE for training and validation. A rapid decrease followed by stabilization was observed in both curves, indicating adequate convergence of the model without signs of overfitting.

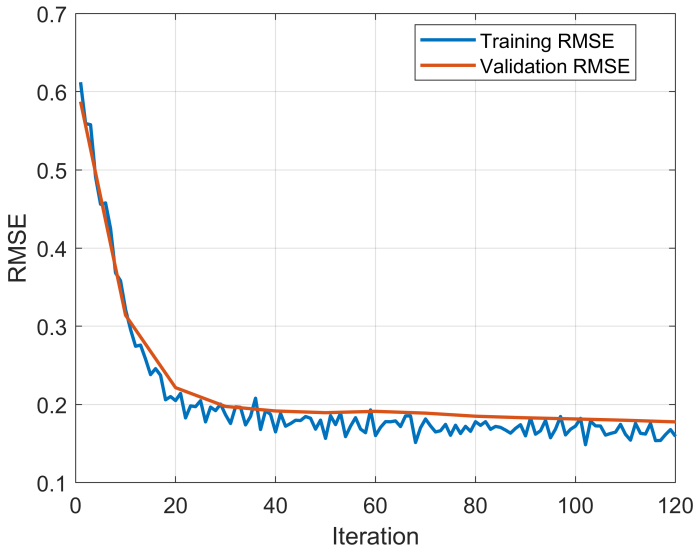


Figure 5.2: MSE evolution across epochs for training (blue) and validation (red) stages in the proposed multilayer perceptron.

Figure 5.4 presents the scatter plots comparing the predicted versus true normalized a_{CT} activations for three participants (CM006, CM008, and CM009). The dashed blue line represents the unity correspondence between the predicted and reference intramuscular envelopes. The value R ranged between 0.81 and 0.94

across participants, suggesting a strong linear relationship between the estimated and true activations. The MAE and RMSE values are summarized in Table 5.1. The lowest prediction error was obtained for participant CM006 ($R = 0.94$, RMSE = 0.092), while CM009 exhibited slightly higher dispersion around the unity line ($R = 0.82$, RMSE = 0.163).

Overall, the MLP model effectively reproduced the temporal and amplitude behavior of the intramuscular activation envelope based on surface-derived CST and acoustic features (f_0 and SPL). Despite the promising estimation results, it is important to note that the current model was trained on a limited number of windowed samples derived from only four phonatory tasks, which prevents generalization at this stage.

Table 5.1: Performance metrics of the MLP model for each participant. Results correspond to the R, MAE, and RMSE.

Participant	**R**	**MAE**	**RMSE**
CM006	0.936	0.071	0.092
CM008	0.838	0.103	0.142
CM009	0.815	0.121	0.163

The performance of the network was then evaluated using the complete dataset formed by merging all samples from the three participants into a single multi-

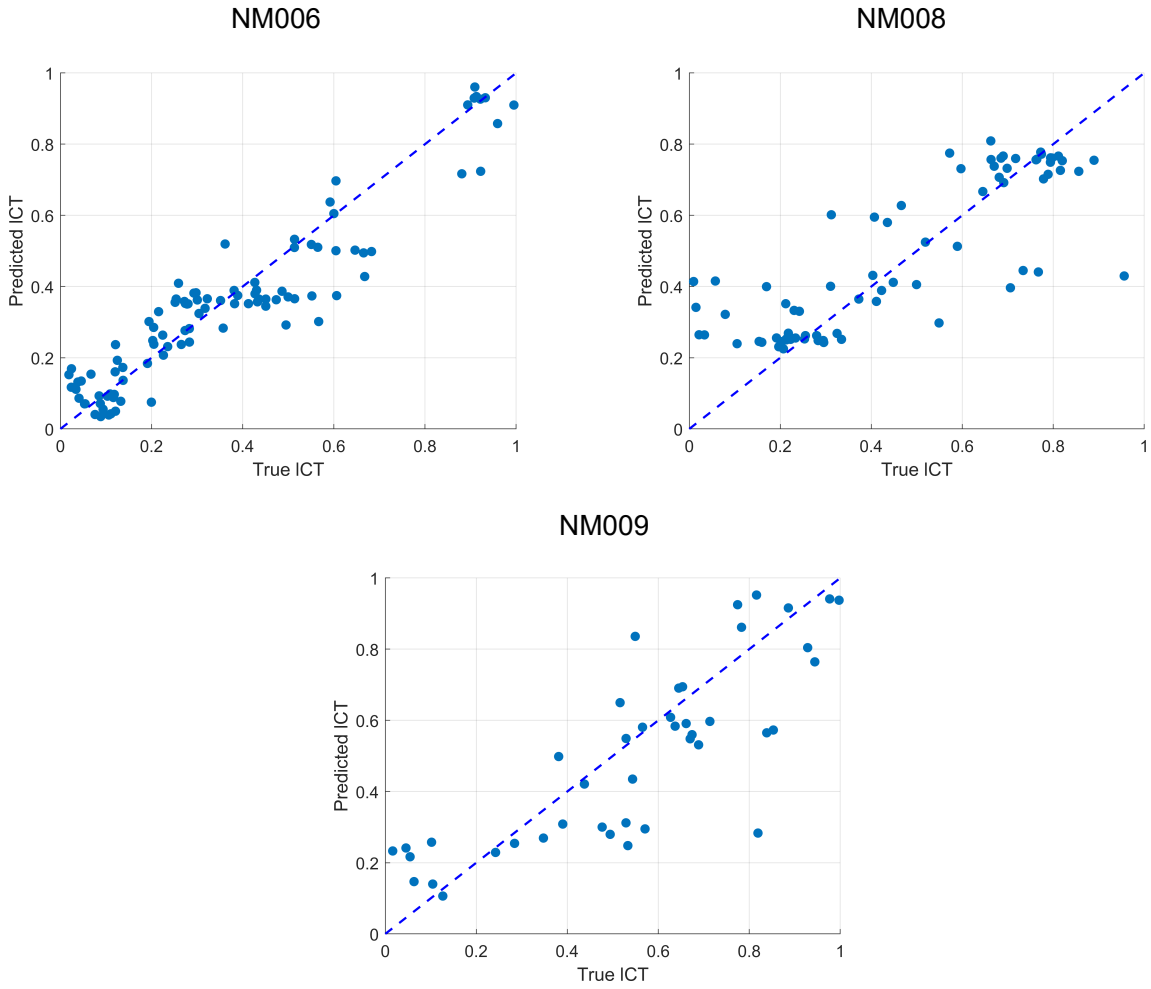


Figure 5.3: Scatter plots comparing the predicted versus true normalized CT activation for three representative participants (CM006, CM008, and CM009). Each point represents the mean activation value obtained from a 50 ms window with 50% overlap. The dashed line indicates the 1:1 correspondence between predicted and true values.

subject database. Table 5.2 summarizes the performance of the multi-subject model trained on this combined dataset. The resulting metrics indicate that the network was able to reproduce the normalized CT activation with a high degree

of correspondence, as reflected by an overall correlation coefficient of $R = 0.84$, together with low absolute and quadratic errors (MAE = 0.103, RMSE = 0.141). These values suggest that the joint use of CST, f_0, and SPL provides sufficient information for the model to generalize the relationship between surface-derived activation and intramuscular activation across different subjects. Figure 5.4 illustrates the relationship between predicted and true activation values for the multi-subject configuration. Each point corresponds to the mean activation value extracted from a 50 ms window. A clear positive trend is observed, with most predictions distributed around the unity line.

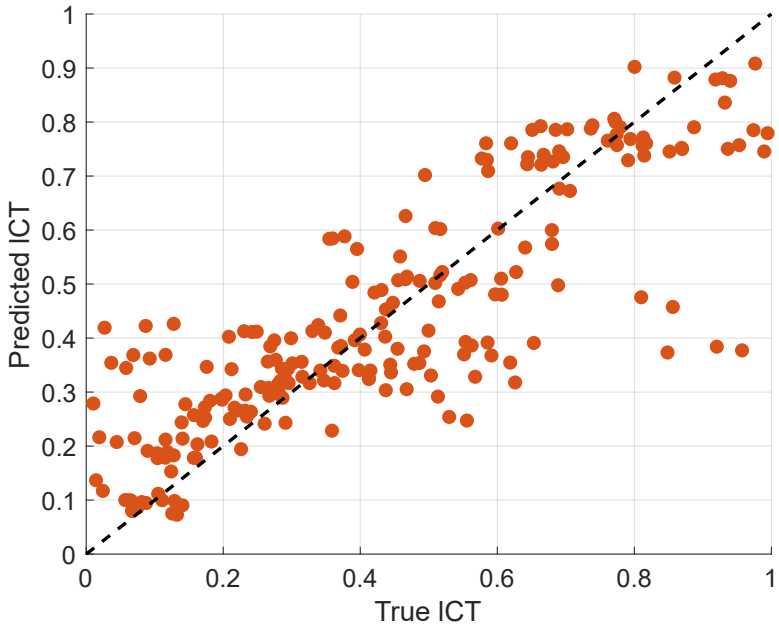


Figure 5.4: Predicted vs. true CT activation for the multi-subject MLP model.

The performance obtained in this work is consistent with previous subject-

specific neural network approaches reported by Ibarra et al. [123],in which high prediction accuracy for subglottal pressure and intrinsic laryngeal muscle activation was achieved, with R^2 values up to 0.95 and low RMSE values. In comparison, the multi-subject model proposed in this study (Figure 2) yielded a correlation coefficient of $R = 0.84$ and an RMSE of 0.14. It is important to note that the approach by Ibarra et al. [123] relied on a larger set of input features, including seven physiological and acoustic variables and direct airflow measurements obtained using a Rothenberg mask, whereas the proposed framework uses only two acoustic features extracted from a microphone recording. Despite this reduced input space, the obtained metrics demonstrate that the proposed approach preserves a substantial level of accuracy, supporting its applicability in scenarios where minimal instrumentation and ambulatory use are required.

Table 5.2: Performance metrics of the multi-subject model.

	R	**MAE**	**RMSE**
Multi-subject	0.841215	0.102942	0.141251

Table 5.3 presents a comparison of the different MLP architectures evaluated in this work. Increasing network depth produced only moderate improvements, with models containing four and five hidden layers achieving slightly higher correlation values and lower error measures. Although the relative differences across configurations were not substantial, the results show that increasing model ca-

pacity contributes to a somewhat more accurate representation of the nonlinear relationship between surface-derived features and normalized intramuscular activation. It can be observed that beyond four hidden layers the improvement becomes negligible, with the best results obtained around this configuration. The RMSE value could not be reduced below 0.13543, which highlights a remaining gap between the estimated and true activation values and reflects the challenge of further decreasing this error.

Table 5.3: Performance comparison for different network architectures.

Architecture	R	**MAE**	**RMSE**
2 Hidden Layers	0.82568	0.10578	0.14660
3 Hidden Layers	0.83630	0.09948	0.14229
4 Hidden Layers	0.85679	0.09444	0.13434
5 Hidden Layers	0.85289	0.09269	0.13543

5.3 Chapter Conclusions

This chapter presented a neural network framework for estimating the normalized activation of the CT muscle using surface EMG-derived and acoustic features.

The proposed MLP architecture successfully integrated the CST, f_0, and SPL to predict intramuscular activation. The selection of these features is supported by their high correlation with CT muscle activation, as demonstrated in Figures 3 and 5 of Section 2 and in previous iEMG studies of the CT muscle. By incorporating both acoustic features, changes in muscle activation associated with variations in pressure and pitch during speech can be captured, while also overcoming the amplitude-related limitations identified in the previous chapter. In addition, the use of features that only require microphone recordings enables ambulatory applications, in contrast to other approaches that rely on a larger number of phonatory features and require the use of a Rothenberg mask in laboratory settings.

The subject-specific training strategy yielded high correlation values and low error metrics across participants, confirming that individual modeling enhances the reliability of EMG-to-activation mapping. The MSE evolution showed stable convergence, indicating that the network generalized well to unseen data without signs of overfitting. These findings validate the feasibility of using high-density surface EMG and acoustic cues as complementary information sources for the estimation of intrinsic laryngeal muscle activity.

Previous subject-specific neural network approaches for estimating CT muscle activation, such as the work by Ibarra et al., [123] have reported higher performance, with larger correlation coefficients and lower RMSE values. However, these results were obtained using a larger number of input features. Therefore,

future work should evaluate the incorporation of additional inputs to approach these performance levels while preserving the simplicity and practicality of the proposed framework.

The multi-subject evaluation further demonstrated that the model is capable of generalizing the mapping across participants, although the improvements associated with increasing model depth were modest. In particular, architectures with four or more hidden layers produced only marginal gains, and the RMSE could not be reduced below 0.13543. This result indicates a persistent gap between the estimated and true activation values, suggesting an inherent limitation in the current representation of the nonlinear relationship between the surface features and the intramuscular activation.

Despite the promising performance, it is important to note that the present results are based on a limited dataset, consisting of three subjects and four phonatory tasks. Therefore, the generalization of the model across larger populations and more diverse voice conditions remains to be evaluated. Future work should explore cross-subject frameworks and transfer learning approaches to reduce inter-subject variability and extend the estimation to additional muscles beyond the CT. The framework establishes a foundation for future developments in surface-based muscle activity modeling, with potential applications in voice physiology, diagnostics, and rehabilitation.

6 Conclusions

This thesis had as its purpose to investigate the feasibility of estimating intrinsic laryngeal muscle activity by means of HDsEMG, at least of the CT muscle. The general objective was to establish a non-invasive framework capable of capturing physiologically significant muscle activation patterns that are traditionally obtained through invasive iEMG recordings. Through the integration of HDsEMG signal and machine learning techniques, this work provides a methodology to advance laryngeal muscle evaluation.

In Chapter 2, the physiological fundamentals of muscle activation and motor unit behavior were reviewed, together with existing approaches to estimate intrinsic laryngeal muscle activity from EMG signals. It started with the relevance of motor unit recruitment and firing rates as fundamental factors for the generation of muscle force and neural drive. This chapter established the theoretical bases to link EMG-derived features with physiological models of voice production and highlighted the limitations of conventional surface EMG techniques when applied to small and deep laryngeal muscles. Other estimation works using different methods, such as neural networks, were also addressed, being possible topics to be tackled in the future.

Chapter 3 focused on the characterization of intrinsic laryngeal muscle activity through intramuscular EMG recordings, which served as a physiological

reference throughout this thesis. An acoustic normalization framework using f_0 and SPL was proposed to reduce inter-session and inter-task variability. The results demonstrated that acoustic normalization improves intermuscular and inter-session consistency, allowing meaningful comparisons of muscle activation across different phonatory tasks. These normalized iEMG signals constituted the basis to validate subsequent HDsEMG-based methods. The results were also expanded to new patients, allowing to observe the robustness of the method, by observing values of approximately 0.4 for frequencies of 150 Hz in the 3 patients. It is known that there is much pending work with respect to normalization in women (CF002), which shows a much larger frequency range, therefore it is a pending topic that will be addressed in the future.

In Chapter 4, a decomposition-based framework was proposed to estimate intrinsic muscle activation from HDsEMG recordings. Motor unit activation activity was extracted using established decomposition algorithms and validated through a two-source approach with simultaneously recorded intramuscular EMG. When observing the validation data, it is demonstrated that part of the CT activation is reaching the surface, allowing to demonstrate the focus of the second objective. The spatial distribution of motor units was analyzed through CoG estimation and depth modeling, which allowed identifying superficial and deep activation groups within the CT muscle. CSTs were generated for each group to represent their respective neural drives. Correlation and coherence analyses demonstrated that

CSTs derived from deeper motor unit groups showed greater synchronization with iEMG signals, supporting the physiological relevance of the proposed approach.

The relationship between motor unit depth and CST synchronization is based on the assumption that deeper motor units are more likely to originate from the target intrinsic muscle (i.e., the CT), whereas more superficial sources may include contributions from overlying muscles such as strap muscles or platysma. As a result, CSTs derived from deeper clusters are expected to better reflect the neural drive of the CT muscle and therefore exhibit higher synchronization with intramuscular EMG signals. In our results, CSTs associated with deeper activation groups showed consistently higher correlation and coherence with iEMG recordings compared to more superficial groups. This supports the physiological relevance of the proposed spatial decomposition approach, as it suggests that the depth estimation and clustering procedure can differentiate motor unit populations that are more closely related to the target muscle.

However, the results of Chapter 4 also revealed important limitations. While CSTs were able to reproduce the temporal trends of intramuscular activation, discrepancies in amplitude were consistently observed, as reflected by RMSE values, mainly in tasks such as pitch glides and low pitch. These findings indicated that, although HDsEMG decomposition effectively captures neural synchronization information, it may not be sufficient to fully reproduce the normalized amplitude of muscle activation using CSTs alone. This limitation motivated the exploration

of neural network-based strategies in the following chapter.

Chapter 5 addressed this limitation by incorporating neural network models to estimate normalized CT activation. A subject-specific MLP was trained using CSTs together with acoustic features (f_0 and SPL) as inputs, and intramuscular EMG envelopes as the target output. The results demonstrated that the inclusion of acoustic correlates significantly improved the estimation of muscle activation amplitude, as reflected by reduced MAE and RMSE values and high R values. The multi-subject model further confirmed that the combined feature set provides sufficient information for the network to generalize across participants. In comparison with previous subject-specific approaches, such as the work by Ibarra et al. [123], where higher performance metrics were obtained using a larger number of input features and additional instrumentation, the present framework achieves comparable trends while relying only on two acoustic features extracted from a microphone recording. This result points to a clear direction for future work, in which the exploration of additional features could further improve performance. However, further validation is required through the analysis of more diverse datasets and larger populations.

Overall, the objectives of this thesis were achieved, and the proposed hypotheses were validated through the experimental and modeling results presented across the different chapters. Specific Aim 1 was fulfilled by developing an acoustic-based normalization framework that reduced inter-session variability in intramuscular

EMG recordings and enabled consistent muscle–acoustic mappings across phonatory tasks. Specific Aim 2 and 3 was also achieved by demonstrating that HDsEMG decomposition, validated through a two-source approach with simultaneous iEMG, can recover physiologically meaningful temporal and spatial components associated with intrinsic laryngeal muscle activity.

Regarding the hypotheses, Hypotheses 1 and 2 were supported by the results, confirming improved normalization consistency, as reflected by reduced CV and MAE metrics, and a measurable correspondence between surface and intramuscular motor unit discharge patterns, evidenced by ROA values greater than 25% in the majority of subjects, as shown in Table 4.2. Hypothesis 3 was only partially supported: while HDsEMG-based estimates successfully reproduced the temporal trends observed in intramuscular CT activity, the resulting RMSE values indicated notable amplitude discrepancies, particularly during dynamic phonatory conditions. In contrast, Hypothesis 4 was supported by the results obtained using subject-specific neural network models, which achieved acceptable RMSE and R values that were comparable to those reported in previous studies.

Despite these encouraging results, several limitations of the proposed framework must be acknowledged. First, intramuscular decomposition did not always achieve optimal performance during phonatory tasks characterized by high levels of motor unit superposition, such as high-pitch phonation. Under these conditions, strong CT muscle activation led to increased overlap of motor unit action

potentials, limiting decomposition accuracy. This limitation is partly due to the use of a single intramuscular recording channel, which restricts the spatial resolution available to separate overlapping motor unit activity. Second, the EMG decomposition methods employed in this work showed sensitivity to noise, which negatively affected performance under poor recording conditions. Factors such as motion artifacts, electromagnetic interference, and variations in electrode–skin impedance can degrade signal quality and reduce the reliability of motor unit identification. This sensitivity underscores the need for robust preprocessing strategies and improved hardware configurations in future studies.

Third, anatomical variability among participants influenced the quality of HD-sEMG recordings. In particular, subcutaneous adipose tissue in the neck region acted as a low-pass spatial filter, attenuating high-frequency components of the EMG signal. This effect was more pronounced in female participants and in men with higher fat levels, who tend to present higher levels of cervical adiposity due to genetic and anatomical factors. As a result, decomposition performance and spatial localization accuracy were reduced in these cases, highlighting an important limitation of surface-based approaches in laryngeal EMG assessment. Another limitation arises from the spatial coverage of the electrode grid used in this study. While high-density arrays provide improved spatial resolution compared to conventional surface electrodes, the grid employed was not large enough to capture the full range of laryngeal movement in all phonatory tasks. Under certain pitch

conditions, the CT muscle likely moved partially outside the electrode coverage area, leading to incomplete sampling of muscle activity. This limitation suggests that larger or adaptive electrode configurations may be required to fully monitor laryngeal dynamics.

Future work should address these limitations through several complementary lines. Expanding the study to include pathological populations, such as individuals with vocal hyperfunction, unilateral vocal fold paralysis, or other laryngeal disorders, would allow evaluation of whether the proposed methods can capture clinically relevant differences in muscle activation patterns. Such studies would also provide insight into the robustness of HDsEMG decomposition under altered neuromuscular conditions. In addition, increasing the number of recording channels, both intramuscular and surface, could substantially improve decomposition accuracy and spatial resolution. Multichannel intramuscular recordings would allow better separation of overlapping motor units, especially during high-activation tasks. Similarly, larger or multi-grid HDsEMG configurations could improve coverage of laryngeal muscles during dynamic phonatory conditions. Future investigations should also explore simultaneous intramuscular recordings of multiple intrinsic laryngeal muscles, such as the TA, to extend the proposed framework beyond the CT muscle. This would enable the study of intermuscular coordination and the development of multimuscle activation models, which are fundamental to understanding the complex mechanisms of voice production.

Finally, the development of computational models capable of simulating HD-sEMG signals of laryngeal muscles represents a promising future research direction. Such models could be used to validate decomposition algorithms, optimize electrode configurations, and explore the effects of anatomical variability on surface recordings. The integration of these models with machine learning frameworks may further enhance the interpretability and generalization of non-invasive estimates of laryngeal muscle activation.

Bibliography

[1] Orhan Asya, Ömer Tarık Kavak, Hatice Ömercikoğlu Özden, Dilek Günal, and Necati Enver. Demographic and clinical characteristics of our patients diagnosed with laryngeal dystonia. *European Archives of Oto-Rhino-Laryngology*, 281(8):4265–4271, 2024.

[2] Seyyedeh Maryam Khoddami, Noureddin Nakhostin Ansari, Farzad Izadi, and Saeed Talebian Moghadam. The assessment methods of laryngeal muscle activity in muscle tension dysphonia: a review. *The Scientific World Journal*, 2013(1):507397, 2013.

[3] Robert T Sataloff, Steven Mandel, Eric A Mann, and Christy L Ludlow. Practice parameter: laryngeal electromyography (an evidence-based review). *Otolaryngology–Head and Neck Surgery*, 130(6):770–779, 2004.

[4] Emiro J. Ibarra, Julián D. Arias-Londoño, Juan I. Godino-Llorente, Daryush D. Mehta, and Matías Zañartu. Subject-specific modeling by domain adaptation for the estimation of subglottal pressure from neck-surface acceleration signals. *Biomedical Signal Processing and Control*, 106:107681, 2025.

[5] Robert T Sataloff, Phurich Praneetvatakul, Reinhardt J Heuer, Mary J Hawkshaw, Yolanda D Heman-Ackah, Sarah Marx Schneider, and Steven

Mandel. Laryngeal electromyography: clinical application. *Journal of Voice*, 24(2):228–234, 2010.

[6] Gerd Fabian Volk, Rudolf Hagen, Claus Pototschnig, Gerhard Friedrich, Tadeus Nawka, Christoph Arens, Andreas Mueller, Gerhard Foerster, Mira Finkensieper, Ruth Lang-Roth, Christian Sittel, Claudio Storck, Maria Grosheva, M. Nasser Kotby, Carsten M. Klingner, and Orlando Guntinas-Lichius. Laryngeal electromyography: a proposal for guidelines of the European Laryngological Society. *European Archives of Oto-Rhino-Laryngology*, 269(10):2227–2245, October 2012.

[7] Christy L Ludlow. Central nervous system control of the laryngeal muscles in humans. *Respiratory physiology & neurobiology*, 147(2-3):205–222, 2005.

[8] Thomas Gay, Marshall Strome, Hajime Hirose, and Masayuki Sawashima. Electromyography of the Intrinsic Laryngeal Muscles during Phonation. *Annals of Otology, Rhinology & Laryngology*, 81(3):401–409, June 1972. Publisher: SAGE Publications Inc.

[9] Ingo R. Titze, Erich S. Luschei, and Minoru Hirano. Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice*, 3(3):213–224, September 1989.

[10] Anne Tabard-Fougère, Kevin Rose-Dulcina, Vincent Pittet, Romain Dayer,

Nicolas Vuillerme, and Stéphane Armand. Emg normalization method based on grade 3 of manual muscle testing: Within-and between-day reliability of normalization tasks and application to gait analysis. *Gait & posture*, 60:6–12, 2018.

[11] Adrian Burden. How should we normalize electromyograms obtained from healthy participants? What we have learned from over 25 years of research. *Journal of electromyography and kinesiology*, 20(6):1023–1035, 2010.

[12] Cara E Stepp, Robert E Hillman, and James T Heaton. Use of neck strap muscle intermuscular coherence as an indicator of vocal hyperfunction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):329–335, 2010.

[13] Evelyne Van Houtte, Sofie Claeys, Evelien D'haeseleer, Floris Wuyts, and Kristiane Van Lierde. An examination of surface emg for the assessment of muscle tension dysphonia. *Journal of Voice*, 27(2):177–186, 2013.

[14] Roberto Merletti and PJJEK Di Torino. Standards for reporting emg data. *Journal of Electromyography and Kinesiology*, 9(1):3–4, 1999.

[15] Roberto Merletti and Silvia Muceli. Tutorial. surface emg detection in space and time: Best practices. *Journal of Electromyography and Kinesiology*, 49:102363, 2019.

[16] Andrew J Fuglevand, DAVID A Winter, and AFTAB E Patla. Models of recruitment and rate coding organization in motor-unit pools. *Journal of neurophysiology*, 70(6):2470–2488, 1993.

[17] Rositsa Raikova, Jan Celichowski, Magdalena Pogrzebna, Hristo Aladjov, and Piotr Krutki. Modeling of summation of individual twitches into unfused tetanus for various types of rat motor units. *Journal of Electromyography and Kinesiology*, 17(2):121–130, 2007.

[18] Eike Petersen and Philipp Rostalski. A comprehensive mathematical model of motor unit pool organization, surface electromyography, and force generation. *Frontiers in physiology*, 10:176, 2019.

[19] Rositsa Raikova, Jan Celichowski, Silvija Angelova, and Piotr Krutki. A model of the rat medial gastrocnemius muscle based on inputs to motoneurons and on an algorithm for prediction of the motor unit force. *Journal of neurophysiology*, 120(4):1973–1987, 2018.

[20] Jim R Potvin and Andrew J Fuglevand. A motor unit-based model of muscle fatigue. *PLoS computational biology*, 13(6):e1005581, 2017.

[21] Rodrigo Manriquez, Sean D. Peterson, Pavel Prado, Patricio Orio, Gabriel E. Galindo, and Matías Zañartu. Neurophysiological Muscle Activation Scheme for Controlling Vocal Fold Models. *IEEE transactions on*

neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society, 27(5):1043–1052, May 2019.

[22] RE Burke and P Tsairis. Anatomy and innervation ratios in motor units of cat gastrocnemius. *The Journal of physiology*, 234(3):749–765, 1973.

[23] Ch K Thomas, BRUCE H Ross, and BLAIR Calancie. Human motor-unit recruitment during isometric contractions and repeated dynamic movements. *journal of Neurophysiology*, 57(1):311–324, 1987.

[24] Francesco Negro and Claudio Orizio. Robust estimation of average twitch contraction forces of populations of motor units in humans. *Journal of Electromyography and Kinesiology*, 37:132–140, 2017.

[25] HS Milner-Brown, Richard B Stein, and R Yemm. Changes in firing rate of human motor units during linearly changing voluntary contractions. *The Journal of physiology*, 230(2):371, 1973.

[26] 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. *Journal of Electromyography and Kinesiology*, 18(5):741–751, 2008.

[27] Rositsa Raikova, Piotr Krutki, and Jan Celichowski. Skeletal muscle mod-

els composed of motor units: A review. *Journal of Electromyography and Kinesiology*, 70:102774, 2023.

[28] CJ De Luca, RS LeFever, MP McCue, and AP Xenakis. Behaviour of human motor units in different muscles during linearly varying contractions. *The Journal of physiology*, 329(1):113–128, 1982.

[29] Paola Contessa and Carlo J De Luca. Neural control of muscle force: indications from a simulation model. *Journal of neurophysiology*, 109(6):1548–1570, 2013.

[30] Hong-Shik Choi, Gerald S. Berke, Ming Ye, and Jody Kreiman. Function of the posterior cricoarytenoid muscle in phonation: In vivo laryngeal model. *Otolaryngology–Head and Neck Surgery*, 109(6):1043–1051, 1993.

[31] Karin Cox, Fariborz Alipour, and Ingo Titze. Geometric Structure of the Human and Canine Cricothyroid and Thyroarytenoid Muscles for Biomechanical Applications. *The Annals of otology, rhinology, and laryngology*, 108(12):1151–8, December 1999.

[32] Andrew M. Vahabzadeh-Hagh, Zhaoyan Zhang, and Dinesh K. Chhetri. Hirano's cover–body model and its unique laryngeal postures revisited. *The Laryngoscope*, 128(6):1412–1418, 2018.

[33] Dinesh K. Chhetri, Juergen Neubauer, and David A. Berry. Neuromuscular

control of fundamental frequency and glottal posture at phonation onset. *The Journal of the Acoustical Society of America*, 131(2):1401–1412, February 2012.

[34] Dinesh K. Chhetri, Juergen Neubauer, Elazar Sofer, and David A. Berry. Influence and interactions of laryngeal adductors and cricothyroid muscles on fundamental frequency and glottal posture control. *The Journal of the Acoustical Society of America*, 135(4):2052–2064, April 2014.

[35] Dinesh K. Chhetri and Juergen Neubauer. Differential Roles for the Thyroarytenoid and Lateral Cricoarytenoid Muscles in Phonation. *The Laryngoscope*, 125(12):2772, July 2015.

[36] Dinesh K. Chhetri and Soo Jin Park. Interactions of subglottal pressure and neuromuscular activation on fundamental frequency and intensity. *The Laryngoscope*, 126(5):1123–1130, May 2016.

[37] Minoru Hirano, John Ohala, and William Vennard. The Function of Laryngeal Muscles in Regulating Fundamental Frequency and Intensity of Phonation. *Journal of Speech and Hearing Research*, 12(3):616–628, September 1969. Publisher: American Speech-Language-Hearing Association.

[38] Allen D. Hillel. The Study of Laryngeal Muscle Activity in Normal Human

Subjects and in Patients With Laryngeal Dystonia Using Multiple Fine-Wire Electromyography. *The Laryngoscope*, 111(S97):1–47, 2001.

[39] Christopher J. Poletto, Laura P. Verdun, Robert Strominger, and Christy L. Ludlow. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *Journal of Applied Physiology*, 97(3):858–866, September 2004. Publisher: American Physiological Society.

[40] Bernard Roubeau, Claude Chevrie-Muller, and Jean Lacau Saint Guily. Electromyographic activity of strap and cricothyroid muscles in pitch change. *Acta oto-laryngologica*, 117(3):459–464, 1997.

[41] Karen Ann Kochis-Jennings, Eileen M Finnegan, Henry T Hoffman, and Sanyukta Jaiswal. Laryngeal muscle activity and vocal fold adduction during chest, chestmix, headmix, and head registers in females. *Journal of Voice*, 26(2):182–193, 2012.

[42] Eileen M Finnegan, Erich S Luschei, and Henry T Hoffman. Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *Journal of Speech, Language, and Hearing Research*, 43(4):934–950, 2000.

[43] Minoru Hirano, William Vennard, and John Ohala. Regulation of register,

pitch and intensity of voice: An electromyographic investigation of intrinsic laryngeal muscles. *Folia phoniatrica et logopaedica*, 22(1):1–20, 1970.

[44] Anders Löfqvist, Nancy S McGarr, and Kiyoshi Honda. Laryngeal muscles and articulatory control. *The Journal of the Acoustical Society of America*, 76(3):951–954, 1984.

[45] Robert E Hillman, Eva B Holmberg, Joseph S Perkell, Michael Walsh, and Charles Vaughan. Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech, Language, and Hearing Research*, 32(2):373–392, 1989.

[46] Margaret A Redenbaugh and Alan R Reich. Surface emg and related measures in normal and vocally hyperfunctional speakers. *Journal of Speech and Hearing Disorders*, 54(1):68–73, 1989.

[47] I Hocevar-Boltezar, Martin Janko, and M Zargi. Role of surface emg in diagnostics and treatment of muscle tension dysphonia. *Acta oto-laryngologica*, 118(5):739–743, 1998.

[48] Murray D Morrison, Linda A Rammage, Gilles M Belisle, C Bruce Pullan, and Hamish Nichol. Muscular tension dysphonia. *The Journal of otolaryngology*, 12(5):302–306, 1983.

[49] Evelyne Van Houtte, Kristiane Van Lierde, and Sofie Claeys. Pathophysi-

ology and treatment of muscle tension dysphonia: a review of the current knowledge. *Journal of Voice*, 25(2):202–207, 2011.

[50] Cara E Stepp, James T Heaton, Tara K Stadelman-Cohen, Maia N Braden, Marie E Jetté, and Robert E Hillman. Characteristics of phonatory function in singers and nonsingers with vocal fold nodules. *Journal of voice*, 25(6):714–724, 2011.

[51] Patricia Maria Mendes Balata, Hilton Justino Silva, Leandro Araújo Pernambuco, Geová Oliveira Amorim, Renata Souto Maior Braga, Elthon Gomes Fernandes da Silva, Leilane Maria de Lima, and Sílvia Regina Arruda Moraes. Electrical activity of extrinsic laryngeal muscles in subjects with and without dysphonia. *Journal of voice*, 29(1):129–e9, 2015.

[52] Ruiying Ding, Charles R Larson, Jeri A Logemann, and Alfred W Rademaker. Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: normal and during the mendelsohn manuever. *Dysphagia*, 17(1):1–12, 2002.

[53] Mingxing Zhu, Wanzhang Yang, Oluwarotimi Williams Samuel, Yun Xiang, Jianping Huang, Haiqing Zou, and Guanglin Li. A preliminary evaluation of myoelectrical energy distribution of the front neck muscles in pharyngeal phase during normal swallowing. In *2016 38th Annual International Con-*

ference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 1700–1703. IEEE, 2016.

[54] Gea Drost, Dick F Stegeman, Baziel GM van Engelen, and Machiel J Zwarts. Clinical applications of high-density surface emg: a systematic review. *Journal of Electromyography and Kinesiology*, 16(6):586–602, 2006.

[55] David J Bracken, Gladys Ornelas, Todd P Coleman, and Philip A Weissbrod. High-density surface electromyography: A visualization method of laryngeal muscle activity. *The Laryngoscope*, 129(10):2347–2353, 2019.

[56] Mingxing Zhu, Faya Liang, Oluwarotimi Williams Samuel, Shixiong Chen, Wanzhang Yang, Lin Lu, Haiqing Zou, Peng Li, and Guanglin Li. A pilot study on the evaluation of normal phonating function based on high-density semg topographic maps. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1030–1033. IEEE, 2017.

[57] Mingxing Zhu, Lin Lu, Zijian Yang, Xin Wang, Zhenzhen Liu, Wenhao Wei, Fei Chen, Peng Li, Shixiong Chen, and Guanglin Li. Contraction patterns of neck muscles during phonating by high-density surface electromyography. In *2018 IEEE international conference on cyborg and bionic systems (CBS)*, pages 572–575. IEEE, 2018.

[58] K Roeleveld, DF Stegeman, HM Vingerhoets, and A Van Oosterom. The motor unit potential distribution over the skin surface and its use in estimating the motor unit location. *Acta physiologica scandinavica*, 161(4):465–472, 1997.

[59] Yasuhiro Nakajima, Saran Keeratihattayakorn, Satoshi Yoshinari, and Shigeru Tadano. An emg-ct method using multiple surface electrodes in the forearm. *Journal of Electromyography and Kinesiology*, 24(6):875–880, 2014.

[60] Enrico Piovanelli, Davide Piovesan, Shouhei Shirafuji, and Jun Ota. Estimating deep muscles activation from high density surface emg using graph theory. In *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, pages 405–410. IEEE, 2019.

[61] Marco Carbonaro, Kristen M Meiburger, Silvia Seoni, Emma F Hodson-Tole, Taian Vieira, and Alberto Botter. Physical and electrophysiological motor unit characteristics are revealed with simultaneous high-density electromyography and ultrafast ultrasound imaging. *Scientific Reports*, 12(1):8855, 2022.

[62] Luca Mesin. Real time identification of active regions in muscles from high density surface electromyogram. *Computers in biology and medicine*, 56:37–50, 2015.

[63] Brad H. Story. An overview of the physiology, physics and modeling of the sound source for vowels. *Acoustical Science and Technology*, 23(4):195–206, 2002.

[64] Tomáš Vampola, Jaromír Horáček, and Ivo Klepáček. Computer simulation of mucosal waves on vibrating human vocal folds. *Biocybernetics and Biomedical Engineering*, 36(3):451–465, 2016.

[65] Michael Döllinger, Zhaoyan Zhang, Stefan Schoder, Petr Šidlof, Boğaç Tur, and Stefan Kniesburges. Overview on state-of-the-art numerical modeling of the phonation process. *Acta Acustica*, 7, 2023.

[66] K. Ishizaka and J. L. Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *The Bell System Technical Journal*, 51(6):1233–1268, 1972.

[67] Brad H. Story and Ingo R. Titze. Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America*, 97(2):1249–1260, 1995.

[68] Matías Zañartu, Luc Mongeau, and George R. Wodicka. Influence of acoustic loading on an effective single mass model of the vocal folds. *The Journal of the Acoustical Society of America*, 121(2):1119–1129, 2007.

[69] Peter Birkholz, Bernd Kröger, and Christiane Neuschaefer-Rube. Articu-

latory synthesis of words in six voice qualities using a modified two-mass model of the vocal folds. 370, 2011.

[70] Matías Zañartu, Gabriel E. Galindo, Byron D. Erath, Sean D. Peterson, George R. Wodicka, and Robert E. Hillman. Modeling the effects of a posterior glottal opening on vocal fold dynamics with implications for vocal hyperfunctiona). *The Journal of the Acoustical Society of America*, 136(6):3262–3271, December 2014.

[71] Ingo R. Titze and Brad H. Story. Rules for controlling low-dimensional vocal fold models with muscle activation. *The Journal of the Acoustical Society of America*, 112(3):1064–1076, September 2002.

[72] Jaromír Horáček, Anne Maria Laukkanen, and Petr Šidlof. Estimation of impact stress using an aeroelastic model of voice production. *Logopedics Phoniatrics Vocology*, 32(4):185–192, 2007.

[73] Gabriel E. Galindo, Sean D. Peterson, Byron D. Erath, Christian Castro, Robert E. Hillman, and Matías Zañartu. Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds. *Journal of Speech, Language, and Hearing Research*, 60(9):2452–2471, 2017.

[74] Ingo R. Titze and Eric J. Hunter. A two-dimensional biomechanical model

of vocal fold posturing. *The Journal of the Acoustical Society of America*, 121(4):2254–2260, April 2007.

[75] Gabriel A. Alzamendi, Sean D. Peterson, Byron D. Erath, Robert E. Hillman, and Matías Zañartu. Triangular body-cover model of the vocal folds with coordinated activation of the five intrinsic laryngeal muscles. *The Journal of the Acoustical Society of America*, 151(1):17–30, January 2022.

[76] Fariborz Alipour-Haghighi and Ingo R Titze. Viscoelastic modeling of canine vocalis muscle in relaxation. *The Journal of the Acoustical Society of America*, 78(6):1939–1943, 1985.

[77] Peter Birkholz, Bernd J Kröger, and Christiane Neuschaefer-Rube. Synthesis of breathy, normal, and pressed phonation using a two-mass model with a triangular glottis. In *Interspeech 2011: 12th Annual Conference of the International Speech Communi- cation Association*, pages 2681–2684, 2011. Florence, Italy.

[78] Robert E. Hillman, Cara E. Stepp, Jarrad H. Van Stan, Matías Zañartu, and Daryush D. Mehta. An updated theoretical framework for vocal hyperfunction. *American Journal of Speech-Language Pathology*, 29(4):2254–2260, 2020.

[79] Ingo R. Titze. *The Myoelastic Aerodynamic Theory of Phonation*. National Center for Voice and Speech, 2006.

[80] Ingo R. Titze. Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *The Journal of the Acoustical Society of America*, 111(1):367–376, 2002.

[81] Karen Ann Kochis-Jennings, Eileen M Finnegan, Henry T Hoffman, Sanyukta Jaiswal, and Darcey Hull. Cricothyroid muscle and thyroarytenoid muscle dominance in vocal register control: preliminary results. *Journal of Voice*, 28(5):652–e21, 2014.

[82] Kristin K Baker, Lorraine Olson Ramig, Shimon Sapir, Erich S Luschei, and Marshall E Smith. Control of vocal loudness in young and old adults. *Journal of Speech, Language, and Hearing Research*, 44(2), 2001.

[83] Minoru Hirano and John Ohala. Use of Hooked-Wire Electrodes for Electromyography of the Intrinsic Laryngeal Muscles. *Journal of Speech and Hearing Research*, 12(2):362–373, June 1969. Publisher: American Speech-Language-Hearing Association.

[84] Thomas Rossing. *Springer handbook of acoustics*. Springer Science & Business Media, 2007.

[85] Dik J Hermes. Measurement of pitch by subharmonic summation. *The journal of the acoustical society of America*, 83(1):257–264, 1988.

[86] Hendrik Adriaan Dewald, Platon Lukyanenko, Joris M Lambrecht, James Robert Anderson, Dustin J Tyler, Robert F Kirsch, and Matthew R Williams. Stable, three degree-of-freedom myoelectric prosthetic control via chronic bipolar intramuscular electrodes: a case study. *Journal of neuro-engineering and rehabilitation*, 16(147):1–13, 2019.

[87] Madeleine M Lowery, Nikolay S Stoykov, and Todd A Kuiken. A simulation study to examine the use of cross-correlation as an estimate of surface emg cross talk. *Journal of Applied Physiology*, 94(4):1324–1334, 2003.

[88] Tyler D Chuang and Stacey M Acker. Comparing functional dynamic normalization methods to maximal voluntary isometric contractions for lower limb emg from walking, cycling and running. *Journal of electromyography and kinesiology*, 44:86–93, 2019.

[89] Loretta M Knutson, Gary L Soderberg, Bryon T Ballantyne, and William R Clarke. A study of various normalization procedures for within day electromyographic data. *Journal of electromyography and kinesiology*, 4(1):47–59, 1994.

[90] Anil Palaparthi, Simeon Smith, and Ingo R Titze. Mapping thyroarytenoid

and cricothyroid activations to postural and acoustic features in a fiber-gel model of the vocal folds. *Applied Sciences*, 9(21):4671, 2019.

[91] Weili Jiang, Biao Geng, Xudong Zheng, and Qian Xue. A computational study of the influence of thyroarytenoid and cricothyroid muscle interaction on vocal fold dynamics in an mri-based human laryngeal model. *Biomechanics and Modeling in Mechanobiology*, 23(5):1801–1813, 2024.

[92] Mohammadreza Movahhedi, Biao Geng, Qian Xue, and Xudong Zheng. Effects of cricothyroid and thyroarytenoid interaction on voice control: Muscle activity, vocal fold biomechanics, flow, and acoustics. *The Journal of the Acoustical Society of America*, 150(1):29–42, 2021.

[93] Francesco Negro, Silvia Muceli, Anna Margherita Castronovo, Ales Holobar, and Dario Farina. Multi-channel intramuscular and surface emg decomposition by convolutive blind source separation. *Journal of neural engineering*, 13(2):026027, 2016.

[94] Chenyun Dai, Henry Shin, Bradley Davis, and Xiaogang Hu. Origins of common neural inputs to different compartments of the extensor digitorum communis muscle. *Scientific Reports*, 7(1):13960, 2017.

[95] Melissa R Mazzo, Logan E Weinman, Valerio Giustino, Bailey Mclagan, John Maldonado, and Roger M Enoka. Changes in neural drive to calf mus-

cles during steady submaximal contractions after repeated static stretches. *The Journal of Physiology*, 599(18):4321–4336, 2021.

[96] Daniele Borzelli, TMM Vieira, A Botter, M Gazzoni, F Lacquaniti, and A d'Avella. Independent synaptic inputs to motor neurons driving antagonist muscles. *bioRxiv*, pages 2022–08, 2022.

[97] Miaojuan Xia, Chen Chen, Yang Xu, Yang Li, Xinjun Sheng, and Han Ding. Extracting individual muscle drive and activity from high-density surface electromyography signals based on the center of gravity of motor unit. *IEEE Transactions on Biomedical Engineering*, 70(10):2852–2862, 2023.

[98] JR Florestal, PA Mathieu, and KC McGill. Automatic decomposition of multichannel intramuscular emg signals. *Journal of Electromyography and Kinesiology*, 19(1):1–9, 2009.

[99] Xiaogang Hu, William Z Rymer, and Nina L Suresh. Reliability of spike triggered averaging of the surface electromyogram for motor unit action potential estimation. *Muscle & nerve*, 48(4):557–570, 2013.

[100] Aleş Holobar, Marco Alessandro Minetto, Alberto Botter, Francesco Negro, and Dario Farina. Experimental analysis of accuracy in the identification of motor unit spike trains from high-density surface emg. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):221–229, 2010.

[101] Gijs van Elswijk, Bert U Kleine, Sebastiaan Overeem, Bertine Eshuis, Karin D Hekkert, and Dick F Stegeman. Muscle imaging: Mapping responses to transcranial magnetic stimulation with high-density surface electromyography. *Cortex*, 44(5):609–616, 2008.

[102] Jonathan Lundsberg, Anders Björkman, Nebojsa Malesevic, and Christian Antfolk. Inferring position of motor units from high-density surface emg. *Scientific Reports*, 14(1):3858, 2024.

[103] Dario Farina and Francesco Negro. Common synaptic input to motor neurons, motor unit synchronization, and force control. *Exercise and sport sciences reviews*, 43(1):23–33, 2015.

[104] Dario Farina, Francesco Negro, and Jakob Lund Dideriksen. The effective neural drive to muscles is the common synaptic input to motor neurons. *The Journal of physiology*, 592(16):3427–3441, 2014.

[105] John G Semmler. Motor unit synchronization and neuromuscular performance. *Exercise and sport sciences reviews*, 30(1):8–14, 2002.

[106] Maoqi Chen, Xu Zhang, and Ping Zhou. Automatic multichannel intramuscular electromyogram decomposition: progressive fastica peel-off and performance validation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(1):76–84, 2018.

[107] Carlo J De Luca, Alexander Adam, Robert Wotiz, L Donald Gilmore, and S Hamid Nawab. Decomposition of surface emg signals. *Journal of neurophysiology*, 96(3):1646–1657, 2006.

[108] Víctor M. Espinoza, Matías Zañartu, Jarrad H. Van Stan, Daryush D. Mehta, and Robert E. Hillman. Glottal Aerodynamic Measures in Women With Phonotraumatic and Nonphonotraumatic Vocal Hyperfunction. *Journal of Speech, Language, and Hearing Research*, 60(8):2159–2169, August 2017. Publisher: American Speech-Language-Hearing Association.

[109] Esmaeil Seraj and Karthiga Mahalingam. Essential motor cortex signal processing: an erp and functional connectivity matlab toolbox–user guide version 2.0. *arXiv preprint arXiv:1907.02862*, 2019.

[110] R Merletti and GL Cerone. Tutorial. surface emg detection, conditioning and pre-processing: Best practices. *Journal of Electromyography and Kinesiology*, 54:102440, 2020.

[111] Dario Farina, Roberto Merletti, and Roger M Enoka. The extraction of neural strategies from the surface emg. *Journal of applied physiology*, 96(4):1486–1495, 2004.

[112] Kevin C McGill. Optimal resolution of superimposed action potentials. *IEEE Transactions on Biomedical Engineering*, 49(7):640–650, 2002.

[113] Kevin C McGill, Zoia C Lateva, and Hamid R Marateb. Emglab: an interactive emg decomposition program. *Journal of neuroscience methods*, 149(2):121–133, 2005.

[114] Roberto Merletti, Aleš Holobar, and Dario Farina. Analysis of motor units with high-density surface electromyography. *Journal of electromyography and kinesiology*, 18(6):879–890, 2008.

[115] François Hug, Alessandro Del Vecchio, Simon Avrillon, Dario Farina, and Kylie Tucker. Muscles from the same muscle group do not necessarily share common drive: evidence from the human triceps surae. *Journal of applied physiology*, 130(2):342–354, 2021.

[116] Maoqi Chen, Xu Zhang, Zhiyuan Lu, Xiaoyan Li, and Ping Zhou. Two-source validation of progressive fastica peel-off for automatic surface emg decomposition in human first dorsal interosseous muscle. *International journal of neural systems*, 28(09):1850019, 2018.

[117] Dario Farina, Aleš Holobar, Roberto Merletti, and Roger M Enoka. Decoding the neural drive to muscles from the surface electromyogram. *Clinical neurophysiology*, 121(10):1616–1623, 2010.

[118] Jakob L Dideriksen, Francesco Negro, Deborah Falla, Signe R Kristensen, Natalie Mrachacz-Kersting, and Dario Farina. Coherence of the surface

emg and common synaptic input to motor neurons. *Frontiers in Human Neuroscience*, 12:207, 2018.

[119] Hélio V Cabral, J Greig Inglis, Alessandro Cudicio, Marta Cogliati, Claudio Orizio, Utku S Yavuz, and Francesco Negro. Muscle contractile properties directly influence shared synaptic inputs to spinal motor neurons. *The Journal of Physiology*, 602(12):2855–2872, 2024.

[120] Howard B Demuth, Mark H Beale, Orlando De Jess, and Martin T Hagan. Neural network design. 2014.

[121] Stuart W Perry. Handbook of neural network signal processing. 2002.

[122] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[123] Emiro Jose Ibarra Sulbaran. *Integrating Machine Learning and Physiological Modeling Tools for the Assessment of Vocal Function Using Neck Surface Acceleration*. Phd dissertation, Universidad Técnica Federico Santa María, Valparaíso, Chile, May 2024.